# IMPERIAL COLLEGE LONDON

## DOCTORAL THESIS

---

**Investigating the systems complexity of cellular information processing at and across different levels**

**Biochemical modification systems, cellular networks, and populations**

---


Author

*Vaidhiswaran Ramesh*

Supervisor

*Dr. J. Krishnan*


*A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy*

*in the*


Sargent Centre for Process Systems Engineering

Department of Chemical Engineering


November 24, 2022

# Abstract

Cells continuously process information, simultaneously reacting to environmental cues, interacting with other cells and maintaining the numerous intra-cellular processes required to function and sustain life. The cell undertakes this task by leveraging internal reaction networks whose defining feature is their multi-level and multi-functional nature. These factors present formidable systems challenges. We aim to address these systems challenges associated with information processing at and across different levels by focussing on the following key studies, 1. Exploring intrinsic information processing capabilities of substrate modification systems, 2. Dissecting multi-level interactions between network regulation and substrate modification systems, 3. Understanding the collective behavior of cells (with their internal networks) and populations.

We first focus on the complexities within a given level. Substrate modification provides functionality to a protein and such modifications ubiquitous within a cell. We studied their intrinsic information processing capacity by, 1. Using the new vista of symmetry in reaction kinetics to explore symmetry breaking and its associated features, and 2. Building an unified approach to study their propensity to enable biphasic dose responses. Such modification systems however are seldom modular and often interact with the wider cellular network. Building on work by Dr. *Thapanar Suwanmajo*, we then study the behavior of these networks across levels, focusing on their interplay with canonical network regulation such as feedback and feedforward. Finally, beyond the cell, we study the role of multi-level cross-population interactions in enabling the collective behavior of cells and populations at different levels.

Throughout we utilize a consolidated methodology (computational, analytical and semi-analytical approaches), that allows us to tackle these systems challenges and characterize the effect of parameters. We provide strict characterization (as a function of parameters) of multiple emergent behaviors, trace transitions and origins of behavior, and provide guarantees for presence/absence of behavior. Our analysis provides multiple evidences of novel behavior arising from the multi-level and systems interplay. At the substrate modification level, we show symmetry breaking leading to multistability and absolute concentration robustness of different kinds, and our unified approach to biphasic doses, reveals how various systems can present enzyme or substrate biphasic or a combination of both. Across levels we show substrate modification together with network regulation can present novel behavior (multistability from negative feedback) and the interplay can subvert expected response from each level. At the level of cells and populations, we show collective behavior arising from multi-level and multi-population interactions, that cannot be attributed or realized from one level or one population. Together, these studies represent a first step in tackling the complexity of multi-level cellular information processing.

# Statement of Originality

I, Vaidhiswaran Ramesh, declare that this thesis titled "*Investigating the systems complexity of cellular information processing at and across different levels: Biochemical modification systems, cellular networks, and populations*" and the work contained within are my own.

I confirm that:

- This work was done wholly while in candidature for a research degree at Imperial College London.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at Imperial College London or any other institution, this has been clearly stated.
- Where I have consulted or quoted the published work of others, this is always clearly attributed. With the exception of such quotations, this thesis is entirely my own work.
- Where the thesis is based on work done by myself in collaboration with others, I have made clear exactly what was done by others and what I have contributed myself.

In particular, Chapters 1 to 4 and 6 are entirely independent contributions. Work discussed in Chapter 5 was built on contributions from ex-research group member Dr. *Thapanar Suwanmajo* [1]. His particular contributions are highlighted when included in the chapter (to help setup the narrative context for the overall study). Unless explicitly stated in such manner, all work presented was done entirely by myself.

Vaidhiswaran Ramesh

London, November 2022

# Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

# Acknowledgement

As in any thesis, these pages reflect without any redaction or curation, the honest thoughts and acknowledgements of the candidate to the elements of support in this long process. While it's often treated as a perfunctory exercise (with hints of soft diplomacy), I would like to keep this as an honest transcript thanking in measure the people who have enabled me to be where I am today - not just by virtue of their position and thus the obligatory influence in my life... but by venturing that extra mile for others.

But before, a word to the ethereal constants in my life. *My mother, family, and Madras*, first and foremost deserves acknowledgement in this space - But to tie and measure their influence to a few meagre words within this manuscript is a mistruth that I will not attempt; for it is to them that I owe it all. All until now and all after.

Then, there are some folks who are separated by decades and centuries but whose works have buoyed me nonetheless during these emotive few years and I'd like to acknowledge them just as any other. *T.M. Krishna, Ilayaraja, Gustave Flaubert, Amor Towles, Penelope Lively, Salman Rushdie, Francois Sagan, Gabriel Garcia Marquez, Ferdinand Celine, Isabel Allende, Marcel Proust, Richard Linklater, Ranjini & Gayathri, Schubert, The Schumanns, Vivaldi* and finally my old Baroque pal *J.S. Bach*.

The fabled halls in *London* that hold the Old Masters works, and the Victorian bookshops dotted along alleyways and basements; with your crumbling parchments and yellowed out wallpapers... thanks for greeting me within those the secret corners that I alone knew for a time. To *Raphael, Titian, Rubens, Turner, Rembrandt, Wallace, Tate Britain, Quinto and Apsley*... and *Napoleon*!

Speaking of whom, a brief mention to *Paris*, which has indeed been a second spiritual abode towards the end, realized in truth with those few intense (and expensive) visits across the channel; To the pied-a-terre in *Saint Germain des Pres*; To the innumerable rues that *Haussmann* lifted out of memory to pander to the angst and yearnings of 20 year old me more than a century later! To the *Oath of the Horatti* and to *Palais du Luxembourg*! To *Messoinoir* and to the memory of that night!

There are then those that are yet to grace me with their presence but nonetheless deserve mention here; To *St. Petersburg* and those cold cold nights; To *Moscow, Tretyakov*

*Gallery* and the *Bolshoi*! To bees from *Nizhny Novgorod* bearing the sweet scent of its unique honey - and finally to the thistle that enamoured *Tolstoy* and the rest of us!

As pontificious as it may seem to accord such words to works and places as far removed from my own thesis and life as something can be, I must confess that the impetus derived from feeling humbled in the presence of something larger than these pages and even us is as worthy of praise, gratitude, and prose as any living person or being! For these halls, works, and places reflect the greatest extent of mans' imagination and his zeal to will elements into being. If not for them - I believe this work - and all aspects of my life might have withered and crumbled long ago.

Anyway, now that we have dispensed acknowledgement to the ethereal - on to the material world. I have had the good fortune in my life to cross paths with better souls than myself and I consider myself blessed even to call some of them my best friends. Few and measured, these people are equally paramount to me being the person I am today! Again in no order of preference or partiality - *Christopher James Wheeler, Alessia Negro, Charles de Grange, Leonore Evans, Manoah Ashirvadham, Sai Darshan, Govind Menon* and... *Zofia Piszka*! Thank you for the memorable years and for making *London* & *Madras* all the more wishful. I am also indebted beyond words to the entire *Tournant* family for making me feel at home in grey *London* and in *Kensington*.

Any thesis reaching this stage is the fruit of a combined effort of the PI and the candidate. In this accord and for much much more - I am thankful to my supervisor *J. Krishnan* for his support and for encouraging a holistic intellectual education over the years. From the numerous musings over cricket to a shared love for *Russian* literature and classical music, it has indeed been a pleasure to 'work' with him in these few years.

Now, before signing off, I'd like to use the words of souls better endowed and blessed than me to make the following comment - One that better summarises the journey and the immense gratitude I owe for all along the way!

> In our journey to the end of the night in search of lost time... we envy, miss and regret a lot of things. But we never envy someone as we do those that knew rapturous love and passionate enslavement... Never miss someone as we do those to whom we have shared... And never regret that which we have the ignorance to forget! And we blame it all on god. After all not all of us are fortunate in our circumstance to believe in chance rather than god!
>
> - Having said that -
>
> The Rubicon is crossed with this writ and our youth has gone to the ends of the earth to die in the absence of the truth; We venture out then - Sei Soli - before memory of that summer betrays us - like boats beating against the current, borne back ceaselessly into the past... just as those races condemned to 100 years of solitude who didn't get a second chance or a future on this earth!

# List of Papers

## List of papers from this thesis

1. **Ramesh V**, Krishnan J. Symmetry breaking meets multisite modification. Elife. 2021 May 21;10:e65358. DOI: 10.7554/eLife.65358. PMID: 34018920; PMCID: PMC8439660.
2. **Ramesh V**, Suwanmajo T, Krishnan J, Network regulation meets substrate modification chemistry *(Under revision at Journal of Royal Society Interface)*
3. **Ramesh V**, Krishnan J, Biphasic cellular signaling: from enzymatic modifications to signalling networks (To be submitted - end of November, 2022)

## Related work

1. Suwanmajo, T., **Ramesh, V**. & Krishnan, J. Exploring cyclic networks of multisite modification reveals origins of information processing characteristics. Sci Rep 10, 16542 (2020). DOI: 10.1038/s41598-020-73045-9

# Contents

**Abstract** **3**

**Statement of Originality** **5**

**Copyright** **7**

**Acknowledgement** **9**

**Nomenclature** **20**

**1 Introduction** **23**

**2 Symmetry and symmetry breaking in multisite modification** **31**
2.1 Introduction . . . . . 31
2.2 Methodology . . . . . 34
2.3 Results . . . . . 35
2.4 Conclusions and Discussions . . . . . 48
2.5 Appendix . . . . . 57
Executive summary of analytical results and resources . . . . . 66

**3 Biphasic responses in cell signalling: A unified approach** **67**
3.1 Introduction . . . . . 67
3.2 Models and Methods . . . . . 69
3.3 Results . . . . . 71
3.4 Discussion . . . . . 81
3.5 Appendix . . . . . 87
Executive summary of analytical results and resources . . . . . 91

**4 Network regulation meets substrate modification chemistry** **95**
4.1 Introduction . . . . . 95
4.2 Models and Methods . . . . . 99
4.3 Results . . . . . 101
4.3.1 Computational results. . . . . 101
4.3.2 Analytical results. . . . . 111

4.4 Discussion . . . 114
4.5 Appendix . . . 122
Executive summary of analytical results and resources . . . 126

**5 Multi-level complexity of cells and populations 131**
5.1 Introduction . . . 131
5.2 Methodology . . . 133
5.2.1 Computational methodology . . . 136
5.3 Single population study . . . 138
5.3.1 Mathematical model . . . 138
5.3.2 Scenarios . . . 142
5.3.3 Questions . . . 144
5.3.4 Results and Discussion . . . 145
5.4 Two populations study . . . 158
5.4.1 Mathematical model . . . 158
5.4.2 Scenarios . . . 162
5.4.3 Questions . . . 166
5.4.4 Results and Discussion . . . 167
5.5 Emergent multi-population behavior . . . 177
5.6 Conclusions and future work . . . 192

**6 Conclusions 195**

**A Parameters 207**
A.1 Symmetry and symmetry breaking in multisite modification . . . 207
A.2 Biphasic responses in cell signalling: A unified approach . . . 210
A.3 Network regulation meets substrate modification chemistry . . . 213
A.4 Multi-level complexity of cells and populations . . . 217

**B Mathematical models and analytical results 219**
B.1 Symmetry and symmetry breaking in multisite modification . . . 219
B.2 Multi-level complexity of cells and populations . . . 246

**Bibiliography 248**

**C Extended Appendix 269**
Symmetry and symmetry breaking in multisite modification . . . 269
Biphasic responses in cell signalling: A unified approach . . . 397
Network regulation meets substrate modification chemistry . . . 525

# List of Figures


2.1 Schematic representation of the various multisite phosphorylation networks considered for symmetry breaking and the three different classes of symmetries they accommodate 36
2.2 Case 1 and case 2 symmetry breaking in various double site phosphorylation networks . 43
2.3 Case 3 symmetry breaking in various double site phosphorylation networks . . . . . 46
2.4 Schematic representation of realization of symmetry and symmetry breaking in multisite modification networks through the interplay of basic biochemistry of post translational protein modification and network symmetry, and Summary . . . . . 49
2.5 Schematic representation of the mixed random ordered DSP network . . . . . 57
2.6 Case 3 symmetry breaking in random ordered DSP with common kinase and common phosphatase acting on each modification site . . . . . 58
2.7 Case 1 symmetry breaking in the random ordered DSP with common kinase and common phosphatase acting on each modification site, even when one of the legs is incapable of breaking symmetry by itself . . . . . 59
2.8 Case 3 symmetry breaking through a supercritical pitchfork bifurcation in the random ordered DSP with separate kinase and separate phosphatase acting on each modification site . . . . . 60
2.9 Features of symmetry and symmetry breaking in systems near symmetric . . . . . 61
2.10 Approximate concentration robustness shown by systems that deviate from exact symmetry . . . . . 62
2.11 Total enzyme concentrations are an additional lever (apart from basic network kinetics) that can independently tune symmetry breaking behavior in MSP networks . . . . . 63
2.12 Case 1 symmetry breaking in distributive ordered TSP network . . . . . 63
2.13 Absolute concentration robustness in ordered DSP . . . . . 64
2.14 Detailed model description of the various MSP networks used in this paper . . . . . 65

3.1 Enzyme mediated substrate modification: Snapshot - from modification mechanism to common substrate modification systems and their role in enabling network regulation . 68
3.2 Presence and absence of substrate and enzyme biphasic dose responses in the commonly observed enzymatic cellular signaling systems. . . . . 75
3.3 Role of biphasic interactions in network and motif response . . . . . 78
3.4 Presence and consequence of biphasic dose responses in cellular signaling arising from enzyme mediated substrate modification . . . . . 85
3.5 Coexistence of multi-stability and biphasic dose response in various enzymatic models for the same underlying kinetic regime . . . . . 87

3.6 Biphasic responses in random double site modification network with separate enzymes effecting each modification . . . . . 88
3.7 Effect of biphasic response in interactions within feedback network motifs (open systems) 88
3.8 Effect of biphasic response in interactions within integral feedback control motif: Perturbation of homeostatic response . . . . . 89
3.9 **Schematic representation of the various models considered in the manuscript.** Models M0-M1 represent the basic unit of (de)modification, the covalent modification cycle and its simple extension with the substrate forming an additional complex with the enzyme. Models M2-M5 represent different systems of ordered distributive double site modification of a substrate (by different or common enzymes acting on each modification site). Models M6-M7 represent other architectures of modification systems in chemical reaction networks, such as cascaded system (with different and common enzymes affecting (de)modification) and coupled covalent modification networks (coupled through enzyme sharing in (de)modification). Models N0-N5, represent different networks and network motifs with biphasic dose-response subsumed in an interaction. . . . . 90

4.1 Schematic description of signaling networks involving two levels, the network regulation level and the substrate modification level . . . . . 98
4.2 Impact of positive feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system . . . . . 102
4.3 Impact of negative feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system . . . . . 104
4.4 Impact of feedback mediated by the partially modified substrate $A_p$ (through activation or inactivation of kinase) on the behavior of the double site modification system . . . . 106
4.5 Impact of incoherent feedforward network regulation on the double site modification system . . . . . 108
4.6 Impact of coherent feedforward network regulation on the double site modification system 110
4.7 Conceptual summary of the multi-level interaction between network regulation and substrate modification levels studied, the methodology used, and a synthesis of results and insights . . . . . 117
4.7 Negative feedback from $A_{pp}$ can enable a combination of multistability and oscillations 122
4.8 Incoherent feedforward regulation can generate multiple homeostasis capabilities in substrate through both simple network regulation and by exploiting and tuning intrinsic substrate modification capabilities such as symmetry breaking . . . . . 123

5.1 Schematic representation of multi-level interaction and emergent collective behavior at the level of cells and populations of cells . . . . . 134
5.2 Schematic representation of the computational pipeline used to explore the multiplicity of scenarios of interactions and cellular circuit combinations in a robust scalable manner 137
5.3 Expected circuit response from the intra-cellular circuit motifs . . . . . 141

5.4 Single population study: Population strength, environmental factors and intra & extra-cellular circuit response in the absence of any interactions between the cell, population and the environment . . . 146
5.5 Simple interactions between population level and cellular level allows for information processing behavior to transcend levels . . . 147
5.6 Interactions between the population and environmental factors can allow for the levels to modulate each other . . . 149
5.7 Interactions between population and enviornment factors can also modulate extra-cellular circuit response indirectly (fixed population growth rate $kg_n$ - changing circuit strength $S_{basal}$) . . . 150
5.8 Interactions between population strength and enviornment factors can also modulate extra-cellular circuit response indirectly (fixed circuit strength $S_{basal}$ - changing population growth rate $kg_n$) . . . 151
5.9 Intra-cellular circuit can modulate and lend behaviors to environmental factors . . . 152
5.10 Population level can dampen dynamic circuit response and systems behavior . . . 154
5.11 Active and direct interactions with the circuit level allow for pronounced information transfer and communication for the population and environmental factors . . . 155
5.12 Distance to circuit level dictates strength of information transfer from circuit level to other levels . . . 156
5.13 Role of diffusion coefficient and cell volume on strength of interactions between circuit and other levels . . . 158
5.14 Two populations study: Population strengths and extra-cellular circuit response in the absence of any interactions between the cell, population and the environment . . . 168
5.15 Role of purely population level interactions in modulating cellular response . . . 169
5.16 Emergent collective information processing capabilities when cellular circuits communicate across populations (unidirectional) . . . 170
5.17 Emergent collective information processing capabilities when cellular circuits communicate across populations (bidirectional) . . . 171
5.18 Effect of cellular level interacting at the population level (cross-population & unidirectional from one population) . . . 172
5.19 Effect of cellular level interacting at the population level (cross-population & unidirectional from both populations) . . . 173
5.20 Effect of population affecting circuit response across populations . . . 174
5.21 Effect of bidirectional interactions between cellular level and population level . . . 175
5.22 Cross-level interactions allow for simultaneous 'co-existence' of different steady states at similar concentrations at the population level, which otherwise reflect drastically different responses at the cellular level . . . 176
5.23 Multi-level interactions generate novel qualitative behavior (tristability) at the systems level, which may not be reflected equally in all species and levels . . . 177
5.24 Population dynamics modulating intra & extra-cellular circuit response through dilution 184
5.25 Population dynamics directly modulates circuit activation (cross-population) . . . 186
5.26 Population dynamics can enable reversible switching in circuit motif response (MI motif) 188

5.27 Population dynamics and cross species circuit interactions in predator prey model . . . 190
5.28 Population dynamics can seemingly introduce inaccessible (for given signal activation) behaviors to the circuit motif . . . . . 191

# List of Tables


2.1 Executive Summary - Chapter 2: Symmetry and symmetry breaking in multisite modification . . . . . 66

3.1 Executive Summary - Chapter 3: Biphasic responses in cell signalling: A unified approach 91

4.1 Executive Summary - Chapter 4: Network regulation meets substrate modification chemistry . . . . . 126

5.1 Examples of communication and cellular information processing at the cellular level . . 133

5.2 Scenarios used to analyze role of interactions in enabling population level response . . 142

5.3 Scenarios used to analyze role of interactions in enabling population level response (two populations study) . . . . . 163

5.4 Circuit choice for two population study with complex emergent population level dynamics 181

5.5 Circuit choice for two population study with complex population level dynamics . . . 181

# Nomenclature

| | |
|---|---|
| $\alpha$ | Environmental Factors |
| *ACR* | Absolute Concentration Robustness |
| *Circ.* | Intra-cellular circuit level |
| *Coh. FF* | Coherent Feedforward |
| *cR* | Circuit Motif Response |
| *DSP* | Double site phosphorylation |
| *Env.* | Environmental level |
| *ERK* | Extra-cellular Signal Regulated Kinase |
| *e* | Intra or Extra-cellular Circuit Response |
| *FB* | Feedback |
| *Inc. FF* | Incoherent Feedforward |
| *IPMK* | Inositol Polyphosphate MultiKinase |
| *MAPK* | Mitogen Activated Protein Kinase |
| *MDM2* | Mouse double minute 2 homolog |
| *MI* | Mutual Inhibition |
| *MWC* | Monod Wyman Changeux |
| *NFB* | Negative Feedback |
| *n* | Population Strength |
| *ODE* | Ordinary Differential Equation |
| *PFB* | Positive Feedback |
| *PGI* | Phospho Glucose Isomerase |
| *Pop.* | Population level |
| *PPI* | Protein Protein Interaction |
| *PTM* | Post Translational Modification |
| *Rlx* | Relaxation Oscillator |
| *Sc.* | Scenario |
| *SIC1* | Stoichiometric Inhibitor of Cdk1-Clb complexes |

# Chapter 1

# Introduction

Cells continuously process information. At any given instant, a cell is reacting to environmental cues (physical and chemical), interacting with other cells (of the same type or different types), while also maintaining the numerous intra-cellular processes required to function and sustain life [2]. The cell is able to undertake this monumental task by leveraging a number of internal chemical reaction networks and circuits, including gene regulatory, protein and molecular networks. However the pathways that enable distinct cellular processes are seldom modular (if ever), and are all interconnected. In many instances the circuits that enable a certain process contain components that play a key role in the functioning of a circuit enabling another distinct process [3]. This interconnectedness of the circuits and the multi-functionality of the components contained there in, while posing a systems challenge to understanding its functioning, is precisely what allows for robust control and effective information processing within the cell. The strength of interactions within these circuits also presents a significant parametric challenge (from a modelling perspective) further underscoring the complexity of cellular information processing [4]. Thus, understanding the role of any pathway or process is essentially a substantial systems challenge.

A key component in many intra-cellular information processing pathways are proteins. Following biosynthesis within the cell, proteins take on multiple roles and are responsible for everything from gene regulation to regulating key pathways and processes, including maintaining cellular shape, cell division, cell metabolism and cell cycle phases [5]. The protein *cytochrome C* for example, simultaneously participates in mitochondrial energy production [6], and apoptotic cell death [7, 8, 9]. Proteins are even capable of exhibiting different functions within one poly peptide chain (known as *protein moonlighting*). For example, the protein phosphoglucose isomerase (*PGI*) in bacteria, functions both as a cytoplasmic enzyme in glycolysis and as an extracellular cytokine and growth factor [10, 11, 12]. Such moonlighting and the multiplicity of functionality in proteins is increasingly seen as the norm rather than an exception [13].

Proteins gain and lose such function in a number of ways including ligand and scaffold binding, and with environmental cues such as changes in pH and temperature [2]. One of the key ways in which proteins gain or lose functionality is through the process of post-translational modification [14]. Protein modification (such as phosphorylation, acetylation, ubiquitination, etc) provide additional functionality to the protein, either by covalent addition of a functional group or proteolytic cleavage of regulatory sub units. This gain in functionality allows a protein to take on multiple roles and participate in different pathways [15, 16, 17, 18, 14]. Each modification in this instance is effected by an enzyme (which itself is a protein that can be modified). Given the ubiquity of proteins in cells, the process of post-translational modification is itself a recurring and pervasive process within cells, with over 430 different post-translational modifications having been identified so far [19, 20].

Proteins can also get modified multiple times, sometimes by the same enzyme and sometimes by different enzymes [2]. These series of modifications act akin to a *molecular barcode* for the protein, allowing it to fulfil several functions and participate in different pathways [21]. These modifications can have a sequence or order to them (mandated by other cellular processes or for structural reasons), and in certain other instances can be (and are) effected randomly. For example, the *p53* protein is known to be modified by as many as 50 individual post-translational modifications (by same and different enzymes) [22]. The *Tau* protein, implicated heavily in *Alzheimer's disease*, is another example of a protein that undergoes multiple modifications, with up to 35 percent of the 441 amino acid residues capable of individual modifications [23, 24, 25]. There are several complexities inherent to the protein modification mechanism. The combinations of enzyme action possible (in the case of multisite modification), the coupling between modifications arising therefrom, and the sequestration of enzymes in multiple complexes are just a few examples. Further, the mechanism of multisite modification could also be different. The modifications could be effected by an enzyme in a distributive manner (enzyme dissociates after each modification) or in a processive manner (enzyme is attached to the protein until all modifications are complete) or a mixed manner [26].

All of this while contributing to the extensive non-linearity in the systems behavior (adding to the complexity), also allows for a richness in the information processing capability. Simple systems of protein modification involving as little as one or two modifications have been shown to be capable of complex information processing behavior including switch like response [27], ultra-sensitivity [28], multi-stability [29, 27], oscillations [30, 31] etc. These behaviors and the modality of their occurrence are of great biological relevance. Thus in essence, the protein modification mechanism, in addition to playing a key role in allowing a protein to participate in and regulate cellular processes, is itself a key cellular process capable of complex information processing.

Another distinct aspect of cellular information processing is the recurrence of network structures (within larger networks), called network motifs, that enable a certain sig-

naling response [32, 33, 34, 35]. Cellular information processing can be studied in a number of different ways. Improved experimental (analytical) tools have meant that an increasing number of data rich techniques and high throughput methods are being used to study the relationships between different species within a cell and to discern its information processing capabilities [36, 37].

For the study of many of these interactions and pathways within the cell, the common perspective is that of networks. The focus on networks, building on basic biochemical relationships between different agents, has allowed for us to understand, predict, and intervene in decision making processes of the cell [38]. Virtually every important process is studied through the underlying network of interactions. Many canonical pathways and networks (ranging from gene regulation to protein protein interaction (*PPI*) to metabolic networks) have been identified. The *RAS* and *Calcium signaling* pathways are a couple of such examples [39, 40]. Studies of these networks has revealed how they can be telling markers of pathologies. For example, misregulation of the *Tau* proteins and its network/pathways is known to be associated with *Alzheimer's disease* and it's pathologies. Networks however are nonetheless complex at every level [41]. A typical cellular information pathway/network is made of up to 100s of species interacting with each other. These networks are also mired in additional complexity such as abundant redundancy, stochasticity of reactions [42], cross-talk [43], and spatial and localization effects, to name a few [44].

Notwithstanding the messiness of networks and their variety, the study of pathways and the information processing in cells through network motifs remains a fruitful vantage point [32, 45, 46, 47, 48]. Feedback and feedforward networks represent some of the most basic network motifs, and these are seen quite ubiquitously across cellular pathways [49, 50, 51]. Other recurring network motif structures include homeostatic motifs, bistable switches, oscillators, ultrasensitive and adaptive modules [52, 38, 53] The study of these motifs, and the biochemical species participating in them, has been immensely helpful in discerning the behavior of pathways and in understanding cellular regulation. These patterns of interactions are pervasive in biology and are seen across different scales and levels. A node in these networks can range from a single protein, to a consortia of proteins & protein networks [33], to cells and even populations [54, 55]. Such network motifs have been successfully leveraged by synthetic biologists to engineer behavior, both at the intra-cellular and at the level of populations of cells [56, 57].

Encompassing and enabled by this massive arsenal of intra-cellular pathways and networks, cells themselves communicate with other cells and the environment to enable a collective behavior. There is an increased interest from the synthetic biology community to leverage insights from network behavior and cell-cell communication to engineer population level response [58, 59, 60, 61, 62].

The underlying theme across this landscape of cellular signalling, is that signaling and

many biological processes involve multiple levels. Proteins participate in signaling pathways involving gene regulation, protein translation and biosynthesis; signaling pathways talk to each other within cells and across cells and elements in the environment; and cells talk to each other and enable population level response. Cross-level or multi-level interactions is thus a staple of cellular information processing, and the study of these systems cannot be broken down in a hierarchical way. Many interesting behaviors can trace their origins to the systems behavior coming from this multi-level interaction.

In this thesis, we aim to investigate the systems level complexity of cellular information processing at and across these different levels. In particular, we focus on three key areas,

1. Understanding the *inherent signaling capabilities of biochemical modification systems* (post-translational modification)
2. Exploring the systems behaviors that emerge from *multi-level interaction between biochemical modification systems and network regulation*
3. Assessing the multi-level complexity of *cells and populations*

**Understanding the inherent signaling capabilities of biochemical modification systems**

As mentioned above, post-translational modifications are an ubiquitous part of cellular signaling, and are integral to the proper regulation and functioning of virtually every cellular pathway. However given the complexity of the system and non-linearity in the network, there are many important unanswered questions regarding their information processing capabilities, and further, how certain features of these modification networks came to be. For example, a fundamental feature of a multisite modification network is whether it presents with any ordering to the sequence of modifications (for non-structural reasons). As mentioned earlier multisite modifications can be random, ordered, or partially ordered. In fact, such ordering is seen in a wide range of proteins that undergo multisite modifications and understanding their characteristics and origins is of interest to both the systems and synthetic biology communities [63, 64, 65, 66].

In our first study (chapter 2), we employ symmetry in chemical reaction kinetics as a new vista for exploring the intrinsic behavior of modification systems, including the propensity for symmetry breaking and the creation of bias in ordering. Such symmetry in this context can represent multisite protein modification with no bias in the sequence of ordering of modifications.

In chapter 3, we explore the propensity of enzymatic modification mechanisms to present biphasic dose responses (non-monotonous bell shaped response to a monotonous input). Such biphasic dose responses are an ubiquitous signaling behavior seen across

all scales of biology (both dynamic responses and steady state dose response biphasic) [67, 68], and are often attributed to the behavior of network motifs (such as incoherent feedforward). In chapter 3 we studied the propensity of basic networks of enzymatic modification to present with biphasic dose response. We employ an unified framework that explores substrate and enzyme biphasic response together, and we show how biphasic interactions as part of network motifs, modify, perturb or enable standard network motif response. In doing so, we look across all levels of modification systems, and the role of the modification mechanism in enabling concrete networks and pathways to present biphasic dose response.

**Exploring multi-level interaction between biochemical modification systems and network regulation**

Modification systems rarely exist in isolation and are often part of networks and network motifs either implicitly (captured within an interaction between nodes of the network) or explicitly (as part of the nodes themselves). These network motifs are already known to give characteristic behaviors, however, there are many unanswered questions with regard to the emergent systems behavior when such network motifs encompass complex information processors such as protein modifications (within their nodes or arrows).

Building on our limited exploration of multi-level interaction in chapters 2 and 3, and on work by ex-research group member Dr. *Thapanar Suwanmajo* in his thesis, in chapter 4 we explicitly study the interplay between modification systems and characteristic network motifs such as positive and negative feedback, and incoherent and coherent feedforward networks. Both the modification system and the network motifs are capable of rich signaling behavior. In this study we explore how one level affects the response (characteristic or expected) from the other, and the emergent systems level response coming from their interaction (that cannot be attributed to one level or the other).

**Assessing the multi-level complexity of cells and populations**

Motivated by recent progress in single cell studies and efforts to build and engineer responses with synthetic multi-cell/population systems, we studied higher order communication between cells and in populations to explore the emergent systems responses at the level of populations of cells interacting with one and other, in chapter 5. Both the population level and the cellular level are actively engaged in cross-level (and sometimes cross-population), including even the environment to enable a collective behavior. Our focus with this study is on how characteristic interactions from each of these levels, between themselves enable such behavior at the systems level.

## Methodology

The above studies are pursued with the help of a three pronged methodology designed to specifically address the questions at hand.

1. **Numerical computation:** simulations and bifurcation analysis
2. **Analytical:** symbolic computation (and *pen & paper*)
3. **Semi-analytical:** bridging the gap between analytical and semi-analytical

Using this three pronged approach, we

- Show the *existence of qualitative behaviors*
- Provide *unambiguous binary answers* regarding the possibility and impossibility of behavior by a model
- Analyze sufficiency and necessary conditions based on *parameters to obtain or rule out behaviors*
- *Characterize features of the system & behaviors* and, their dependence on parameters
- Provide strategies to *predict or rule out behaviors* based on partial functional (parametric) awareness of the system

Each system studied is modelled mathematically as a system of ODEs. These models are constructed using first principles and any assumption made is clearly stated. The mathematical model descriptions are cross-validated through *COPASI* [69], where the model descriptions are auto-generated through specification of the reaction network schematic. The models are numerically simulated in *MATLAB* [70], and bifurcation analysis is carried out with the help of the *MATLAB* package, *MatCont* [71, 72]. Analytical work (computational algebra) is pursued using the software *Maple* [73]. Any computational code written to generate the figures (including any automation pipeline and *MatCont* dependency package) is available on *GitHub*.

All parameter values used to generate the figures are provided at the end of the thesis (see appendix A). The *Maple* files (containing the proofs) are provided along with thesis. The extended appendix contains a PDF version of these *Maple* files for easy accessibility.

## Organization of the thesis

In chapter 2 (Symmetry and symmetry breaking in multisite modification), we study the information processing capacity of protein modification, in particular how symmetry breaking can provide features to and enable behavior in multisite protein modification. We first describe the models of multisite systems considered, along with the classes of symmetries present in these systems. The computational and analytical results are then presented followed by a discussion of their relevance. Additional supplementary figures relevant to the discussion are presented at the end of the chapter.

Detailed mathematical description of the models and a brief description of the analytical proofs mentioned are provided in appendix B.1. Detailed proofs and supporting analytical results are presented in the appendix C (*Maple* PDF).

In chapter 3 (Biphasic responses in cell signalling: A unified approach), we explore the propensity of substrate modification networks to present with steady state biphasic dose responses (to both substrate and enzyme doses). We begin by describing the different network topologies involving protein modification as a basic building block. The computational and analytical results regarding the capacity of these models to enable biphasic dose response are then presented. The effect of such biphasic interactions when they emerge as part of network motifs are then studied before studying the concrete exemplar of *ERK* pathway. Detailed mathematical descriptions of the models used and the analytical proofs are presented in appendix C (*Maple* PDF).

In chapter 4 (Network regulation meets substrate modification chemistry), we study the multi-level interaction present between multisite modification systems and network motifs more explicitly. We first describe the different network motifs and modification systems considered, before analyzing the systems behavior that emerge from their interaction. A dedicated methodology developed to tackle this is presented followed by the relevant results. Mathematical descriptions of the models used and the detailed proofs are presented in appendix C (*Maple* PDF).

In chapter 5 (Multi-level complexity of cells and populations), we study the multi-level interactions inherent in cellular communication at the level of populations of cells (single and two populations), and how they enable the collective response across levels. We initially study complexity inherent to a single population before studying that of two populations. In each case we first describe; the models, the different interactions considered, choice of cellular circuit motifs and the assumptions made. Then we build a framework of scenarios considering different combinations of interactions and circuit motifs. We then describe the computational automation pipeline created to generate bifurcation plots for the scenarios before presenting and discussing the results. Circuit model descriptions are presented in appendix B.2.

Chapters 2 to 4 have an executive summary at the end of the chapters that summarizes the key computational and analytical results. It details references to the location in appendix C (*Maple* PDF) to access the relevant proofs along with references to the computational results.

# Chapter 2

# Symmetry and symmetry breaking in multisite modification

## 2.1 Introduction

Reversible phosphorylation, and post-translational modification of proteins in general constitutes a basic way of conferring functionality to proteins in cells. This basic unit (covalent modification) is built upon in many different ways to result in the complex biochemical pathways encountered in cells.

A particular elaboration of this mechanism which is widely encountered is the reversible multisite modification of proteins by enzymes. Here, a number of basic variations are possible depending on whether the enzymes involved in distinct modification steps are different or if a common enzyme effects multiple modifications. In the latter case there are variations depending on whether the enzymatic mechanism is distributive (enzyme dissociating from substrate after every modification) or processive (enzyme remains attached). Finally there are variations depending on whether a specific ordering of modifications occurs (ordered mechanism) or not (random mechanism).

In addition to being a widely encountered way in which substrates are reversibly modified to confer functionality (and consequently of broad interest), interest in multisite modification stems from the fact that the basic modification mechanisms are capable of acting as complex molecular information processors [31]. Various studies have highlighted the possibilities of these mechanisms exhibiting switching and threshold behaviour [27], bistability/multistability [29, 74], oscillations [75, 26, 76], biphasic dose response curves [77], and other complex behaviour [78]. A range of studies have delineated the ingredients required (from the above possible variations) to enable or pre-

vent such behaviour [79, 80, 81]. We emphasize that this rich repertoire of behaviour emerges from the most basic considerations and aspects of enzymatic modification of substrates, and that this behaviour is a feature and a consequence of the modification network (rather than a single modification). Information processing capabilities are also at the heart of different strands of work in synthetic biology engineering multisite modification, and reaction networks more broadly [65, 82, 83, 84]. This chapter focusses on a distinct aspect of information processing of multisite modification: symmetry and symmetry breaking.

Symmetry and symmetry-breaking are themes encountered across different scales and levels in biology, ranging from the cell population, to the cellular, to the molecular level. A fundamental theme in developmental biology is the breaking of symmetry to generate patterns. The basic questions here centres around how an apparently homogeneous field of cells can differentiate to exhibit a basic pattern which serves as a precursor for subsequent development. Modelling, experiments and concepts from self-organization have been used to probe this generation of form, which breaks spatial symmetry. The underlying mechanisms invoked involve many variations on the classical Turing mechanism, or the interplay of mechanics and chemistry [85, 86]. This can be significantly complicated by the presence of many layers of regulation. Strong experimental evidence for such mechanisms present at the core of developmental regulation have been demonstrated in multiple model systems [87]. Symmetry breaking as a basis of generating form at the cellular level, for instance polarization and polarized or other strongly inhomogeneous patterns of concentration of species, has been explored in a range of contexts. Examples include polarity generation in fungi and plant cells [88], and in neutrophil chemotaxis [89]. Symmetry has also been invoked as a key ingredient in the development of the MWC model which has been used to explain allostery in biomolecular information processing [90].

While the theme of symmetry in chemistry is well recognized especially at the molecular structure level [91], there are relatively few studies of symmetry-breaking at the molecular reaction level. In chemical reaction systems, symmetry is encountered in the context of chirality in racemic mixtures. Racemic mixtures comprise equal amounts of the two enantiomeric forms of a chiral molecule with opposite chiralities, and a central question is how a dominant orientation (chirality) of the molecular mixture can emerge from this. Some studies explain this as an emergent behaviour of the reaction network system governing the two forms of the molecules: even if the network/reaction system is symmetric allowing for equal amounts of the two forms, this symmetry can break giving rise to a dominant form. A recent study [92, 93] evaluated and demonstrated the feasibility of such symmetry breaking in a number of potential reaction systems. Chiral symmetry breaking has been experimentally observed in crystallization of nanoparticles [94], fibril formation from racemic mixtures [95] and in the Soai reaction [96]. Such symmetry breaking is of particular importance in prebiotic evolution and biology, where bio-polymers and bio-molecules are characterized by a specific chirality

and orientation, even though the original non-life chemical world was chirally symmetric [97]. The establishment of such chirality has been postulated to be important in understanding the origins of life [98].

This chapter focusses on a specific aspect of symmetry breaking at the junction of the biological and the chemical: the breaking of symmetry in basic multisite phosphorylation systems.

The motivation for studying symmetry and symmetry-breaking in the context of multisite modification stems from different sources: conceptual insights, relevance to systems biology and potential application in synthetic biology. In this regard we note that: (i) Many of the basic modification networks accommodate different types of symmetries, as we discuss below. (ii) Certain symmetries, for example resulting in equal concentrations of different partial phosphoforms of a given level (the Case 2 Symmetry, discussed below) are not only plausible in vivo, but have also been assumed in multiple contexts sometimes implicitly. (iii) An asymmetric state currently observed may have its genesis traced back to a symmetric state in evolution, which broke symmetry. (iv) Other symmetries (the Case 3 Symmetry, discussed below) have been found to be particularly desirable in enabling oscillatory behaviour: in fact a thorough parametric analysis of oscillatory behaviour in certain random double-site modification networks reveals clusters in parameter space centred around parameter sets representing networks with this symmetry [99]. The Case 3 symmetry involves a combination of two symmetries (Case 1 and Case 2) which we individually study as well. (v) Symmetry-breaking can confer new functionality and information processing characteristics enriching the repertoire of multisite phosphorylation. (vi) Our study of symmetric systems allows us to draw important insights about multisite modification even when exact symmetry does not hold good. Thus it is also relevant to networks which are approximately symmetric. (vii) In this sense the symmetric scenarios also serve as valuable (and sometimes non-obvious) vantage points from which to investigate important aspects of multisite modification. Furthermore, while studying modification networks of larger numbers of modifications, the symmetric networks may represent one of the few tractable vantage points from which to study and elucidate the behaviour of such networks. (viii) These serve as interesting candidates for engineering multisite modification in synthetic biology with desirable features.

We examine basic models of multisite phosphorylation and evaluate the possibility of spontaneous symmetry breaking in basic and canonical reaction pathways / circuits / networks of multisite phosphorylation. We discuss the consequences of the results which emerge, for multisite modification networks which may not possess an exact symmetry. We then discuss the various consequences of such behaviour for biological systems, and cellular signalling pathways and networks which contain multisite modification. The ordering of multisite modification is a fundamental aspect of substrate modification and its regulation, and the deployment of modified substrates in various processes. It has been the focus of different studies [100, 82, 66, 65] spanning

canonical pathways, important cellular processes, basic principles and engineering for synthetic biology. We show how symmetry-breaking could provide a natural mechanism for the creation of ordered multisite modification systems from random multisite modification, which could in turn explain the various degrees of ordering encountered in cells

## 2.2 Methodology

Our approach to analyse the different networks in this paper, relies on a careful balance of analytical and computational work. Through these two strands we clearly isolate the presence or absence of classes of symmetry breaking, and where possible elucidate the necessary and sufficient conditions for this to occur. The networks described above as system of ODEs were simulated using the ode15s solver in *MATLAB* [70]. The results of the simulations were complemented and cross-verified using the computational software *COPASI* [69]. *COPASI* automatically generates the system of ODEs based on provision of the network schematic and thus is used to cross validate the *MATLAB* models. Bifurcation analysis was carried out using the computational package *MatCont* [71] in *MATLAB*, and the symbolic software package *Maple* [73] was used to cross verify the analytical results.

The bifurcation analysis presented in the paper is performed by varying the total substrate concentration $A_{Total}$. $A_{Total}$ is a natural choice for a bifurcation parameter in this context as varying this parameter still accommodates the different classes of symmetries (Case 1-3). However the symmetry breaking in all cases reported can also be isolated from bifurcation along any kinetic parameter (or parameter pair) or total enzyme concentrations, as long as the original symmetric structure is maintained.

We comment here that echoes of the symmetry breaking observed here are also seen when exact symmetry is not maintained (as shown in figs. 2.9 and 2.10). Further discussion on approximately symmetric systems and echoes of 'symmetry-breaking' in such contexts is presented in the main text.

**Parameters:** The parameters used to generate the figures in the paper are presented in appendix 2. Our results focus on instances of symmetry breaking and associated behavior. The parameters used are generic and are in ranges commensurate with values typically reported in literature. We emphasize that we use computation here primarily as a tool to complement analysis from analytical work and to show the presence of symmetry breaking and associated features in a model. Thus, in this spirit, the values used are only representative and the behavior is seen in a well-defined region of the parameter space (as we demonstrate in the analytical work). Analytical work also explicitly reveals basic features about the symmetry broken state seen computationally. In networks where symmetry-breaking does not occur, this fact is established analytically. The analytical results concerning features of symmetry breaking and the asym-

metric states are cross-validated by computational analysis (bifurcation analysis). This cross-validation is presented along with the parameters used in the *Maple* file and in the extended appendix (see appendix C).

Mass action kinetic descriptions are used for the elementary reactions involved as this does not make any assumptions regarding regimes of enzyme activity and complex formation (unlike other reaction descriptions such as Hill kinetics and Michaelis-Menten kinetics). Thus the results and model behavior shown are not artifacts of the choice of kinetic description used and are representative of the true functionality of these networks.

**Computational Resources:** A *Maple* file with the model descriptions and detailed analytical proofs is presented along with this text. A PDF copy of the entire *Maple* document is also provided in the extended appendix (see appendix C), with the files printed in the same order as they appear within the *Maple* document, for easy accessibility.

**Analytical Proofs:** The chapter appendix attached at the end of the chapter is in two parts. Part 1 (appendix B.1) contains the model descriptions for the various systems modelled in the chapter, and part 2 (appendix B.1), details the analytical argument for the various proofs described within the chapter. This is in addition to the extended appendix.

## 2.3 Results

We begin by discussing basic aspects of the models we employ and the way they are analyzed before proceeding to the results. We discuss the multisite modification networks we study, and the possible symmetries they may exhibit (with further details in appendix B.1).

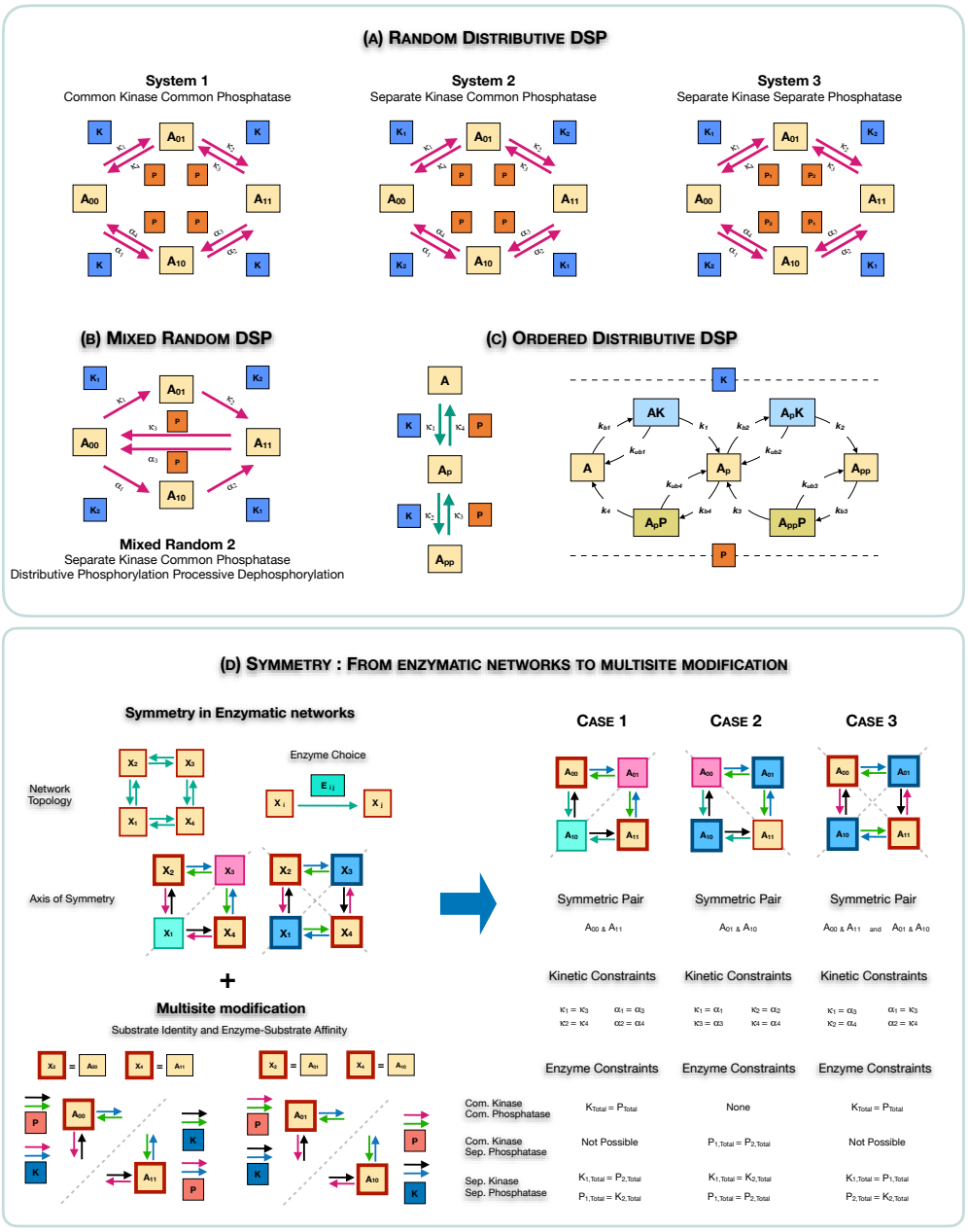



**Figure 2.1:** (A-C). **Schematic representation of the various multisite phosphorylation networks considered in the chapter.** (A) depicts the core networks while (B,C) serve as suitable contrasts to illustrate basic points. The labels $\kappa_i$, $\alpha_i$ in the schematic represent the triplet of binding, unbinding and catalytic rate constants involved in the enzyme modification for the $i^{th}$ modification (on each leg of the network). Detailed model description is provided in fig. 2.14. (A) shows the various random (distributive) double site phosphorylation (DSP) networks (the focal point of the study) with different combinations of enzyme action (common kinase and common phosphatase, separate kinase and common phosphatase, and separate kinase and separate phosphatase). (B) shows the random DSP with distributive phosphorylation and processive dephosphorylation (depicted for simplicity as direct arrows from $A_{11} \rightarrow A_{00}$) with separate kinase and common phosphatase (Model: Mixed Random 2). (C) shows the ordered distributive DSP with common kinase and common phosphatase, and an expanded description of reaction mechanism showing in detail the binding, unbinding and catalytic action for each modification step. (D). **Schematic representation of the three different classes of symmetries in the random DSP networks considered in this study.**

Continued on next page...

The different symmetries are depicted in the right hand panel. In each case, the axis of symmetry is depicted, and nodes of the network on either side of this axis (enclosed in a boundary of the same colour) have equal concentration. Identically coloured arrows in schematics are indicative of equal kinetic rate constants (for the corresponding triplet of binding/unbinding/catalysis reaction) and equal total concentrations of enzymes involved. The associated kinetic and enzymatic requirements required for enforcing symmetries are also listed. These are key ingredient in establishing symmetry in the reaction network. The origins of these symmetries can be conceptualized and visualized by examining an enzymatic network with a "square" topology (left hand panel), where every reaction is mediated by an enzyme. Such networks can have symmetry about one axis or two axes. Now examining the single axis reflection symmetry in multi-site modification results in two possibilities of symmetric nodes (the pair $(A_{00}, A_{11})$ and $(A_{01}, A_{10})$). In each case - symmetry requires different pairs of reactions (depicted by identically coloured arrows) to have equivalent rates and enzyme amounts. Importantly, in the $A_{11} = A_{00}$ case these pairs of reactions are associated with enzymes of different kinds, while in the $A_{01} = A_{10}$ case they are associated with enzymes of the same kind. While this is depicted for a single kinase and a single phosphatase it applies to any combination of common/separate kinase and phosphatase. This dichotomy underscores the difference between case 1 and case 2 symmetry. Overall this conceptualization allows us to obtain the three symmetries along with the kinetic and enzymatic requirements shown in the right hand panel.

**Models of multi-site modification.** Our primary focus is on random mechanisms of multisite modification, and we study the case of double-site modification, as a tractable, representative case. Figure 2.1A represents random mechanisms of modification (i.e. modifications can proceed in either order), and depicts cases where the kinases and phosphatases effecting individual modifications could either be the same or different. Taken together, these networks span a range of basic cases of multisite modification, including the possibility of an enzyme performing multiple modifications (seen in many biological contexts) and the possibility that this may be associated with one modification direction, but not the other (due to the fact that kinases significantly outnumber phosphatases, as seen in genome-wide studies [101]). When a common enzyme is involved in effecting multiple modifications, the modification mechanism is assumed to be distributive, unless otherwise stated. We note here that such modification circuits are encountered in multiple cellular contexts, and can be viewed as building blocks of more complex multisite modification networks. Such networks have been the focus of detailed studies in contexts such as circadian oscillations [102] (involving the common kinase common phosphatase network depicted: the substrate represents the Per proteins), with additional studies on temperature compensation in this context [103, 104]. They have also been used to evaluate design principles for both oscillatory and pattern forming behaviour more broadly [99, 105]. For purposes of contrast and elucidating basic effects, we also examine two related modification networks: (a) an ordered double-site modification mechanism (Figure 2.1(C)) mediated by a common kinase and common phosphatase. The specific ordering of modification involves the phosphorylation order being opposite to that of dephosphorylation resulting in one partial phosphoform. This has been extensively studied in the literature (eg. [29, 31]). (b) A random mechanism where the dephosphorylation is processive (fig. 2.1B and fig. 2.5A).

**From networks to models.** The model for each of these networks (depicted in fig. 2.1(A-C), fig. 2.5(A)) is built up from widely-used descriptions of individual substrate modification by an enzyme (involving reversible complex formation and irreversible modification to give the product - see fig. 2.14). Such a description makes no a priori assumptions about the kinetic regimes of modifications. Further details are presented in the appendix B.1. Throughout this chapter we work with these canonical modification circuits, where the substrate forms are denoted by $A$, with subscripts denoting the type of modification. Depending on the context these could represent different proteins.

**Associated network symmetries.** In order to understand and visualize the types of symmetries we will examine, it is fruitful to examine a "square" reaction network, which has the same network topology as that of the multisite modification networks above. Note that in this depiction, the nodes of the network represent substrates, while the enzymes are implicitly present in the arrows: both substrates and enzymes together constitute an enzymatic reaction network of this type. As depicted in fig. 2.1D (left panel) there are two types of symmetries which can be encountered: (i) In the first case, the two "legs" of the network are symmetric (about the axis of symmetry depicted), This means that the rates of reaction for corresponding reactions on either side of the axis are the same. The associated pair of species (nodes of this network representing substrates, on either side of the axis of symmetry) is expected to behave identically (assuming the same initial conditions). (ii) In the second case, the symmetry is associated with two pairs of species simultaneously and can be viewed as a simultaneous occurrence of two of the previous symmetries, along different axes (see fig. 2.1D). Viewed from a general network perspective, in each case the symmetry is a consequence of rates of different pairs of reactions (intrinsic reaction rate constants as well as total enzyme concentrations) being identical, thus giving rise to the symmetry. Thus enabling such symmetries establishes correspondences/constraints between different pairs of enzymes. Note that in this network, we have not made any restrictions on which enzymes may be involved in specific steps. Establishing the structural requirements for symmetry then allows us to examine when and how the multisite modification networks we study exhibit different symmetries.

**Network Symmetry meets multisite phosphorylation.** We now focus on the network symmetries in the specific instance of the biochemistry of multisite modification. In so doing, we discuss different types of symmetries which multisite modification networks can exhibit. While some of these symmetries may appear to be more natural biologically, it is useful to examine all these together to obtain a comprehensive systems understanding. Furthermore, some of these have been postulated explicitly or invoked implicitly in multiple different contexts.

Symmetries in such networks require basic conditions/constraints on the kinetics and enzyme amounts (refer fig. 2.1D), right panel). In particular, equivalence between two reactions (as represented in the schematic) requires that the rate constants of their constituent elementary reactions (binding, unbinding and catalytic) remain equal. The

first two cases of symmetries correspond to a scenario where the two "legs" of the network are symmetric. The difference between them is what the symmetric nodes of the network correspond to in the context of multisite modification along with the fundamentally distinct pairings of enzymes in each case (discussed further below).

**Case 1 symmetry:** $[A_{00}] = [A_{11}]$**.** In this case the nodes involved in either leg of the symmetric network are $A_{00}$ and $A_{11}$. In such a case, the requirement of a symmetry implies that for these two phosphoforms, the action of an enzyme (kinase) on one of these substrates ($A_{00}$) has the same rate as that of another enzyme (phosphatase) on the other substrate ($A_{11}$) (this is seen by the corresponding reaction arrows on the two legs of the network). Furthermore, an analogous requirement applies to the production of each of these species from the partial phosphoforms. With these requirements a symmetry between $A_{00}$ and $A_{11}$ is maintained. We further note that such a requirement (of having certain rates of kinase-mediated reactions being equal to that of other phosphate-based reactions) places a constraint on total enzyme amounts as well. The Case 1 Symmetry is of interest both as a basic independent symmetry, and as a constituent of the Case 3 Symmetry discussed below.

**Accommodating the requirements for symmetry.** 1. The above requirements can be accommodated both in the common-kinase common-phosphatase case, and the separate-kinase separate-phosphatase case, but not in the separate kinase common phosphatase case (discussed in appendix B.1). 2. We also note that a simpler network, that corresponding to ordered double site modification, also accommodates a symmetry of this type (while only possessing a single partial phosphoform). Here too, this is accommodated in the common-kinase common-phosphatase and separate-kinase separate-phosphatase cases.

**Case 2 symmetry:** $[A_{01}] = [A_{10}]$. In this case the nodes involved in either leg of the symmetric network are $A_{01}$ and $A_{10}$. Such a symmetry is realized if the following pairs of reactions have the same rates, (i) phosphorylation of $A_{00}$ to produce the respective partial phosphoforms, (ii) dephosphorylation of $A_{11}$ to produce the respective partial phosphoforms, (iii) the phosphorylation of the respective partial phosphoforms, and (iv) the dephosphorylation of the two partial phosphoforms. Note that equal rates of reaction requires the same intrinsic kinetic rate constants (for binding, unbinding and catalysis of substrate by enzymes) as well as total enzyme amounts. This is characterized by saying that the rate of modification of all substrates of a given level of modification are the same (and likewise for demodification) and this ensures that progression in substrate modification is equally balanced between the pathways associated with each partial phosphoform (a feature explicitly/implicitly assumed in multiple instances in the literature). This symmetry can be accommodated in all cases of separate/common kinases and separate/common phosphatases.

**Difference between Case 1 and Case 2 Symmetries.** Case 1 and Case 2 Symmetries involve different pairs of symmetric nodes. As noted earlier, the symmetries require both

intrinsic rate constants and enzyme amounts to be equal for different pairs of enzymes. The essential difference between the two cases is the essentially different enzyme pairs associated with this. In Case 1 Symmetry, the pairing is between enzymes of different types (a kinase and a phosphatase), while in the Case 2, it is between enzymes of the same type (between kinases and between phosphatases). This is exactly why the Case 1 symmetry is not possible in the separate kinase common phosphatase network, while the Case 2 symmetry is.

**Case 3 symmetry:** $[A_{00}] = [A_{11}]$ **and** $[A_{01}] = [A_{10}]$ simultaneously. This involves the combination of the earlier cases. Here the action of a kinase enzyme on a substrate occurs at the same rate as that of the phosphatase enzyme on its associated substrate, in the diagonally opposite modification leg. Thus the action of the kinase on $A_{00}$ modifying it to $A_{01}$ is the same as that of the phosphatase action on $A_{11}$ modifying it to $A_{10}$. This applies to the modification of all substrates in the network. Such symmetries have been implicated in oscillatory networks of multisite modification underlying circadian oscillators [99]. Again, similar to the Case 1 Symmetry, the separate-kinase common-phosphatase case cannot accommodate this symmetry.

**The basic questions.** The basic questions we address below are (i) Are these symmetries always maintained or can they be broken? (ii) What network features determine whether or not symmetry breaking is possible? (iii) What kind of capabilities does symmetry-breaking contribute? (iv) When symmetry breaking is possible, can the parameter regimes for symmetry breaking be established?

**Methods of Analysis.** To address the above questions, we employ two approaches in tandem (i) computational analysis, involving simulations and bifurcation analysis, where we demonstrate the possibility of such behaviour occurring. We note here that our bifurcation parameter is the total substrate concentration, though it could apply to other parameters (ii) analytical work which rules out the possibility of symmetry breaking in networks irrespective of kinetic parameters, bringing to the fore structural features which prevent the occurrence of the behaviour. Analytical work is also used to demonstrate necessary conditions for symmetry breaking (in terms of kinetic parameters and total enzyme and substrate amounts) and further that in these cases these conditions are sufficient to guarantee the presence of symmetry breaking. Additionally, analytical work also reveals important characteristics of symmetry-broken states.

We present the results for Case 1, Case 2 and Case 3 for the different random modification networks below. The ordered double-site modification network can exhibit Case 1 symmetry, as noted above. Therefore, in presenting the Case 1 symmetry, we start with this simpler network, before proceeding to the Random Modification Networks.

**Analysis of a simpler ordered mechanism reveals the origins of case 1 Symmetry breaking.** We first analyze the scenario of Case 1 Symmetry. It is instructive to examine an ordered mechanism (Figure 2.2A) in this regard, as it is simpler while exhibiting the same behaviour encountered in random mechanisms. The system has a symmetric

steady state which is characterized by (i) equal concentrations of unmodified ($A$) and fully modified ($A_{pp}$) phosphoforms and (ii) equal concentrations of free kinase and phosphatase (note that the total concentrations of these enzymes need to be the same for symmetry to be present). This steady state simply represents an absence of bias in the direction of modification (i.e. between the unmodified and fully-modified phosphoforms). However as the total substrate concentration is varied, we find that this steady state loses stability via a supercritical pitchfork bifurcation [106]. A pair of asymmetric steady states emerge which are stable. These correspond to either $[A_{pp}] > [A]$ or the other way around, and unequal free enzyme concentrations. Interestingly on each of these steady state branches the value of the intermediate phosphoform $A_p$ remains fixed at the level at the bifurcation point (fig. 2.2A, lower panel). The presence of asymmetric states, as well as the fact that the partial phosphoform concentration is fixed on the branches of asymmetric steady states is established analytically (see appendix B.1, and appendix C).

**Conditions for symmetry-breaking.** The asymmetric steady states represent the establishment of overall directionality in the reaction network output, even in the absence of any a priori bias (in terms of reaction rates and enzyme concentrations). Analytical work also reveals the necessary conditions for symmetry breaking to occur in this system: $k_2 > k_1$. In other words, the catalytic rate constant for the second phosphorylation is greater than that of the first phosphorylation step. Note that this does not involve binding or unbinding constants for enzyme/substrate interactions. Further analysis indicates that in such a parameter regime, a symmetry-broken state is guaranteed to exist for some value of the bifurcation parameter $A_{Total}$ (see appendix B.1 and appendix C). It is worth emphasizing here that (i) the nonlinearity responsible for the symmetry breaking arises purely from sequestration effects with no explicit feedback present (ii) an analogous case of single site modification is incapable of intrinsically exhibiting such behaviour.

**The random mechanism with common kinase and phosphatase.** The symmetry-breaking observed in this ordered mechanism is seen in the random modification with common kinase and phosphatase (Figure 2.2B). The random modification network can be thought of as two connected "legs" of ordered modification networks, and consequently echoes of the behaviour seen previously are observed here. In this instance, beyond the bifurcation point, the concentrations of both partial phosphoforms remain fixed at their values at the bifurcation point. This is established analytically (see appendix B.1 and appendix C).

**Conditions for symmetry-breaking.** An analytical necessary criterion for the presence of symmetry breaking in this system is presented in appendix B.1 and appendix C. It takes the form $(k_2 - k_1) + \alpha(k_2/a_2)(a_2 - a_1) > 0$. Here $k_1, k_2$ are the catalytic rate constants associated with phosphorylation along one leg of modifications ($A_{00} \rightarrow A_{01} \rightarrow A_{11}$), while $a_1, a_2$ are the catalytic rate constants associated with phosphorylation along the other leg of modifications ($A_{00} \rightarrow A_{10} \rightarrow A_{11}$). Further $\alpha$ is a positive constant de-

termined in terms by rate constants (including binding and unbinding). Further work in appendix C establishes the fact that this is a sufficient condition for the generation of asymmetric states for some value of $A_{Total}$.

**A comparison with ordered mechanisms reveals additional flexibility available for symmetry-breaking in random mechanisms.** We can make multiple inferences from this condition, in relation to the corresponding condition for ordered mechanisms discussed above. (i) If both $k_2 > k_1$ and $a_2 > a_1$ then the requisite condition is satisfied. This means that if each leg (viewed as an ordered mechanism) satisfies the conditions for symmetry breaking in ordered mechanisms, this guarantees the possibility of symmetry breaking in the random mechanism. (ii) For the same reason, if neither leg satisfies the condition, then symmetry breaking is precluded in the random modification network (iii) Interestingly if only one leg satisfies the criterion for symmetry breaking, it is possible for the entire random network to break symmetry (an example of this is presented in fig. 2.7). In such a case, the random network can be viewed as being comprised of two sub-networks, only one of which is the driver of this behaviour.

**Symmetry-breaking is possible even if the enzymes performing each modification are different.** Random mechanisms with different kinases and phosphatases can also exhibit the same type of symmetry (this places constraints on total concentrations of "corresponding" enzymes, in addition to the kinetic constraints already discussed). As seen in Figure 2.2C, this system also exhibits a similar symmetry breaking behaviour as encountered above, and here again the concentration of partial phosphoforms is fixed beyond the bifurcation. The case of different kinases and phosphatases represents a much more general case (not requiring any enzyme to perform more than one modification) with significantly reduced nonlinearity (for the same reason), which is nonetheless sufficient for symmetry breaking.

**Conditions for symmetry-breaking reveal requirements on both modification legs.** The necessary conditions for symmetry breaking here are established analytically in the see appendix C, where it is also shown that these conditions guarantee the existence of a symmetry-broken state. This takes the form $k_2/k_1 > P_{2Total}/P_{1Total}$ and $a_2/a_1 > P_{1Total}/P_{2Total}$ (the equation could also be written in terms of the total concentrations of kinases). The main difference here is that (i) there are two such conditions to be satisfied and (ii) they also involve total enzyme amounts. (iii) When $P_{2Total} = P_{1Total}$, this amounts to $k_2 > k_1$ AND $a_2 > a_1$, which is a requirement for each of the legs of the modification network. (iv) When $P_{2Total} \neq P_{1Total}$, this amounts to a tighter requirement on one of the legs (where enzyme $P_2$ is involved in the first dephosphorylation step) and a weaker requirement on the other (where $P_2$ is involved in the second dephosphorylation step).

**Symmetry-breaking cannot be observed in an ordered mechanism constituting a single leg of the modification if all modifications are effected by different enzymes.** Each leg of the modification corresponds to an ordered modification mechanism with

different kinases and phosphatases, which is incapable of symmetry breaking (as shown in appendix C) and multistability in general. Thus the observed symmetry breaking is an emergent behaviour of the entire network with both modification legs.

**Implications.** The implication of Case 1 Symmetry-breaking is that it is possible to establish a directionality to the modification even if none existed, resulting in an establishment of relative dominance of fully modified phosphoforms vis-a-vis fully unmodified forms or the other way round. Case 1 Symmetry is also a constituent of the Case 3 Symmetry, and this has implications in that situation as well.

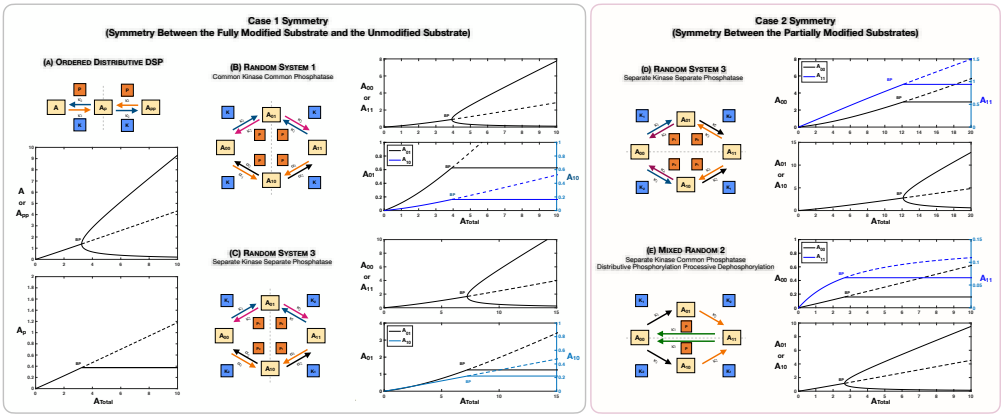



**Figure 2.2: Case 1 and case 2 symmetry breaking in various double site phosphorylation networks.** (A-C). Case 1 symmetry breaking in distributive DSP: (A), ordered DSP with common kinase and common phosphatase, (B), random DSP with common kinase and common phosphatase, (C), random DSP with separate kinase and separate phosphatase. Note that in these cases the concentrations of the partially modified substrates are fixed after symmetry breaking (in the symmetry broken state) in the bifurcation diagrams. (D-E). Case 2 symmetry breaking: (D), random DSP with separate kinase and separate phosphatase, (E), mixed random DSP with distributive phosphorylation through separate kinases and processive dephosphorylation through common phosphatase. Note that in these cases the concentrations of the fully modified and unmodified substrate are fixed after symmetry breaking (in the symmetry-broken state) in the bifurcation diagrams. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Dashed lines in the schematic represent axis of symmetry of the network. BP: (branch point) Pitchfork bifurcation]

**Case 2 Symmetry: When can it break?** Case 2 symmetry implies that there is no bias in the ordering of modifications and this is manifest in the equal steady state concentrations of the partial phosphoforms $A_{01}$ and $A_{10}$. Examining all the cases of random networks together (fig. 2.1($A-C$)) reveals the following insights: (i) The case of common kinase and common phosphatase will not lead to the breaking of this symmetry. (ii) The case of different kinase and common phosphatase will also not lead to the breaking of this symmetry. In both cases, this can be shown directly analytically, by demonstrating that for any steady states (irrespective of parameters), the partial phosphoforms necessarily must have equal concentrations (see appendix C). Incidentally, an identical conclusion can be drawn for the common kinase common phosphatase case, irrespective of the number of modification sites. (iii) On the other hand, the case of different

kinase and different phosphatase can indeed exhibit the breaking of this symmetry (fig. 2.2D) via a supercritical pitchfork bifurcation: here the asymmetric steady states are characterized by fixed values of concentrations of unmodified and fully modified phosphoforms. This is established analytically (see appendix C).

**Conditions for symmetry breaking.** Analytical results provide further insights. The necessary requirements for an asymmetric state are: $k_1/k_4 > P_{Total}/K_{Total}$ and $k_2/k_3 < P_{Total}/K_{Total}$. Note here that $K_{Total}$ denotes the total concentration of each of the kinase enzymes while $P_{Total}$ denotes the total concentration of each of the phosphatase enzymes (the equality, a requirement of symmetry). $k_1, k_2$ denote the catalytic constants for phosphorylation of $A$ and the partial phosphoforms (the constants being equal), while $k_3, k_4$ denote the catalytic constants of dephosphorylation of $A_{11}$ and the partial phosphoforms. Further analysis shows that this is sufficient for the existence of an asymmetric state. From this we can infer that (a) If $k_1/k_4 < k_2/k_3$ then this behaviour is precluded. (b) On the other hand if $k_1/k_4 > k_2/k_3$, then by suitable choices of total enzyme concentrations (quantities easy to manipulate), these conditions can be satisfied. (iii) In such a case there is only a finite range of ratio of total enzyme concentrations (between $k_2/k_3$ and $k_1/k_4$) which can result in symmetry-breaking. (iv) The presence of multiple kinases and phosphatases proves to have a combination of both sufficient nonlinearity as well as sufficient flexibility (from the multiplicity of enzymes) to enable such behaviour. As seen previously with multiple kinases/phosphatases, the parameters need to satisfy two inequalities. A combination of the naturalness of the symmetry and the widely encountered modification scenario suggests that symmetry breaking here may be encountered quite broadly.

**Case 2 Symmetry breaking in a separate kinase common phosphatase network with processive dephosphorylation.** We aimed to get further insights into the factors driving such behaviour by comparing it with other related networks. To do this, we examined associated random modification networks where the dephosphorylation was processive (Figure 2.1(B), fig. 2.5(A)). Note that this implies that the phosphatase has to be the same for all modifications. Here we find that if the kinase is common to different phosphorylation steps, the symmetry does not break; however, surprisingly when the kinases are different, symmetry-breaking does indeed happen, as shown in fig. 2.2E (again reinforcing the flexibility provided by having different enzymes in enabling such behaviour). This is supported analytically (see appendix C). A direct comparison with the different kinase common phosphatase random mechanism above reveals a new dimension: having processive dephosphorylation while reducing the nonlinearity in the network, actually enables this behaviour which was otherwise precluded. Elsewhere we have noted how having processive dephosphorylation, could readily enable other behaviour (oscillations) which was not observed when the modification was distributive [26]. We also note that different conditions in the cell (or stimuli) could effect a transition from distributive to processive mechanisms, as demonstrated [107, 108, 100].

**Implications.** Case 2 Symmetry-breaking ultimately results in the biasing of one or-

dered sequence of modifications over another. The implications of this as a key step for generating ordering of modifications is discussed subsequently.

**Case 3 Symmetry breaking reveals the simultaneous breaking of two symmetries.** The scenario of Case 3 symmetry is examined in Figure 2.3. This involves a combination of the earlier cases. Figure 2.3A (and fig. 2.6) focuses on the common kinase common phosphatase case. Here too, the symmetric steady state (characterized by $[A_{01}] = [A_{10}]$ and $[A_{00}] = [A_{11}]$), can lose symmetry to a pair of asymmetric steady states, via a supercritical pitchfork bifurcation (fig. 2.6). The noteworthy point here is that both symmetries necessarily break together in general. In addition to previously observed bifurcation patterns, new possibilities arise. One is the possibility of a subcritical pitchfork bifurcation Figure 2.3(A). In such a case the branches of asymmetric states are unstable at the point of inception, but following a saddle-node bifurcation, they become stable. As a direct consequence of this, when the parameter (total concentration of substrate) is varied, a sudden change from a symmetric to asymmetric steady state is observed in simulations, the latter exhibiting a pronounced asymmetry, which is not a small perturbation of the symmetric state. The asymmetric steady states are characterized by high levels of $A_{11}$ and one of the partial phosphoforms, and low levels of $A_{00}$ and the other phosphoform, or the other way around. Furthermore the grouping in which each partial phosphoform is present is determined by baseline parameters (and changing baseline parameters can alter the grouping).

Irrespective of the nature of the pitchfork bifurcation (supercritical or subcritical), we find that the sum of partial phosphoform concentrations remains fixed on the asymmetric branches, and this is established analytically (see appendix B.1 and appendix C). The implications of this distinct type of invariant are discussed in the next section. It is interesting to contrast the invariant in this Case 3 Symmetry breaking with invariants in symmetry breaking in the constituent symmetries (in this instance of common kinase common phosphatase, the Case 1 Symmetry, since Case 2 symmetry does not break). In the Case 1 Symmetry breaking, the invariants are the individual partial phosphoform concentrations while here it is the sum.

**Implications.** This simultaneous breaking of symmetries and pairing of partial phosphoform and fully modified (or fully unmodified phosphoform) has a transparent interpretation. Symmetry breaking simultaneously imposes directionality to the modification (i.e. relative dominance of fully modified vs fully unmodified phosphoform) as well as a particular route of modification (via one of the two phosphoforms).

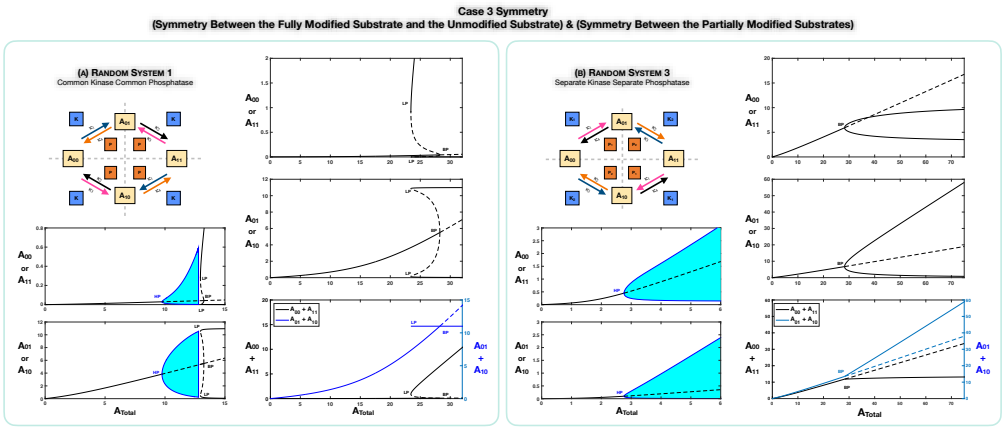



**Figure 2.3: Case 3 symmetry breaking in various double site phosphorylation networks.** (A) shows case 3 symmetry breaking in random DSP with common kinase and common phosphatase. The symmetric steady state is capable of losing stability either through a Hopf bifurcation leading to oscillations, which is followed by symmetry breaking through a subcritical Pitchfork bifurcation eventually leading to stable asymmetric steady states (Column 1), or just by breaking symmetry leading to asymmetric branches through a subcritical Pitchfork bifurcation which eventually becomes stable (Column 2). As seen in the plots, both symmetries simultaneously break. Note that the sum of the concentrations of the partially modified substrates is fixed in the asymmetric states in the bifurcation diagrams. Symmetry-breaking via a supercritical pitchfork bifurcation, as seen previously in other cases, can also be seen (Figure 2.6). (B) shows case 3 symmetry breaking in random DSP with separate kinase and separate phosphatase. The symmetric steady state is capable of losing stability either through a Hopf bifurcation leading to oscillations (Column 1), or by breaking symmetry leading to two stable asymmetric branches through a supercritical Pitchfork bifurcation (Column 2). Note that the sum of the concentrations of the completely modified and completely unmodified substrates is approximately constant in the asymmetric steady states in the bifurcation diagram. (However case 3 symmetry breaking in this network is also capable of providing approximate concentration robustness in the sum of the concentrations of partial substrate forms - see main text and Figure 2.8). [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during such oscillations. BP: (branch point) Pitchfork bifurcation, HP: Hopf bifurcation, LP: (limit point) Saddle node bifurcation.]

**The presence of oscillations.** Another behaviour which is observed in a different parameter regime is the emergence of oscillations, via a Hopf bifurcation, and this precedes the subcritical pitchfork bifurcation (Figure 2.3A). The oscillations do not preserve the symmetry of the original system. Instead we see correlated changes between the corresponding pairs of substrates at different time intervals. As the total substrate concentration increases, the period of oscillations increases, as the periodic trajectory comes close to a steady state in the phase space (Figure 2.6). Oscillations in such networks can occur without symmetry-breaking and in fact oscillations emerging from such random modification networks have been the focus of earlier studies [99], Here we show that in the presence of symmetry (a condition recognized as a desirable criterion for oscillations), oscillations may be present in conjunction with symmetry breaking, which affects the oscillatory range and characteristics of oscillations.

**Conditions for symmetry breaking.** Analytical work in the case of common kinases and common phosphatases reveals a necessary condition for the realization of an asymmetric state (which is shown to be sufficient as well). This takes the form $c_3(1-k_3/k_2)+c_1(1-k_1/k_4) > 0$. Here $c_1$ and $c_3$ are known positive constants, which depend on kinetic parameters. As in the situation of Case 1 symmetry, this hinges on two ratios of catalytic constants, though in contrast to that case it is the phosphorylation and dephosphorylation rate constants associated with the inter-conversion between $A_{00}$ and $A_{01}$ ($k_1/k_4$) and dephosphorylation and phosphorylation rate constants for the interconversion between $A_{01}$ and $A_{11}$ ($k_3/k_2$). As before we can make a range of conclusions: (a) $k_3/k_2 > 1$ and $k_1/k_4 > 1$ will preclude an asymmetric state (b) $k_3/k_2 < 1$ and $k_1/k_4 < 1$ will ensure the possibility of an asymmetric state (c) A combination of parts of the above conditions can allow for an asymmetric state depending on the constants $c_1, c_3$. In this regard we also point out that these results show when symmetry-breaking is precluded, and this combined with [99] yield conditions under which oscillations can occur without interference from symmetry-breaking (and in fact multistability).

**The case of different kinases and phosphatases.** Here (Figure 2.3B), we again find symmetry breaking via supercritical pitchfork bifurcations (but not subcritical pitchfork bifurcations), and the symmetry broken states (which necessarily have both symmetries broken: see appendix C) are characterized by having a higher level of one pair of substrates (one partial phosphoform, and a completely modified/unmodified phosphoform) and a lower level for the other pair.

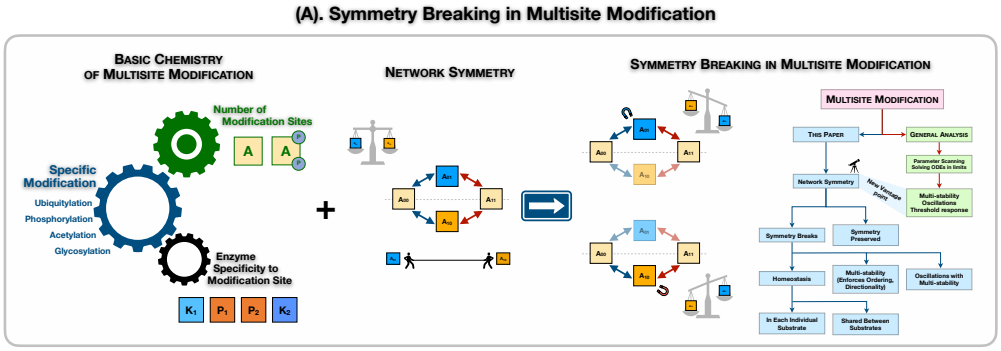
Here an exact invariant of the form examined previously does not hold: instead we find that (depending on parameters) either the sum of partial phosphoform concentrations, or the sum of concentrations of $A_{00}$ and $A_{11}$ is approximately constant (see Figure 2.3, Figure 2.8, appendix C). It is worth noting as a contrast that fixed individual concentrations for pairs of species $(A_{01}, A_{10})$ and $(A_{00}, A_{11})$ are obtained in this case for symmetry breaking of the constituent Case 1 and Case 2 symmetries respectively. The possibility of oscillations (stable over a broad range of parameters) emerging is also seen here (Figure 2.3(B)), though we have never found it occurring side-by-side with symmetry-breaking (seen earlier in the common kinase common phosphatase network). The presence of oscillations expands on and complements [99], by revealing oscillations in this network which is desirable as it has additional tuneable dials (multiple total enzyme amounts).

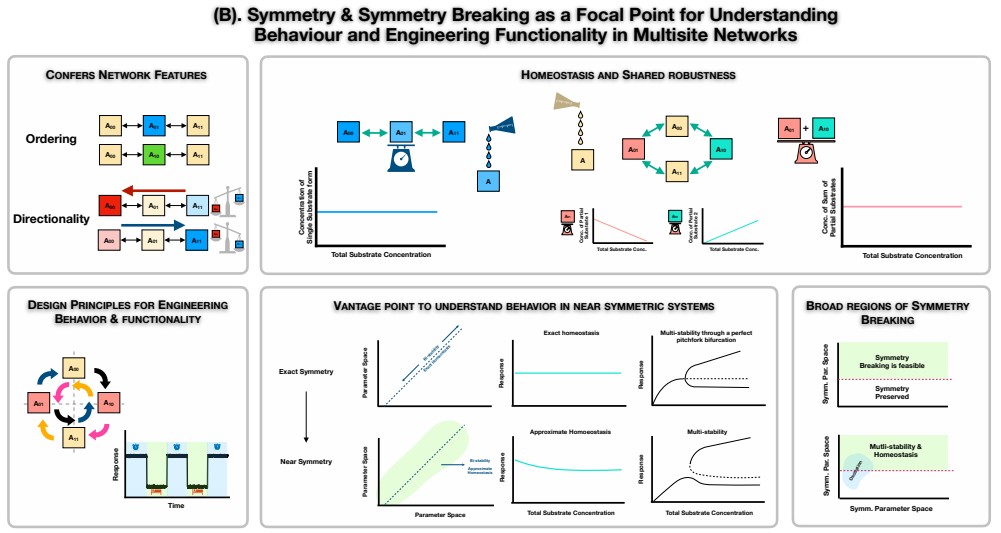
**Implications for inferences based on measurements.** In the Case 3 symmetry breaking in both the common-kinase common-phosphatase and separate kinase separate phosphatase, we find that in the symmetry broken state, the concentration of one partial phosphoform and either unmodified or fully modified form may be significantly elevated relative to its counterpart. This disparity between the two pairs can become pronounced, and easily be misinterpreted as suggesting (a) the nonexistence of other active modifications or (b) a significant disparity in intrinsic kinetic rates of modification in the two legs of the network.

## 2.4 Conclusions and Discussions

This chapter has focussed on symmetry-breaking behaviour in multisite phosphorylation systems (summarized in Figure 2.4). The wide prevalence and relevance of multisite phosphorylation is well-established, but why focus on symmetry?

There are multiple reasons for this: (i) The structure of basic modification networks for multisite phosphorylation (the topology as well as positions of enzyme action therein) admit to the possibility of symmetries. Additionally these have been sometimes implicitly or explicitly assumed in the literature [109, 110]: for example Case 2 symmetries where different partial phosphoforms behave in a similar way. In such instances, our results indicate that even if there is no biasing in the network interactions, the two phosphoforms in such a case may end up behaving very differently. Thus a simple, plausible assumption may have far-reaching and unexpected consequences. (ii) Certain symmetries may indeed naturally exist, for example the possibility that modification/demodification can proceed at an equal rate, irrespective of the modification site under consideration (the Case 2 Symmetry). In other instances (the Case 3 symmetry), exhaustive parametric analysis of random double site modification networks reveals the fact that oscillatory behaviour occurs in clusters centred around these symmetric networks (i.e. parameter sets which enable Case 3 symmetry) [99]. Thus networks which possess these symmetries (or represent small to moderate deviations therefrom) represent those enabling oscillations, suggestive of a basic design principle. The Case 3 Symmetry involves the Case 1 and Case 2 Symmetries as basic building blocks. Our analysis in both instances indicates distinct, unexpected outcomes which can emerge in terms of system behaviour and information processing characteristics (iii) The breaking of symmetries may have been exploited during the process of evolution: in such cases the presence of observed asymmetric networks may have its origins in symmetric cases encountered in evolution. In particular, as discussed further below, we show how multisite modification possesses natural ingredients for creating ordering by (Case 2) symmetry breaking. (iv) The insights we obtain are also relevant to systems where the exact symmetry may not strictly hold, but which are not large deviations of the symmetric case. In the latter case, clear echoes of the behaviour we discuss may continue to be observed (Figures 2.9 and 2.10). In such cases, the symmetric scenario provides a key vantage point from which to understand the origins of the behaviour. This behaviour may manifest itself as multistability in these cases, but in contrast to multistability which may be more broadly seen in parameter space, the origins and characteristics of the steady states in these instances can be traced back to the symmetric case and symmetry-breaking. This is also an example of how having clear-cut reference cases, allows us to elucidate how and why certain behaviour may arise, going beyond parameter scanning based model analysis and data analysis. We further emphasize that as the number of modifications increases, the underlying modification networks become considerably more complex (with an exponential increase in the size of the network, in the absence of ordering) and symmetric networks represent one of the few tractable




(C). Summary of Symmetry and Symmetry Breaking Behavior in MSP Networks

| | Case 1 | | Case 2 | | Case 3 | |
|---|---|---|---|---|---|---|
| **Ordered DSP** (Common Kinase Common Phosphatase) | Yes | Partial substrate form ($A_{10}$) exhibits concentration robustness post symmetry breaking | | | | |
| **Random System 1** (Common Kinase Common Phosphatase) | Yes | Partial substrate forms ($A_{00}$ & $A_{10}$) individually exhibit concentration robustness post symmetry breaking | No | | Yes | Sum of the partially modified substrates ($A_{01}$ + $A_{10}$) exhibit concentration robustness post symmetry breaking |
| **Random System 2** (Separate Kinase Common Phosphatase) | | | No | | | |
| **Random System 3** (Separate Kinase Separate Phosphatase) | Yes | Partial substrate forms ($A_{01}$ & $A_{10}$) individually exhibit concentration robustness post symmetry breaking | Yes | Completely modified and the unmodified substrates ($A_{00}$ & $A_{11}$) individually exhibit concentration robustness post symmetry breaking | Yes* | No exact robustness exhibited (Approximate concentration robustness in pairs of substrate possible - see text) |
| **Mixed Random 1** (Common Kinase Common Phosphatase Distributive Phosphorylation Processive Dephosphorylation) | | | No | | | |
| **Mixed Random 2** (Separate Kinase Common Phosphatase Distributive Phosphorylation Processive Dephosphorylation) | | | Yes | Completely modified and the unmodified substrates ($A_{00}$ & $A_{11}$) individually exhibit concentration robustness post symmetry breaking | | |

**Figure 2.4:** A. Schematic representation of realization of symmetry and symmetry breaking in multisite modification networks through the interplay of basic biochemistry of post translational protein modification and network symmetry. The analysis of the multisite network through symmetry and symmetry breaking reveals the underpinnings of new network features, and serves as a rich and distinct vantage point to understand information processing behavior in multisite modification networks. B. Symmetry breaking as a new vista for understanding behavior and engineering functionality in multi-site modification networks: Symmetry breaking can confer network features such as ordering and directionality to MSP networks, limit the range of oscillations and enable robust homeostasis for individual or combinations of substrates. It also provides key insights on the origin of behaviors in networks which are not far from symmetry. C. A tabular summary of the presence and absence of symmetry and symmetry breaking in MSP networks, along with features of the symmetry broken states (exact absolute concentration robustness).

vantage points from which to study such networks (v) In addition to revealing distinct new information processing characteristics of multisite modification (for instance exact absolute concentration robustness of different types) of relevance in natural systems, this also serves as a potentially fruitful point of departure in engineering information processing circuits involving multisite modification in synthetic biology.

Our studies primarily focussed on different random mechanisms of double-site modification, with different combinations of common/separate kinase and common/separate phosphatase. These serve as a useful basis for investigation, noting that, (i) these different types of enzyme combinations are widely encountered in cellular biology [24, 82, 111], (ii) the double site mechanism is among the simplest multisite modification system which both exhibits different types of symmetries and symmetry-breaking. This provides a tractable case for understanding this behaviour transparently, serving as a basis for subsequent investigation of more complex modification scenarios. Furthermore, the insights obtained from our analysis suggest natural extensions and generalizations to modification networks with a greater number of modification sites. This is seen in preliminary studies of ordered triple site modification systems (see appendix C and fig. 2.12).

**Symmetries.** The symmetries which arise can be conceptualized and understood, by examining symmetries of the underlying "square" reaction network topology (refer Figure 2.1D). That viewpoint, relevant for a broad class of chemical reaction networks, can then be applied to the specific instance of multisite substrate modification networks. Our studies involved a systematic analysis of different symmetries which emerge (noting the reasons above): symmetry in the modification direction (Case 1 Symmetry: arising from a symmetry in action of the two enzymes on their corresponding substrates), symmetry between different branches of modification (Case 2 Symmetry: arising from the same rates of modification to and from phosphoforms at a given level of modification: see text for distinction from Case 1) and combinations of the two (Case 3 Symmetry: arising from a symmetry in the action of the two enzymes in the modification/demodification at a given site, but on opposite legs). The different types of symmetries are encountered among the different classes of multisite modification networks (with either common/separate kinase/phosphatase) though some networks may not exhibit all symmetries (summarised in fig. 2.4C).

**Which symmetries can be broken?** The Case 1 symmetry can be broken in all the random modification networks where it exists. This symmetry breaking serves as a distinct mechanism for establishing directionality. Additionally, case 1 symmetry breaking can also be observed in a simple *ordered* DSP network (with only a single partial phosphoform), though only in the common kinase, common phosphatase case. This ordered network serves as a simpler network for understanding this symmetry breaking transparently. Case 2 Symmetry breaking is the basis of ordering of modifications. While Case 2 symmetry is possible for all combinations of common/separate kinase and common/separate phosphatase, the symmetry is broken only for the different kinase and

different phosphatase case. A combination of the flexibility afforded by the different sets of enzymes, along with sufficient nonlinearity (due to enzymes participation in multiple complexes) enables this. Interestingly, if the dephosphorylation mechanism is processive (rather than distributive, as assumed throughout), the separate kinase common phosphatase network can also give rise to Case 2 symmetry-breaking, reinforcing how the interplay of processive and distributive modification can enable new behaviour (see [26] for another example). In Case 3, the two symmetries necessarily break together. A distinct behaviour encountered here is the presence of symmetry breaking and oscillatory behaviour. In most cases, the symmetry broken states are associated with transparent invariants which we analytically identify: these represent a behaviour reminiscent of Absolute Concentration Robustness (discussed below). Additionally, the symmetry-breaking manifests as a supercritical Pitchfork bifurcation, while in the Case 3 symmetry with common kinase/common phosphatase, a subcritical Pitchfork bifurcation is observed, along with possible tristability. Our analytical work reveals how symmetry-breaking (in all cases studied) may be accessed in large regions of (symmetric) parameter space, by varying total enzyme/substrate concentrations, which represent easy to manipulate experimental factors (see fig. 2.11).

**Multisite modification and network symmetry breaking.** It is worth viewing the above results from a different perspective: the breaking of symmetry in (potentially general) biochemical networks of the "square" topology (discussed in Figure 2.1D). While symmetry simply imposes the restriction that the two legs of the network have kinetics which are identical, when can the symmetry actually break? Our study presents multiple insights: (i) Firstly a degree of non-linearity is required, and this arises from conservation of species and sequestration of enzymes/substrates in complexes, a fundamental aspect of biochemical systems. All the modification networks we consider have enzymes shared between at least two enzyme-substrate complexes (this stems from the fact that a given modification is effected by only one enzyme, and without any ordering) and this provides the non-linearity. (ii) On the other hand for symmetry breaking to occur, a sufficient flexibility is required in the network to be able to allow for this. This is clearly seen in the Case 2 symmetry in modification networks (with distributive enzyme mechanism), where reduced non-linearity notwithstanding, it is only the separate kinase separate phosphatase modification network that allows symmetry to be broken. In general, there is a trade off between nonlinearity and flexibility (associated with distinct enzymes for different steps), but multisite modification provides many instances of sufficient combinations of both factors to realize symmetry breaking. These insights bringing together basic (bio)chemistry and network features, are broadly relevant in biochemical networks.

**Enzyme Sequestration.** Enzyme sequestration (and competition) provides the key non-linearity for generating symmetry-breaking obviating the need for explicit feedback. Eliminating enzyme sequestration eliminates the possibility of symmetry-breaking. Enzyme competition is a key ingredient in multisite modification, and in general this

could combine with zeroth order ultra sensitivity to generate new behaviour. However, the symmetry-breaking we have found does not require any explicit assumption on the kinetic regime of enzymatic action (as seen from the sufficient conditions we have obtained) and so zeroth order ultra sensitivity is neither necessary nor sufficient for this.

We now discuss the relevance of our results from different vantage points. All of these underscore the fact that information processing is a characteristic and consequence of the modification network (rather than an individual modification) and that symmetry and symmetry breaking provides distinct classes of insights therein (see fig. 2.4).

**Ordering and Directionality in multisite modification.** Multisite modification systems encountered in vivo often exhibit different degrees of ordering ranging from complete ordering of the sequence of modifications, to partial ordering, to a complete absence of ordering (symmetric scenario). Ordering is a fundamental aspect of substrate modification and its deployment in different pathways and processes. In fact ordering of modifications is key to establishing a strictly sequential logic, which is likely to be an important aspect of information processing in those cellular contexts. A range of studies focus on these contexts, basic principles and the potent role in engineering multisite modification [100, 82, 111, 66, 65, 112]. How ordering has emerged is however unclear and there could be multiple contributing factors. Our results indicate that the basic biochemistry of multisite modification by itself provides the basis for creating an ordering, by breaking symmetry. The biasing which emerges can itself be very significant, and with possible additional refinements, give rise to ordering. This demonstrates that a key driving factor could be at the modification network level, rather than at the molecular level. Our analysis of the different symmetry cases allows us to explore the different ways in which both ordering and directionality may be determined. We determine explicit conditions for the occurrence of symmetry-breaking, revealing broad ranges of parameter space where this can happen. In the context of ordering, this, along with the demonstration of sufficiency of the conditions for symmetry breaking, demonstrates the robustness of the mechanism. We further point out that even if the system is not exactly symmetric, an echo of such symmetry breaking may be seen, which is indicated by multiple steady states which strongly bias one pathway over another, in a manner which is not commensurate with the (small) differences in kinetics of the pathways (Figures 2.9 and 2.10).

Given a symmetric (Case 2 Symmetry) or close to symmetric network where different phosphoforms behave (essentially) the same, there are different ways in which evolution could lead to biasing of one modification pathway over the other. One is by effecting local changes in one of the pathways. Symmetry breaking allows for a distinct mechanism whereby changing one easy to manipulate parameter (expression level of substrate), a significant biasing of one pathway over the other is established. This could be further reinforced (if this is a desirable outcome) by local changes in the pathway or increasing substrate amounts further (which further accentuates the biasing). This

can lead to either partial or even complete ordering subsequently. Thus the mechanism could be seen as an efficient way of effecting a substantial change which could be reinforced and consolidated by further tinkering. It can also generate different robustness characteristics.

A similar comment applies to directionality. Case 3 symmetry breaking results in a combination of ordering and directionality, which ultimately manifests itself as elevated combinations of specific partial phosphoforms and unmodified/fully modified forms. Such a behaviour of the network (for instance if observed experimentally) could easily be misinterpreted as suggestive of either some modification being inactive, or there being a strong bias in the intrinsic kinetics, neither of which may be correct. Our results also provide important insights in the cases of larger numbers of modification sites. For instance, analogues of Case 2 symmetry breaking could explain both ordering and partial ordering (some sequences of modifications ordered) in those systems.

**Absolute Concentration Robustness.** Our analysis reveals the presence of (exact) Absolute Concentration Robustness (ACR) of different species in the symmetry-broken state. In this regard we note: (i) The relevant species (in some cases, partial phosphoforms, and in others the fully modified or fully unmodified phosphoforms) exhibit concentration robustness to changes in total substrate concentration and is fixed at a level corresponding to the concentration of these species at the symmetry breaking bifurcation (the inception of the asymmetric branch). (ii) Depending on the network and the type of symmetry broken, this can manifest itself as ACR for pairs of species (Case 1 and Case 2). (iii) In other cases (Case 3), the robustness is in the sum of concentrations of species, either exactly or approximately. From the above points we see that multisite modification contains an in-built mechanism of creating robustness for clusters of species, either individually or collectively, something which represents an appealing characteristic for natural and engineered modification networks. It remains to be seen how this has been exploited in cells. (iv) There are different ways in which ACR may be obtained (for instance in bifunctional enzymes [113, 114]). The mechanism seen here shares a feature of ACR observed in autocatalytic networks, arising from a transcritical bifurcation [115, 116] as being intimately tied to a non-linear dynamic transition arising from the biochemistry. In both cases, there is more than one steady state possible, and one of the steady states exhibits the ACR.

**The origins of Absolute Concentration Robustness.** Based on the above, a natural question is which substrates could exhibit (exact) ACR and whether symmetry is a pre-requisite. We note that in the ACR we have made no assumption/restriction or invoked any particular kinetic regime for enzymatic action. We answer the questions (based on analytical work: see fig. 2.13) relating to ACR in these terms, in the ordered DSP network. (i) Only $A_p$ can exhibit ACR and this occurs only in response to $A_{Total}$ (not $K_{Total}$ or $P_{Total}$). (ii) ACR necessarily requires multiple steady states, with two branches of steady states exhibiting ACR.There is another steady state branch which does not exhibit ACR, but intersects one of the branches in what was computationally

observed to be a transcritical bifurcation (iii) There is a constraint on parameters to enable this, which is weaker than the symmetry condition. (iv) In the case of symmetry, the two ACR branches are symmetric and intersect with the other branch in a pitchfork bifurcation.

**Approximate ACR.** As noted above, networks deviating from exact symmetry can exhibit approximate concentration robustness (refer fig. 2.10). Concentration robustness (approximate) could also be obtained in specific limiting kinetic parameter regimes. In ordered DSP with common kinase common phosphatase we find that (i) $A_{pp}$ and $A$ could also exhibit concentration robustness. This can happen in a regime where the enzyme producing this (from $A_p$) acts in the saturated regime while the action of both enzymes on reactions not involving the species under consideration, act in the unsaturated limit (see appendix B.1). Here approximate ACR occurs without requiring multistability (ii) Similarly approximate ACR can occur in $A_p$ without multistability by (for instance) having phosphorylation of A in the saturated regime, and phosphorylation of $A_p$ and dephosphorylation of $A_{pp}$ in the unsaturated limit (or having phosphatases in excess). Similar insights can enable approximate ACR for one species in the corresponding random network (see appendix B.1). In contrast to such limiting regimes, absolute/approximate concentration robustness via symmetry breaking is present along with a rich repertoire of information processing characteristics.

**Pathways and Modularity.** How does symmetry-breaking in multi-site modification both affect and be affected by the behaviour of a signalling network of which it is a part? (i) The importance of multisite modification stems from the fact that it confers functionality to proteins which can regulate other processes. Here the symmetry breaking allows for both regulation of downstream pathways as well as insulation of some downstream pathways from the effect of total upstream substrate and other upstream perturbations (via ACR for specific substrates in the modification network). This ability to insulate some parts of a network, while not the others is a desirable feature which can be exploited: symmetry-breaking in multisite modification provides a way of realizing this exactly purely from chemistry without requiring elaborate network structures (incorporating adaptation, feedback and so on). (ii) Additionally, we find examples of "Shared Absolute Concentration Robustness" where the sum of two species concentrations is fixed. This represents a case where robustness is applied to a combination of pathways, if the two species regulate different pathways. This may be relevant in multiple cell signalling contexts, by directly incorporating an inbuilt trade-off between the activation of two pathways, for instance for efficient resource allocation. If the two species regulate the same pathway in the same way, this translates into robustness in regulating the pathway, while allowing flexibility through the redundancy. (iii) The effect of sequestration of a substrate species can be to either facilitate or make difficult the possibility of symmetry breaking. The sequestration of a substrate species in a downstream complex is the basis of a retroactive effect in a signalling pathway [117, 118]. In the current case, this retroactive effect can help facilitate the possibility of sym-

metry breaking, and further that this happen in a context-dependent way (iv) Other factors associated with the network, for instance feedback, may also significantly affect the possibility of this happening. These aspects need to be assessed systematically and will be studied in the future. Interestingly an existing study [119] examines sequential multisite modification with two explicit feedbacks: one from the maximally modified phosphoform increasing the probability of (every) modification and the other from the unmodified form increasing the probability of every de-modification. In a stochastic setting, this has been shown to result in breaking a symmetry between phosphorylation and dephosphorylation even with no enzyme sequestration. In contrast to this, all our studies are on the intrinsic behaviour of multisite modification, and in a deterministic setting.

**Relevance to oscillatory enzymatic networks.** Studies of multisite networks have focussed on their capacity of generating oscillations [30, 120], including random networks with common kinase and common phosphatase with a view to their relevance in circadian oscillators [99]. A detailed computational study [99] reveals regions of parameter space which facilitate the presence of oscillations, and the prominent regions are clustered around a symmetric network (the Case 3 network that we have studied). What is the relevance of our analysis here? Our study shows how oscillations occur in such cases, and also how (by changing substrate amounts), both oscillations and symmetry breaking may occur. We can identify different regimes based on our analysis. (a) A regime where symmetry breaking is ruled out. Here our analysis indicates regimes where oscillations can occur without any potential interference from symmetry-breaking. (b) A regime where symmetry breaking is possible, and in fact guaranteed, for some total substrate amount. In the latter case we demonstrate that by varying total enzyme amounts (easily tunable dials) it is possible to obtain, multistability, oscillations or a combination of such behaviour (see fig. 2.11). It indicates how in certain cases symmetry breaking may occur limiting/preventing a range of oscillations. In particular, it indicates that oscillations do not have to be present for an indefinitely large range (suggested in the computations of [99]). We provide further insights with regard to oscillations. (c) The existence of long period oscillations which hover between different symmetric states is also seen ( fig. 2.6). (d) We demonstrate the possibility of oscillations in networks with different kinases and phosphatases, which potentially benefit from a greater tuneability than the common-kinase common phosphatase case. (e) By contrast we find the other symmetries do not readily yield oscillatory behaviour, though further work needs to be done to study this exhaustively. The above points sharpen our understanding of oscillations emerging in random modification mechanisms and reinforce the theme of multisite modification as a complex information processor.

**Experimental signatures and testable predictions.** Our analysis reveals key features associated with symmetry-breaking, which suggest multiple non-trivial signatures which could be seen experimentally: for instance a considerable disparity in partial phospho-

form behaviour, which may be incommensurate with the (minor) differences in kinetics in the legs of modification, characteristic patterns of absolute concentration robustness for specific species or groups . These signatures even if approximate could suggest the presence of symmetry-breaking. On the other hand experiments could be developed to realize this behaviour by constructing underlying modification circuits, either for synthetic purposes or to probe and test the behaviour itself. The multiplicity of enzymes involved allows for the deployment of a broad experimental tool-kit for these purposes. Systems mimicking Case 2 symmetry could be created by engineering different modification sites (of similar properties), with modification by different isoforms of an enzyme. If this is also done for the demodification, then an approximate realization of a Case 2 Symmetry (different kinase different phosphatase) can be realized. Alternatively, in a similar vein it may be possible to engineer a different kinase common phosphatase system with processive dephosphorylation: here the dephosphorylation could be induced to be processive. Other approaches could involve reconstitution of components of existing systems, such as circadian oscillators.

It is worth examining implications and extensions of our study to a larger number of modification sites. Random networks lead to an exponential increase in the number of states. Additionally, the modifications/demodifications can be effected by common enzymes (for all modifications), distinct enzymes (for every modification) or a combination thereof, leading to a further combinatorial explosion in possibilities. Clearly direct analogues of the symmetry breaking seen here (for eg. Case 1 and Case 2) can be encountered here. In addition, new possibilities can emerge. In Case 2, for instance, in addition to the situation where all modification legs behave the same, we can have a situation where some modification legs (or parts thereof) are the same. Furthermore, not surprisingly, new behavioural characteristics can emerge. For instance in the ordered Triple Site Phosphorylation network (Case 1 Symmetry: Common Kinase Common Phosphatase), we find shared robustness (see fig. 2.12), not seen in the ordered double site modification. These aspects need a dedicated study of their own and will be studied in the future. Viewed from the perspective of information storage, symmetry breaking suggests that a symmetric double-site modification network contains a bit of information. We emphasize however that symmetric network encodes a richer set of information, such as simultaneously presenting homeostasis and multiple steady states, an observation relevant to networks of any number of modification sites.

All in all, we have shown how basic biochemistry of multisite modification even within simple modification networks can be the basis of symmetry breaking. Symmetry-breaking in turn can confer ordering, directionality, exact concentration robustness, and this can significantly enhance the repertoire of information processing in multisite modification and regulation in signalling networks where they are present. The insights which arise from a structured systems study is of relevance in multiple contexts spanning the chemical and the biological, from systems biology to systems chemistry with potential in synthetic biology and the engineering of chemical systems,

## 2.5 Appendix

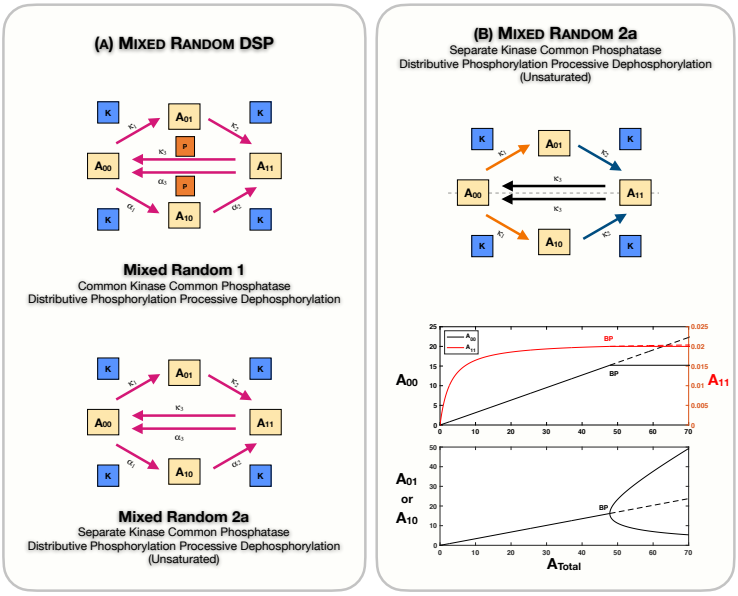



**Figure 2.5:** (A). Schematic representation of the mixed random ordered DSP network with common kinase and common phosphatase effecting distributive phosphorylation and processive dephosphorylation (Mixed Random 1), and the mixed random ordered DSP network with separate kinases and common phosphatase effecting distributive phosphorylation and processive dephosphorylation in the unsaturated regime (Mixed Random 2a). (B) shows case 2 symmetry breaking in the Mixed Random 2a network. Note that the concentration of the fully modified and unmodified substrate forms are fixed in the asymmetric steady states. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation]

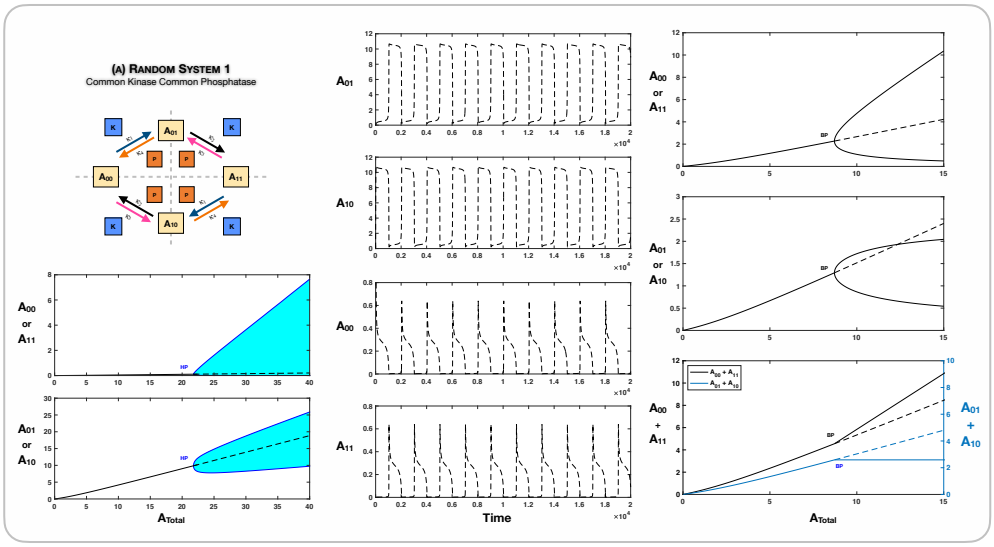



**Figure 2.6:** Case 3 symmetry breaking in random ordered DSP with common kinase and common phosphatase acting on each modification site. Column 1 shows the presence of oscillations emerging through a Hopf bifurcation in the bifurcation diagram along $A_{Total}$. Column 2 shows long period oscillations in the system represented in Figure 2.3(A) from the main text (for a $A_{Total} = 12.95$). Such long period oscillations emerge when the oscillatory branch from the Hopf bifurcation approaches asymmetric stable steady states. Column 3 shows the presence of symmetry breaking through a supercritical Pitchfork bifurcation. Note that the sum of concentrations of the partial substrates is conserved in the asymmetric steady states. This network is also capable of symmetry breaking through a subcritical pitchfork as seen in fig. 2.3. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during such oscillations. BP: (branch point) Pitchfork bifurcation, HP: Hopf bifurcation]

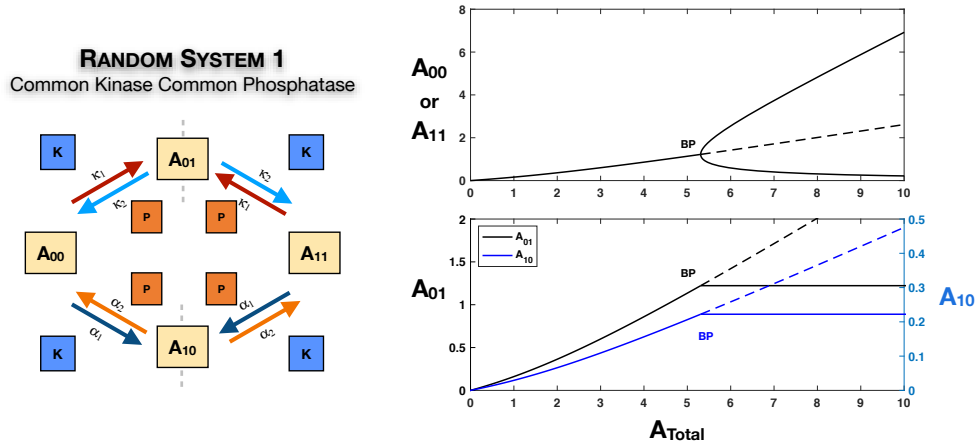



**Figure 2.7:** Case 1 symmetry breaking in the random ordered DSP with common kinase and common phosphatase acting on each modification site, even when one of the legs ($A_{00} \Longleftrightarrow A_{10} \Longleftrightarrow A_{11}$) is incapable of breaking symmetry by itself (i.e. viewed as an ordered mechanism the kinetic parameters of the $A_{00} \Longleftrightarrow A_{10} \Longleftrightarrow A_{11}$ leg forbid independent symmetry breaking - see main text and analytical work for discussion). Homeostasis (Absolute Concentration Robustness) is observed in both partial substrates. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation]

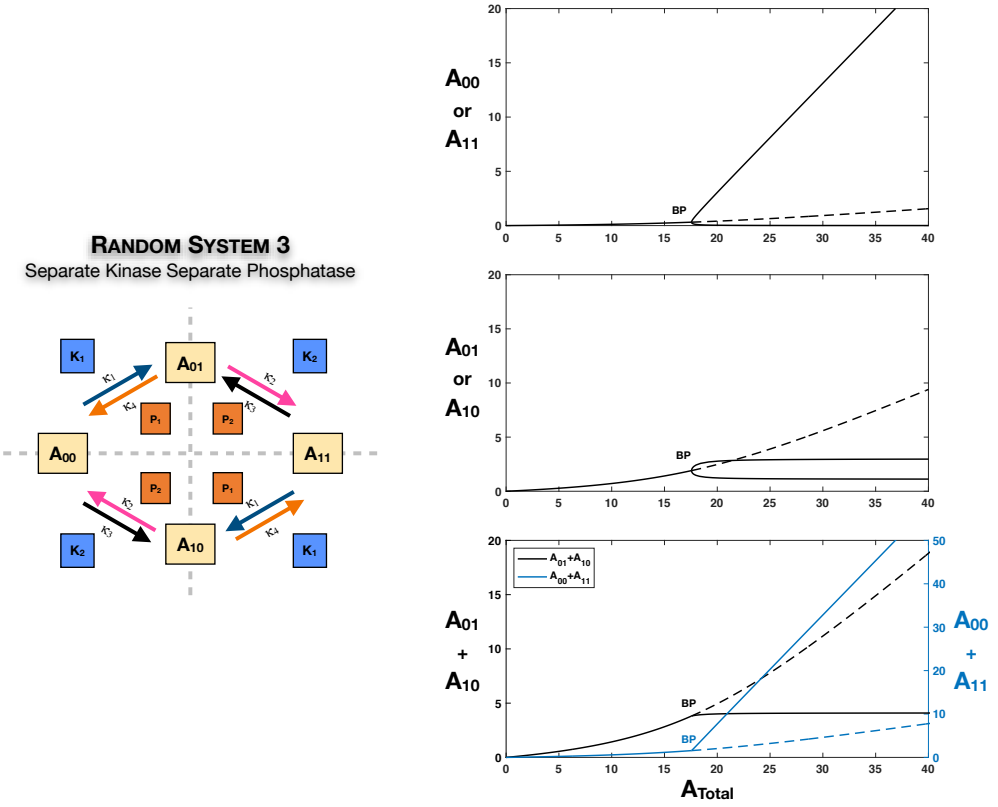



**Figure 2.8:** Case 3 symmetry breaking through a supercritical pitchfork bifurcation in the random ordered DSP with separate kinase and separate phosphatase acting on each modification site. Note that the sum of the concentrations of the partially modified substrates is approximately fixed in the asymmetric steady states in the bifurcation diagram. However, case 3 symmetry breaking in this network is also capable of providing approximate concentration robustness in the sum of the concentrations of the completely modified and unmodified substrate forms - see main text and fig. 2.3. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation]

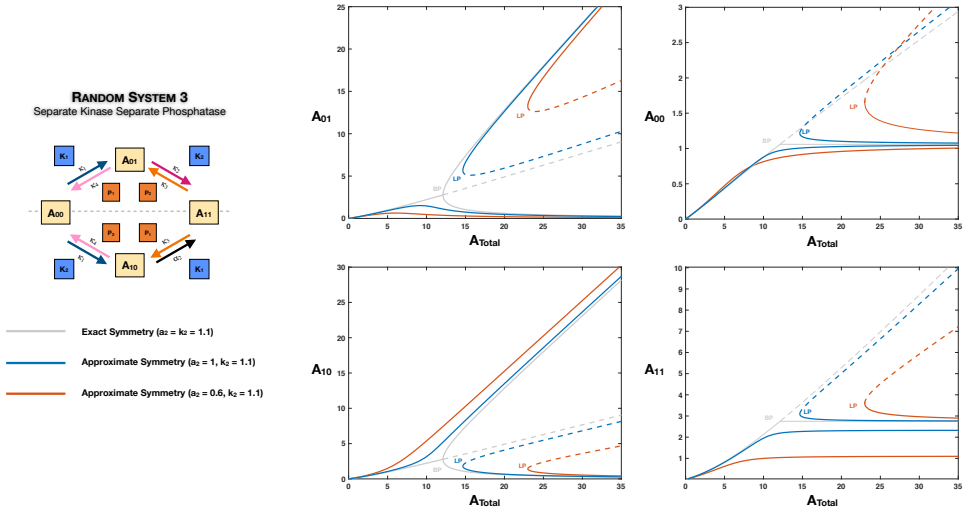



**Figure 2.9: Features of symmetry and symmetry breaking can still manifest when the network is only approximately symmetric.** This is represented through the example of case 2 symmetry breaking in the distributive ordered DSP with separate kinase and separate phosphatase affecting phosphorylation and dephosphorylation respectively. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation; LP: (limit point) Saddle Node bifurcation]

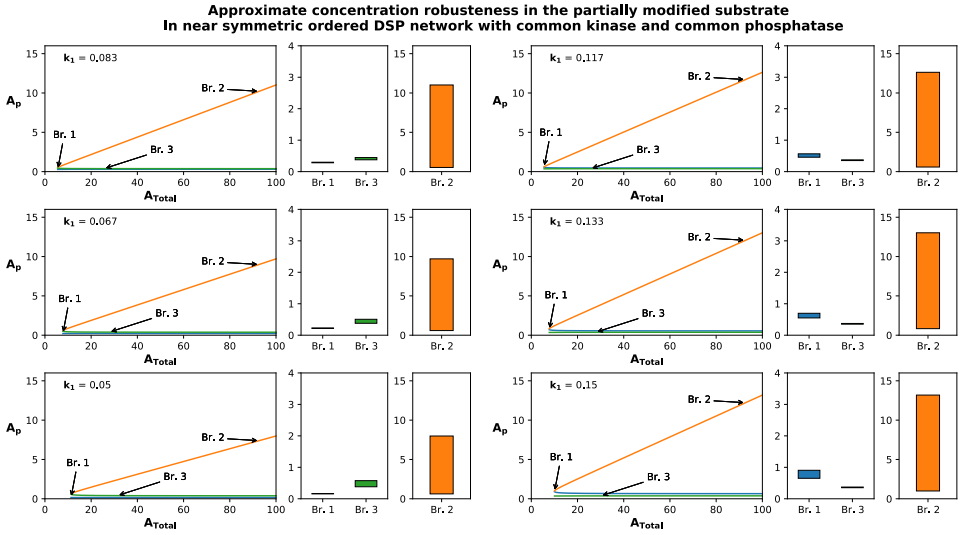



**Figure 2.10: Approximate concentration robustness shown by systems that deviate from exact symmetry**: Figure represents the approximate concentration robustness shown by $A_p$ in the ordered distributive DSP with common kinase and common phosphatase when the network is not symmetric. Note that the system is symmetric when $k_1 = 1$ and is presented in fig. 2.2A. In order to ascertain the behavior of the system and in particular ACR characteristics of $A_p$, we perturb the kinetics ($k_1$) at 6 values between $50\%$ and $150\%$ of the symmetric value (while keeping all other kinetics fixed) and present the result in 6 panels composed of three plots each; The first represents a bifurcation diagram showing the presence of steady state branches in a range of $A_{Total}$ (0 to 100) where multistability is present. The other two are bar graphs representing the norm of the concentration (Max value - Min value) of $A_p$ on the branches 1 and 3, and branch 2 respectively, across the entire range of variation ($A_{Total} = 0 - 100$). Note that in the symmetric network we obtain a perfect pitchfork bifurcation, however when the system deviates from exact symmetry, multistability is obtained through a saddle node bifurcation as shown here and in Appendix 2 - fig. 2.9. The norm of $A_p$ is significantly smaller in magnitude on branches 1 and branches 3 (which would be the perturbed analogues of the ACR branches in the perfectly symmetric system) as compared to the norm in branch 2 (the perturbed analogue of the symmetric branch). The norm on branch 2 scales with increasing range of $A_{Total}$ (Note: the norm shown here is only for the range of $A_{Total} = 0 - 100$) while the norm of branches 1 and 3 are varying negligibly with changing $A_{Total}$. This behavior is depicted for 6 different parameter values between $50\%$ to $150\%$ and is again indicative of approximate concentration robustness seen in near symmetric systems for a large range of total substrate concentration.

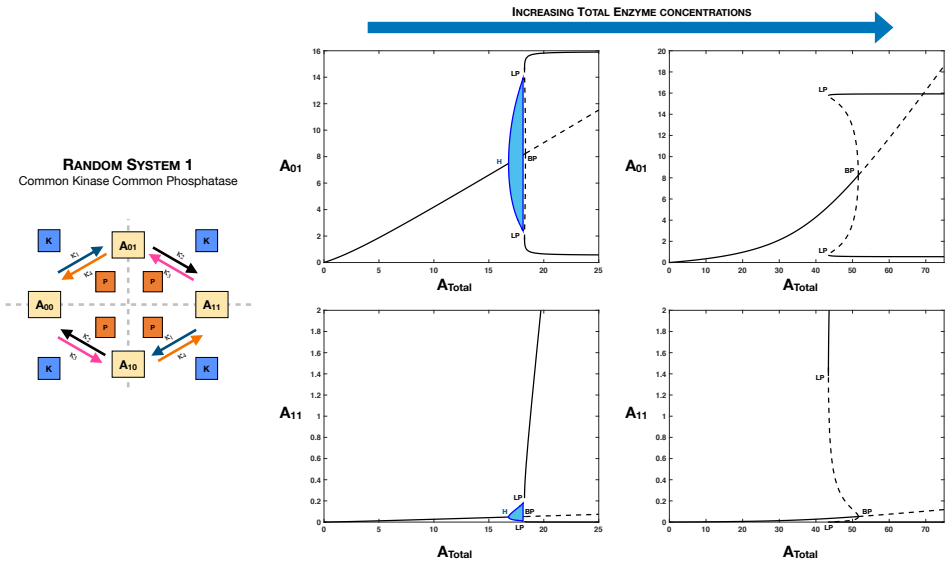



**Figure 2.11: Total enzyme concentrations are an additional lever (apart from basic network kinetics) that can independently tune symmetry breaking behavior in MSP networks.** This is represented through the example of case 3 symmetry breaking in the distributive random DSP with common kinase and common phosphatase). Panel 2 shows how increasing enzyme concentrations (left $\longrightarrow$ right) can lead to loss of oscillatory behavior in the network, for the same basal kinetics parameters.

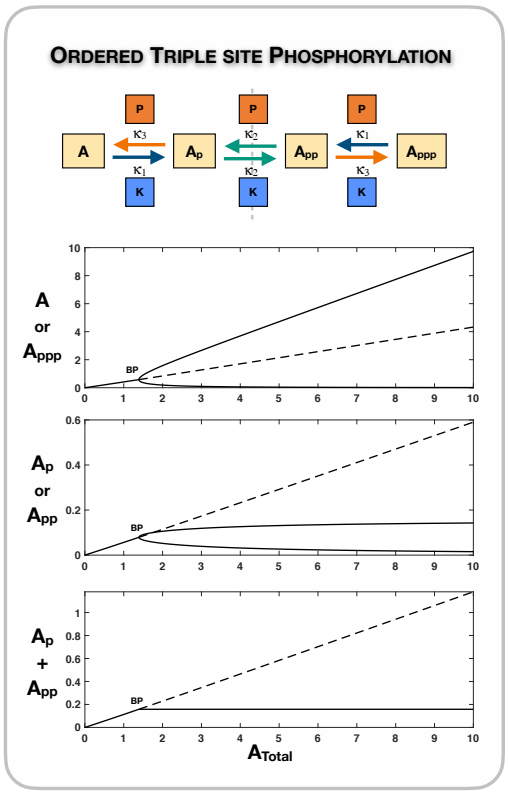



**Figure 2.12:** Case 1 symmetry breaking in the distributive ordered TSP network. Here we observe Absolute Concentration Robustness in the sum of the concentrations of the partially modified substrates. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation]

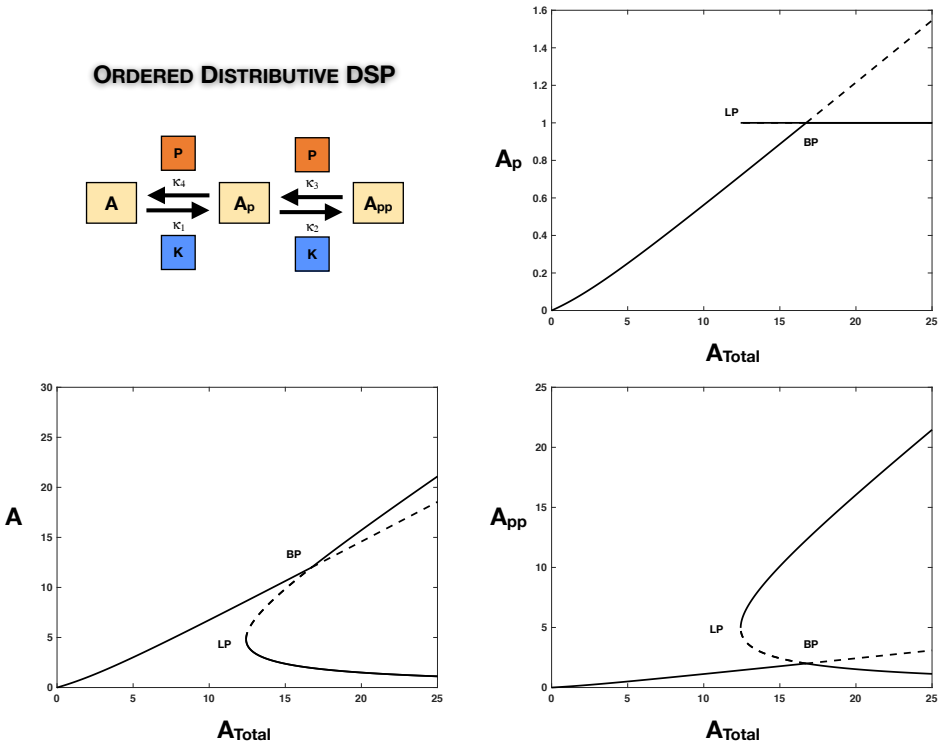



**Figure 2.13: Absolute Concentration Robustness in ordered DSP**: The ordered double site with common kinase common phosphatase is capable of exhibiting (exact) Absolute Concentration Robustness (ACR) in the partially modified substrate ($A_p$) with respect to changing total substrate concentration even in the absence of symmetry in the kinetics or total enzyme amounts (However a weaker constraint is required to enable this). The figure presents a computational example of this (complementing the discussion in the main text and the appendix 1). We observe a single non-ACR exhibiting branch of steady states exchanging stability with branches of ACR exhibiting steady states through a transcritical bifurcation (as opposed to a pitchfork bifurcation as seen in the symmetric examples earlier). The concentration of $A_p$ is fixed to be mathematically exact on these ACR branches as shown in the top-right plot. The unstable ACR branch emerging out of the transcritical bifurcation becomes stable through a saddle node bifurcation as shown. [Dotted lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. BP: (branch point) Pitchfork bifurcation; LP: (limit point) Saddle Node bifurcation]

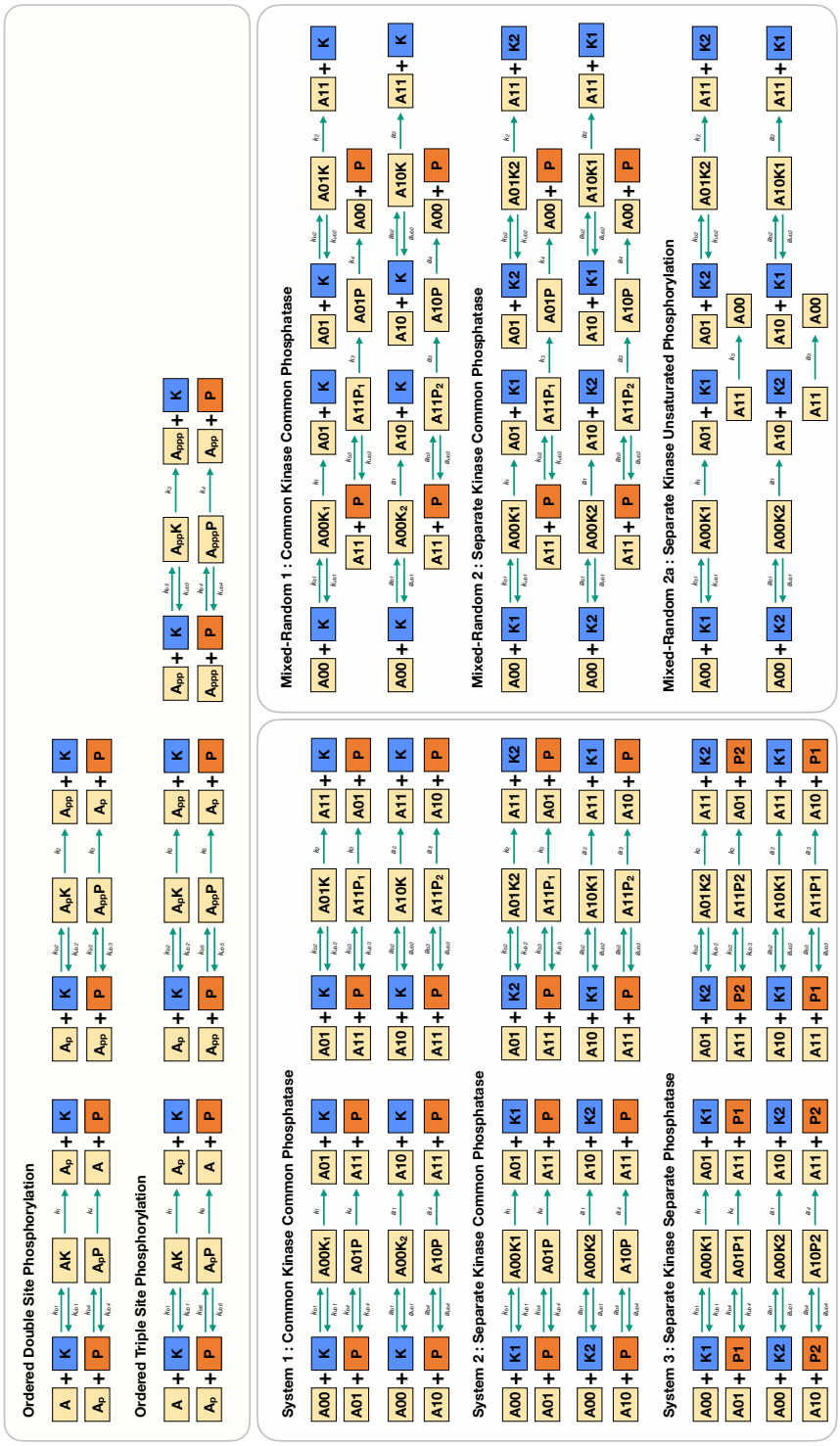



**Figure 2.14: Detailed model description of the various MSP networks used in this paper.** The constituent binding, unbinding and catalytic reactions of each modification step is described in detail and are modelled using mass kinetic description.

## Executive summary of analytical results and resources

**Table 2.1:** Executive Summary - Chapter 2: Symmetry and symmetry breaking in multisite modification

| Model | Case | S. br. | Brief description of analytical result | | Ex. App. | Computation |
|---|---|---|---|---|---|---|
| Ordered DSP (CC) | 1 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. of $A_p$ in S. br. states | C(2.1-2) | 2.2, 2.10 and 2.13 |
| Ordered TSP (CC) | 1 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in sum of $A_p$ + $A_{pp}$ in S. br. states | C(2.3) | 2.12 |
| Random System 1 | 1 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in $A_{01}$ & $A_{10}$ in S. br. states | C(3.1) | 2.2 and 2.7 |
| | 2 | No | Proof of absence of symmetry breaking | | | - |
| | 3 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in sum of $A_{01}$ + $A_{10}$ in S. br. states | | 2.3, 2.6 and 2.11 |
| Random System 2 | 2 | No | Proof of absence of symmetry breaking | | C(3.2) | - |
| Random System 3 | 1 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in $A_{01}$ & $A_{10}$ in S. br. states | C(3.3) | 2.2 |
| | 2 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in $A_{00}$ & $A_{11}$ in S. br. states | | 2.2 and 2.9 |
| | 3 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in sum of $A_{01}$ + $A_{10}$ in S. br. states | | 2.3 and 2.8 |
| Mixed Random 1 | 2 | No | Proof of absence of symmetry breaking | | C(4.1) | - |
| Mixed Random 2 | 2 | Yes | Obtain necc. and suff. cond. for S. br. | Show presence and obtain ACR conc. in $A_{00}$ & $A_{11}$ | C(4.2) | 2.2 |
| | 2 | Yes | Show similar results in system operating with saturated dephosphorylation (Mixed Random) 2a | | C(4.3) | 2.5 |

# Chapter 3

# Biphasic responses in cell signalling: A unified approach

## 3.1 Introduction

The enzymatic modification of substrates is the basic building block of cell signalling and cellular biochemical networks [121, 117, 31]. The basic intuition from the analysis of these systems is that increasing both the concentration of the enzyme performing the modification, and total concentration of the substrate increases the steady state concentration of the modified substrate.

However observing both pathways and networks which are comprised of these building blocks indicates that there may be significant divergence from this basic behaviour. One of the most widespread ways in which this may manifest is through a biphasic dose response [67, 122, 123, 124, 125]. Far from being an accidental occurrence, biphasic dose responses have many advantages, including a built-in attenuation mechanism, a cap on the maximal level of output as well as the presence of a zone of "optimal" activation. This is borne out by the fact that these very same implications have important consequences in different contexts.

A typical way in which biphasic responses in cell signalling networks are studied is to investigate network motifs which generate such behaviour [126, 127, 128]. In this chapter, by contrast, we investigate the sources of biphasic dose responses, by starting with the basic building blocks of substrate modification. By investigating a series of basic building blocks, we reveal the potential for basic modifications to intrinsically give rise to biphasic dose responses, either in response to enzyme concentration or substrate concentration variation. In so doing we obtain a unified synthesis of biphasic responses in response to both enzyme and substrate variation. We then build on this to investigate how these sources of biphasic responses intrinsic to chemical modification

affect network behaviour. The study is rounded out by an analysis of a concrete exemplar system (*ERK* signalling), leading to a re-evaluation of existing hypotheses and assumptions complemented by a unified synthesis of biphasic responses here.

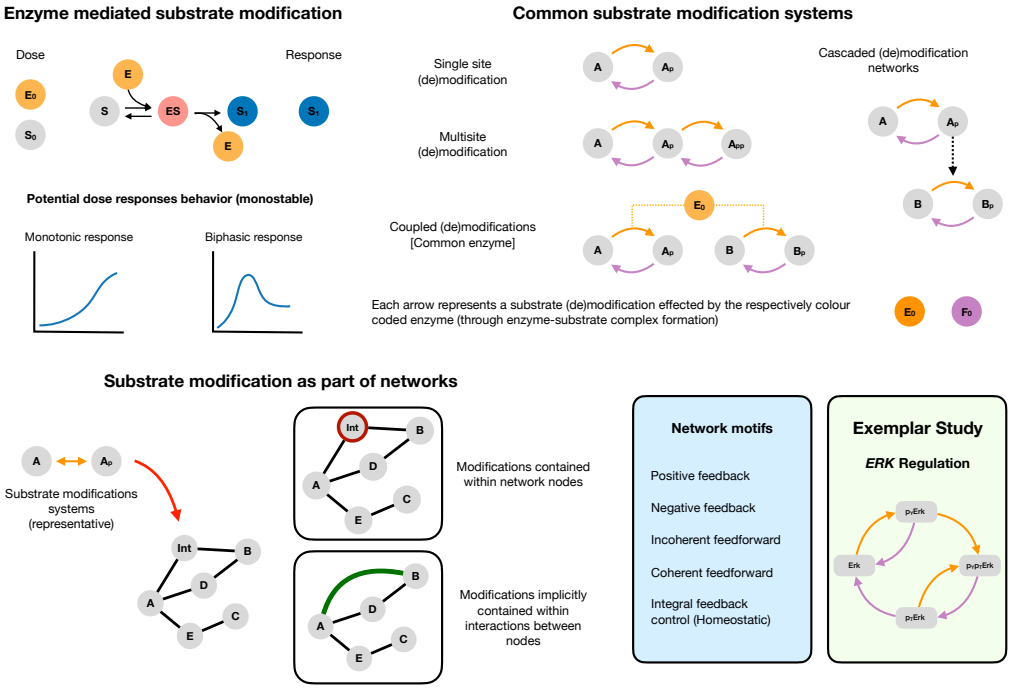



**Figure 3.1: Enzyme mediated substrate modification: Snapshot - from modification mechanism to common substrate modification systems and their role in enabling network regulation.** Enzyme mediated substrate modification features an integral part of cellular signalling and is among its basic building blocks. *Top left* shows one mechanism of distributive substrate modification by an enzyme involving the following steps: the binding (and unbinding) of the substrate ($S$) with the enzyme ($E$) to form a substrate-enzyme complex ($ES$), which finally disassociates to give the modified substrate ($S_1$) and the free enzyme ($E$). Substrate modification has two natural doses for the response (modified substrate $S_1$) - total substrate amount and total enzyme amount ($S_0$ and $E_0$ respectively). The potential dose-response behavior in the simplest instance can thus be either a) monotonic or b) biphasic dependence of the modified substrate with the dose. Both substrate and enzyme doses can potentially present biphasic responses. *Top right* shows the schematics of commonly encountered substrate modification systems in cell signalling. Each arrow indicates the triplet of binding, unbinding and catalytic action by the enzyme. Arrows of the same colour indicate action by a common enzyme type (a kinase and a phosphatase for example). These systems form key components of wider signalling networks in the cell. From a network perspective they can be incorporated in two ways. 1. They can either be inherent to a node in such networks or 2. be implicitly present within an interaction between nodes (arrow, which can be the sum of actions of nodes not explicitly defined). This is depicted in the *bottom row*. A fruitful way the role of enzymatic modification within networks is through the study of recurring signalling motifs (with characteristic input-output responses) such as feedback and feedforward structures. *D.* lists the canonical network motifs (positive and negative feedback, and coherent and incoherent feedforward, integral feedback control) that we study in our work. Further we also consolidate our insights by studying the role of enzymatic modification in enabling biphasic responses in the *ERK* signaling pathway. Detailed mathematical model descriptions of the different substrate modification motifs and network regulation motifs studied are are discussed in the supplementary material, and elaborate reaction schematic of the models considered are shown in fig. 3.9.

## 3.2 Models and Methods

**Networks considered.** In this chapter we consider various basic networks encompassing enzyme mediated protein modification. These networks include single site (de) modification cycles, prototypical examples of double site modification (DSP) networks (ordered distributive), and cascaded networks. We also explore the role of sharing of enzymes, and multi-functionality of enzymes within these systems and as such consider associated network structures such as coupled modification cycles and multiplicities of enzyme actions in the DSP network. Put together, these systems (models M2-M8) represent the basic building blocks of enzymatic networks and are encountered in multiple instances of bio-chemical reactions and pathways.

Building on this, we also consider simple signaling motifs (feedforward and feedback), with biphasic interaction between the nodes of the network motif (models N0, N3-N5) to explore the role of such interactions on expected network motif response. Finally, we also consider complex signal regulation of the DSP with common kinase and common phosphatase (models N1-N2), with biphasic signal regulation activating the kinase and incoherent feedforward regulation of the DSP by a signal. These models allow us to explore the role of biphasic signal regulation and interaction between nodes in networks on the emergent systems level behavior. We finally focus on the *ERK* regulation pathway, and it's propensity to register biphasic dose response (model adopted from [123, 75]). A summary schematic of the various models used is presented in fig. 3.8.

**Mathematical modelling.** The networks considered above are modelled as a system of ordinary differential equations using simple mass action kinetic description for the elementary reactions. Each (de)modification is effected by three elementary reactions, the reversible binding (and unbinding) of the free active enzyme with the substrate to form an enzyme substrate complex, followed by the irreversible catalytic reaction to yield the free active enzyme and the (de)modified substrate (see fig. 3.1). Using such a description for the elementary (de)modification mechanism, the mathematical model for the networks are constructed. Since the total amount of substrate and enzymes are also conserved in the system, the ODEs are also associated with the respective conservation equations.

We present below how the covalent modification cycle (where a substrate gets modified and demodified by a kinase and a phosphatase; model M0) can be described mathematically using such a description.

$$
\begin{aligned}
\frac{d[A]}{dt} &= -k_{b1}[A][K] + k_{ub1}[AK] + k_2[A_pP] \\
\frac{d[AK]}{dt} &= k_{b1}[AK] - (k_{ub1} + k_1)[AK] \\
\frac{d[A_p]}{dt} &= k_1[AK] - k_{b2}[A_p][P] + k_{ub2}[A_pP] \\
\frac{d[A_pP]}{dt} &= k_{b2}[A_p][P] - (k_{ub2} + k_2)[A_pP] \\
\frac{d[K]}{dt} &= -k_{b1}[A][K] + (k_{ub1} + k_1)[AK] \\
\frac{d[P]}{dt} &= -k_{b2}[A_p][P] + (k_{ub2} + k_2)[A_pP]
\end{aligned}
\qquad
\begin{aligned}
A_{Total} &= [A] + [A_p] + [AK] + [A_pP] \\
K_{Total} &= [K] + [AK] \\
P_{Total} &= [P] + [A_pP]
\end{aligned}
$$

The mathematical models for the other networks considered are constructed in a similar manner. In using simple mass action kinetics to represent the kinetics of the (de) modifications, we make no a priori assumption regarding the kinetic regime of the enzyme action, and thus the results are truly representative of the overall behavior of the system. The models are cross-validated with model descriptions generated in *COPASI* [129] (which only requires definition of the reaction schematic for model generation).

*Modelling biphasic interactions between nodes in a network.* The biphasic interactions between nodes $A \to B$ in a network, (where more A implies lower activation of B beyond a certain threshold) was modelled using a Poisson function using the following mathematical expression, $\frac{d[B]}{dt} = k_1[A]e^{-k_2[A]}$. This is in contrast to the direct mass action regulation of $A \to B$, which is modelled as $\frac{d[B]}{dt} = k_1[A]$. The parameter values (for $k_1$ and $k_2$) are chosen such that the range of the interaction realized through either the Poisson (biphasic description) or the linear regulation are similar when averaged over range of bifurcation.

When choosing parameters particularly in the context of systems with biphasic response in the interaction, in order to maintain parity and a likewise comparison between the system with the response and without, we ensure that the relative strengths of the interaction terms remains the same when averaged over the range of total signal activation. Further details on how this is pursued is provided in appendix C.

**Parameters.** The models constructed have numerous parameters, involving both kinetic constants and total amounts of substrate and enzymes. It is useful to draw distinction between these two types for two reasons. 1. The total amounts represent easily accessible levers that can be tuned, while the kinetics are broadly intrinsic to the system. 2. In the context of biphasic dose responses, the total amounts represent doses to the system and thus are natural levers that change, and along which bifurcation analysis are carried out.

**Doses.** Enzyme mediated protein modifications have two natural doses (with respect to the maximally modified substrate form), i.e. 1, total substrate amount, and 2. total enzyme amount.

**Approach.** The approach to assessing biphasic response in the networks is as follows.

We begin by analytically characterizing the presence or absence of the biphasic behavior for the relevant substrate concentration with the dose, and in the process obtain binary answers to the feasibility of observing the behavior in parameter space for a given model. In the instance where the behavior is shown to be present, we show evidence of it computationally with a 1D bifurcation analysis with the relevant dose. We further supplement this with analytical work to establish sufficient and necessary conditions (involving parameters, such as total amounts and kinetic constants) that guarantee the presence of biphasic dose response behavior and conversely conditions that preclude the behavior in parameter space.

*Bifurcation analysis.* 1D bifurcation analysis was carried out computationally using the *MatCont* package in *MATLAB* [72]. In all cases showing the presence of biphasic response, the result is accompanied with analytical work establishing the presence of the behavior in vast regions of the parameter space (see below).

*Analytical approach to assessing biphasic response.* Analytical work exploring the presence or absence of biphasic dose response (for a given choice of kinetic parameters and total amounts) relies on the mathematical requirement for a steady state of the system to satisfy the following expression at some finite concentration of dose $\frac{d[Var]}{d[Dose]} = 0$, should a biphasic response exist. Analytical work proceeds by exploring the models' propensity to satisfying this condition for some feasible steady state of the system. This is done by solving the system of ODEs of the model at steady state and ascertaining an expression for evaluating the above expression ($\frac{d[Var]}{d[Dose]}$) as a function of the steady state concentration of fewer variables and the parameters. This reduction allows us to ascertain whether the behavior is possible/impossible across parameter space, and where possible ascertain necessary and sufficient conditions guaranteeing (and conversely precluding) the behavior. The entirety of this analysis is carried out in the computational platform *Maple* [73], and is presented as a PDF version in appendix C.

Note: Our analysis through out relies on the existence of a biphasic dose response in the steady state of the system of the dose. We however make no comment or further analysis on the stability of the steady state when such a biphasic response is predicted. This however can be easily verified with computations and we supplement each of our analytical results with computational proof of the existence of the behavior for a robust range of parameters.

## 3.3 Results

Enzyme mediated protein (substrate) modification has two natural doses in total amounts (for the modified substrate form); Two natural doses arise when considering dose-responses for enzyme-mediated protein (substrate) modification: 1. The total enzyme amount, and 2. The total substrate amount. In this chapter, using the exemplar of

phosphorylation for such protein modifications, we study the capacity of networks involving such enzyme mediated protein modification to show biphasic dose responses in the concentration of the (maximally) modified substrate form with each of these doses. The biphasic dose responses in the (maximally) modified substrate with increasing total amounts of substrate and enzyme are henceforth called substrate biphasic and enzyme biphasic responses.

The results are organized in the following manner. We first analyze the capacity of simple modules of enzymatic substrate modification (such as the covalent modification cycle and multisite substrate modification) to exhibit biphasic dose response. Following this, we analyze the consequences of such dose responses for the network behavior by analyzing simple network motifs. Finally we study the propensity for and consequences of such biphasic responses in the *ERK* regulation pathway, which represents a concrete and well characterized cellular pathway shown to present biphasic dose response behavior [108].

In each of these cases, we use analytical work (see models and methods) to ascertain the presence or absence of biphasic behavior in the model. The presence of a biphasic response is demonstrated computationally by performing a 1-parameter bifurcation analysis (varying the relevant dose: total amounts of substrates or enzymes). Analytical work is further used to reveal the conditions (intrinsic kinetic parameters and total amounts of substrates and enzymes) under which a given system can exhibit biphasic dose responses.

**The covalent modification cycle with additional complex formation is capable of enzyme biphasic responses.** The simplest enzymatic protein modification system, the covalent modification cycle (model M0) is incapable of exhibiting biphasic responses with either increasing total enzyme amounts or total substrate amounts (see appendix C, section 1). However an additional complex formation of the modified substrate with the kinase enzyme (model M1), allows for the network to exhibit enzyme biphasic responses (see fig. 3.1A). The capacity of this system to exhibit enzyme biphasic dose responses has been studied earlier (albeit with different modelling assumptions by Varusai et al., [128]). Interestingly, enzyme biphasic responses may be obtained for any choice of intrinsic kinetic parameters. However this system is incapable of exhibiting substrate biphasic responses (see appendix C, section 1).

**Multisite modification and commonality of enzyme action promotes biphasic dose responses.** Proteins rarely undergo just one (de)modification and often are (de)modified at different sites (effected by either the same enzyme or by different enzymes). To explore the consequence of such multiplicities in modifications, we consider the simplest extension of a basic covalent modification cycle, with a substrate undergoing two modifications in an ordered fashion, where the order of phosphorylation is opposite that of dephosphorylation and the (de)modifications effected by the same or different enzymes (see models M2-M5 in fig. 3.7). In the instance where a common enzyme effects

multiple modifications, the mechanism is assumed to be distributive.

We observe that the double site modification (DSP) network with common kinase and common phosphatase effecting the modifications (Model M2) is capable of exhibiting both substrate and enzyme biphasic responses in the maximally modified substrate, $A_{pp}$ (see fig. 3.2B) [122]. Analytical work reveals the parametric dependence for this behavior to be observed. Substrate biphasic responses are guaranteed to be observed for some total amount of kinase and phosphatase for any set of intrinsic kinetic parameters of the system. On the other hand, enzyme biphasic responses can only be present when the following condition ($k_1k_3 - k_2k_4 > 0$) is satisfied by the kinetic catalytic constants ($k_1, k_2, k_3$ and $k_4$: $k_1$, $k_2$ associated with phosphorylation, $k_3$, $k_4$ associated with dephosphorylation) of the system. When the condition (which is necessary and sufficient for the behavior) is not satisfied the system is incapable of exhibiting enzyme biphasic responses (see appendix C, section 2).

It is interesting to note that this kinetic condition involving the catalytic constants is also observed elsewhere in studies relating to the presence of multistability in the DSP with common enzymes. Conradi et al., [74] identified that if the condition is violated, then the network is guaranteed to exhibit multistability at some total amounts of substrate and enzyme. These results together suggest that irrespective of the kinetic regime of the system, the DSP with common enzymes is capable of exhibiting either multistability or biphasic dose response with enzyme necessarily (for some total amounts of proteins). This reveals the plurality of information processing behaviors capable with substrate modification systems with only two modification sites, and how catalytic constants (and more broadly intrinsic kinetic constants) can allow us to guarantee and generate these behaviors in such systems.

**Additional specificity of enzymes in DSP reduces capability of biphasic responses.** An additional source of complexity in enzyme mediated substrate modification arises from the specificity of enzyme action. The same enzyme could effect modifications on different proteins or in multisite modification, there could be distinct enzymes effecting each (de)modification. Considering such scenarios with different enzymes effecting each (de)modification in double-site modification (see models M3-M5 in fig. 3.6), we notice that the capacity of the DSP motif to exhibit biphasic dose response disappears, with increasing specificity in enzyme action. The DSP with separate kinases and a common phophatase (model M4) acting on each modification site is capable of only exhibiting enzyme biphasic responses, and only with the variation of total amount of the kinase effecting the second modification (see fig. 3.2D). On the other hand, it is incapable of exhibiting substrate biphasic dose response irrespective of kinetic parameters. Similarly, the DSP with a common kinases and different phosphatases effecting each modification (model M3) is capable of exhibiting substrate biphasic responses (see fig. 3.2C), while it is incapable of exhibiting enzyme biphasic responses with changing total kinase amount. Analytical work further reveals that substrate biphasic responses in the model is guaranteed to be observed for any choice of underlying kinetic param-

eter choice, for some total amounts of kinase and phosphatase. The DSP with separate kinases and phosphatases effecting (de)modifications (model M5) is incapable of exhibiting either substrate or enzyme biphasic (see appendix C, section 3).

Taken together, this shows how complexity at the level of the biochemistry of a node in a signalling network (arising from number of modifications, and enzyme sharing) can enable biphasic dose responses in the (maximally modified) substrate form. Sequestration and the commonality of enzymes are key features that promote biphasic dose response (both enzyme and substrate).

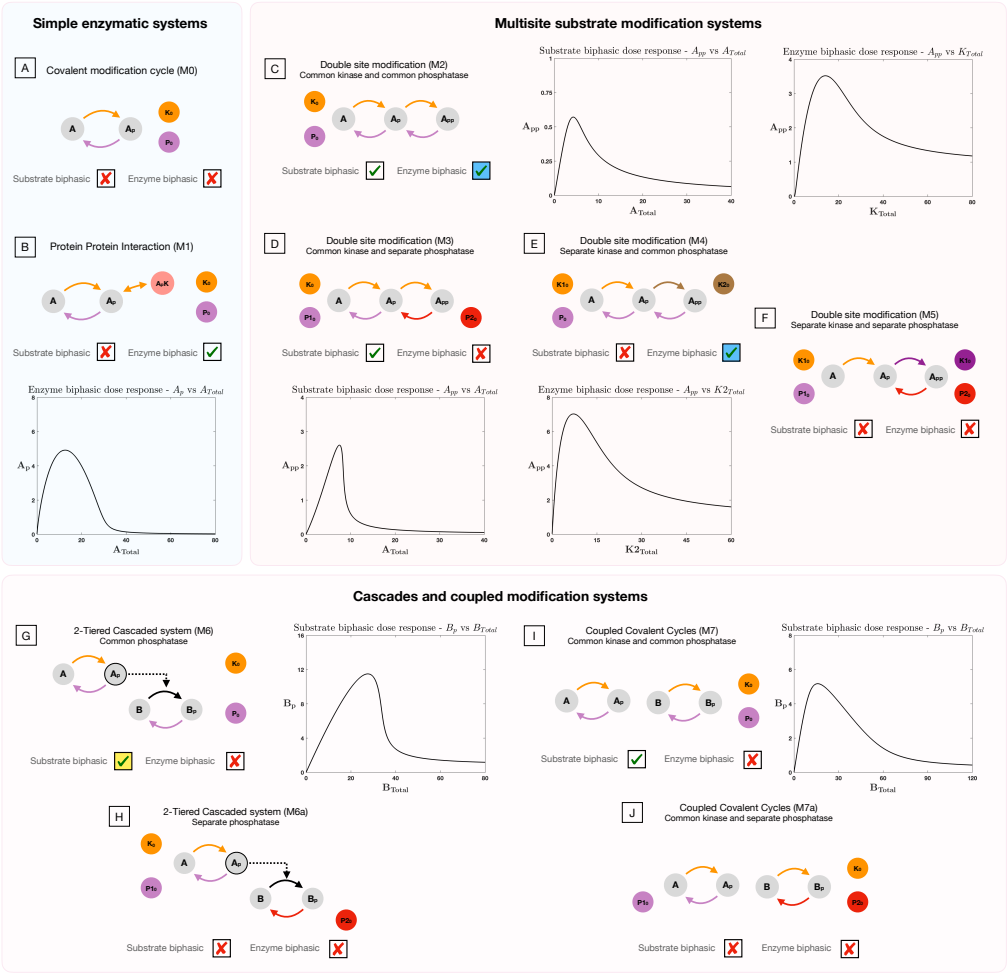



**Figure 3.2: Presence and absence of substrate and enzyme biphasic dose responses in the commonly observed enzymatic cellular signaling systems.** Looking across multiple systems (including combinations of enzyme action) and different doses allows us to clearly ascertain the origin and necessary features of enzymatic modification systems to present both enzyme and substrate biphasic dose-responses. Where results denote a biphasic response is absent - it is absent irrespective of kinetic regime, parameter values and total amount so substrate and enzyme. This is established through analytical work (see supplementary file 1). Where a biphasic dose-response is present, it is shown with a bifurcation diagram, with changes to the relevant dose. The presence of specific dose-response can be characterized in parameter space in the following way, either the behavior is 1. present for *all kinetic regimes* (accessible for some total amounts of substrate and enzyme, transparent boxes), 2. present only for *specific kinetic regimes* (accessible for some total amounts of substrate and enzymes, blue shaded boxes) - see text and analytical work for more details. A. *Covalent modification system*. Absence of substrate and enzyme biphasic response. B. *Protein Protein Interaction*. Enzyme biphasic dose response in the protein-protein interaction model (modelled using mass action kinetics and no assumptions regarding kinetic regime of enzyme action). In contrast to the covalent modification cycle which is incapable of the behavior, this result indicates how a single additional complex formation by the enzyme (and resulting sequestration) can generate enzyme biphasic dose response. C. *Double site modification (DSP): common enzymes*. Presence of enzyme and substrate biphasic response (enzyme biphasic only in specific kinetic regimes - see text). D. *DSP: common kinase and separate phosphatase*. Presence and absence of substrate and enzyme biphasic response respectively.

Continued on next page...

E. *DSP: separate kinase and common phosphatase*. Absence and presence of substrate and enzyme biphasic response respectively (enzyme biphasic only in specific kinetic regimes - see text). F. *DSP: separate kinase and separate phosphatase*. Absence of both substrate and enzyme biphasic response. These results together show how commonality of enzymes promotes biphasic responses with both doses, and in particular how commonality in phosphatase and kinase action enables substrate and enzyme biphasic response. G, H. *Two tiered cascaded systems with common and separate phosphatases respectively*. Presence of substrate biphasic in the second tier substrate of the two-tier cascaded enzymatic modification system with a common phosphatase. The first tier substrate is incapable of substrate biphasic in the same model. Both tier substrates are incapable of substrate biphasic with separate phosphatases, and absence of enzyme biphasic in both systems. I, J. *Coupled covalent cycles with common and separate phosphatases respectively*. Presence and absence of substrate and enzyme biphasic in the system with common phosphatases. The system with separate phosphatases is incapable of substrate and enzyme biphasic. This indicates how simple features of enzyme mediated modification such as coupling and substrate functionality can enable motifs to present with biphasic dose response, even when their constituent sub-modules (covalent modification cycles) are incapable of it. *Note*: These results are also complemented with analytical work characterizing the parametric dependence on observing biphasic response - see fig. 3.4 and appendix C for further details

**Beyond multisite regulation.** Modifications provide additional functionality to proteins, and in instances, a protein undergoing a modification can gain functionality that allows it to effect modifications on another protein [130]. This generates a cascaded network structure of modifications (such as the *MAPK* system [100, 131, 83, 27]). Here too, common/separate enzymes may effect the demodifications of each of the substrates.

To explore the possibility of biphasic responses in such enzymatic modification cascades, we studied the two tier enzymatic cascade with the demodifications being effected by either common or separate enzymes (M6 and M6a respectively, see fig. 3.7).

This enzymatic cascade with separate phosphatases is incapable of exhibiting either substrate biphasic responses (in response to the variation in total substrate amounts in either tier) or enzyme biphasic responses in response to change in total amounts of kinase. On the other hand, while the analogous system with common phosphatases effecting demodifications on each substrate is still incapable of enzyme biphasic responses or substrate biphasic responses with respect to the first tier substrate (see appendix C, section 4), it is capable of exhibiting substrate biphasic responses in the second tier substrate (see fig. 3.2E).

Other forms of coupling are possible with enzymatic modifications. For example, a single enzyme can effect modifications on multiple substrates allowing for a sophisticated degree of cross talk and coupling between networks [132, 133, 134, 135]. We studied the propensity for biphasic responses in the system with enzyme sharing between two covalent modification cycles. The system with a common kinase and common phosphatase for the cycles (model M7) can exhibit substrate biphasic responses in either of the substrate modification cycles (individually, see fig. 3.2F). The system is however incapable of exhibiting enzyme biphasic dose responses. Analytical work further reveals that the coupled covalent modification cycles are necessarily guaranteed to present substrate biphasic responses in e one of the modification cycles (but not both) for some

total amount of kinase, phosphatase and (other) substrate amount. The identity of the modification cycle which is capable of substrate biphasic responses is determined by an explicit analytical criterion which depends on intrinsic kinetic parameters.

The same system with separate enzymes acting on the two modification cycles (either kinase/phosphatase, model M7a or both) is incapable of exhibiting with substrate biphasic responses, showing how enzyme sharing and simultaneous coupling through the both enzymes is required for realizing substrate biphasic (see appendix C, section 4).

Interestingly, in all models capable of biphasic responses studied above, the systems which are capable of exhibiting multistability are also capable of exhibiting it alongside biphasic response in the dose response curve (see fig. 3.5).

**Biphasic interactions within signal regulation.** Biphasic dose responses are thus readily seen with only minimal ingredients (enzyme sharing, sequestration and multi-site modification) in cell signalling networks involving enzyme mediated protein modification. Cellular signal processing is largely dependent on networks involving proteins modifications and thus the consequence of biphasic response at a node within a network can fundamentally alter network response.

To explore this, we begin by analyzing how a simple biphasic signal can regulate a covalent modification cycle (model N0, see fig. 3.3A). This simple interplay of a signal regulation and biphasic transduction can convert a saturated response into a biphasic response with a range of prolonged homeostatic response in the concentration of the maximally modified substrate form, showing the potential for complicated signaling behaviors to emerge out of simple network considerations and biphasic interactions among nodes.

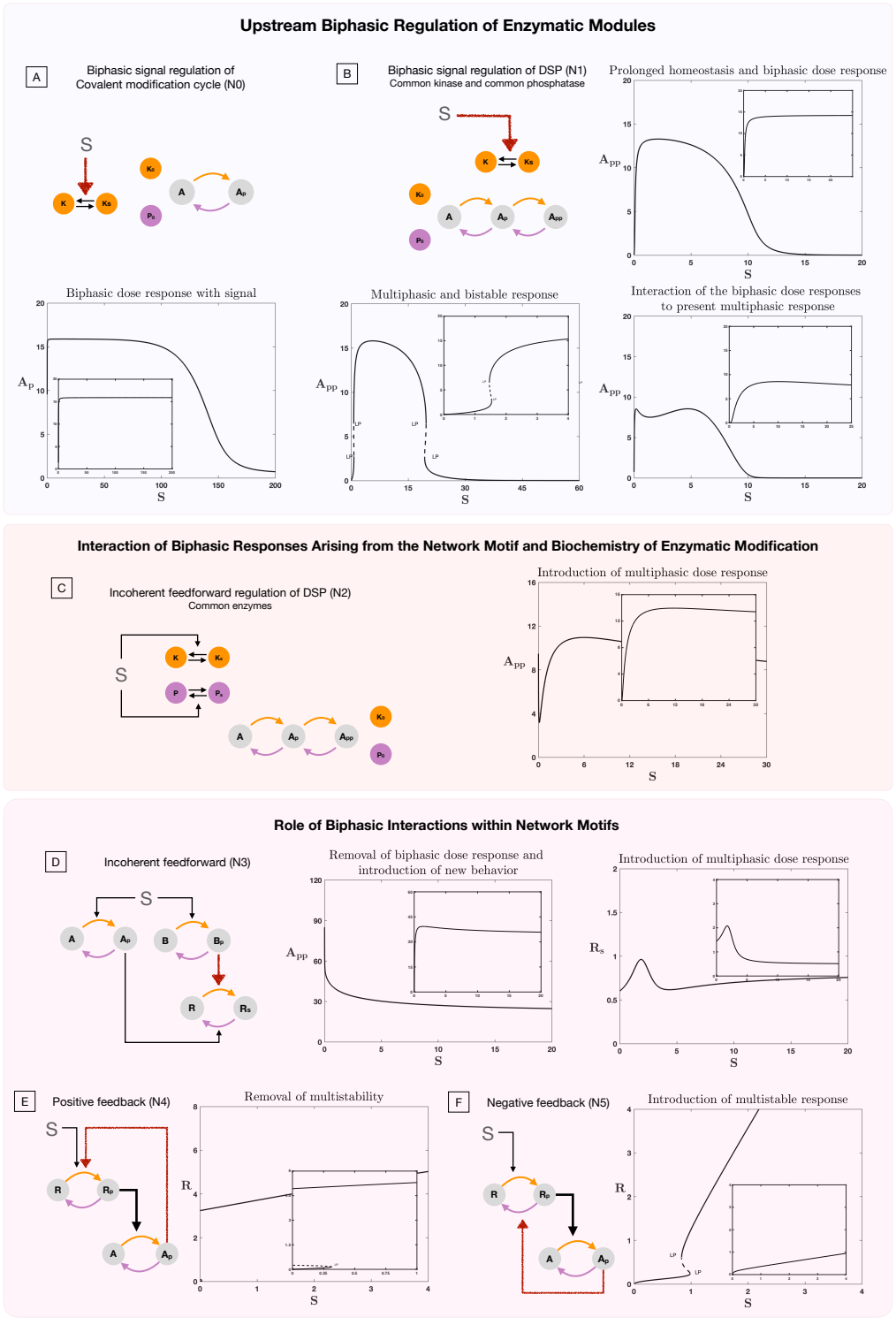



**Figure 3.3: Role of biphasic interactions in network and motif response.** The results presented in this plot show the consequence of biphasic signal regulation and biphasic response within interactions in enabling network response. In each plot the inset image represents the basal behavior without biphasic signal regulation for the same kinetic parameters of the system.

Continued on next page...

*(A-B.)* Role of biphasic signal regulation in perturbing behavior of enzymatic systems, *A.* Covalent modification cycle, and *B.* DSP (common enzymes). *A* shows how biphasic dose responses can enable homeostatic and biphasic response from a covalent modification cycle. *B* shows how biphasic signal regulation can enable multiphasic behavior, and realize distinct regions of biphasic and multistable regimes with signal. *C* contrasts this with the response from incoherent feedforward regulation of the DSP (a driver of biphasic responses), which can similarly present multiphasic response. (inset plot shows behavior without incoherent feedforward, i.e. absent activation of the phosphatase). *(D-F.)* **Effect of biphasic interaction is affecting common network regulation motifs**. Biphasic interaction can alter the fundamental feature of feedforward motifs in multiple ways. *D.* In feedforward networks (incoherent), it can destroy expected biphasic response and also enable the creation of multiphasic response. Similarly, in feedback networks, it can destroy multi-stability (*E*) and homeostatic responses (*F*). [Brown arrows in network schematics represent biphasic interactions. LP: saddle node bifurcation, solid lines and dotted lines denote stable and unstable steady state respectively]

**Impact on network motifs.** To explore this interplay further, we consider standard signaling motifs (feedback and feedforward) with biphasic interactions between the nodes (see models N3-N5 in fig. 3.6) and study how the biphasic interaction can alter expected network behavior. Networks motifs, such as feedforward and feedback networks, are recurring network structures in cellular signaling that are known to be responsible for key information processing characteristics such as multistability, biphasic responses and homeostasis.

The incoherent feedforward network motif (model N3), represents a simple network where a signal activates both a kinase and a phosphatase (and more generally activating opposing effects) and is known to present biphasic responses [128, 126]. This nominal network behavior can be affected by the presence of a biphasic interaction between the nodes in three ways; 1. The behavior can be reinforced or strengthened leading to a more prominent overall biphasic response, 2. The two distinct biphasic responses (from the network structure and from the interaction between the nodes) can combine to give multiphasic response, and 3. The two distinct biphasic responses can cancel out resulting in a simple monotonic dose response (see fig. 3.3D).

**Feedback motifs.** The positive feedback network motif, characteristic of reinforcement, is known for introducing multistability [52, 136]. Biphasic interaction between the nodes in such a network motif is capable of destroying such multistability arising from the feedback structure (see fig. 3.3E). Similarly the negative feedback network motif, characteristic of homeostatic response [137], can have this behavior impacted with the introduction of combinations of behaviors including homeostasis, multistability and monotonic responses (see fig. 3.3F).

The above results indicate how hidden biphasic responses between nodes in networks can not only fundamentally alter expected network responses, but can also lead to incorrect inferences regarding role and capability of the network.

**Complex signal regulation and interaction of biphasic responses from distinct sources.** Cellular signaling often involves both complex signal regulation and additional complexity at the level of a node, with the node itself being capable of complex behaviour

(such as the DSP motif). This implies the possibility (in addition to the cases considered above) of biphasic responses both in interactions as well as within a node in a network. To explore this interplay further, we consider signal regulation of DSP (with common kinase and phophatase). This allows us to explore the interplay of biphasic response arising from two fundamentally different sources, 1. the network (incoherent feedforward signal regulation) or biphasic response within an interaction between nodes, and 2. the intrinsic biochemistry of the node (DSP, is itself capable of exhibiting biphasic response - see fig. 3.3B,C). We incorporate these aspects in simple models N1 and N2 (see fig. 3.6); In the former instance (model N1, fig. 3.3B) the biphasic response is inbuilt into the signal regulation activating kinase, where as in the latter (model N2, fig. 3.3C), the biphasic arises from the incoherent feedforward network structure (signal activating both enzymes).

When the biphasic interaction from the network arises directly from the regulating signal, the combination of biphasic responses from the network and the node can result in(i) introduction of biphasic (when the DSP does not present it inherently) (ii) when the DSP presents biphasic response independently, strengthen the response and also generate distinct zones of multistability and homeostasis (fig. 3.3B). When the biphasic behaviour at the network level arises out of the incoherent feedforward regulation, similarly the behavior could be strengthened or destroyed fig. 3.3C. In both instances multiphasic response is also possible.

Thus beyond simple signal regulation, minimal complexity at the level of the biochemistry of the node and biphasic responses (either within the signalling interaction, or arising from network structure) can lead to a plethora of complicated and non-intuitive signaling behavior.

***ERK* Regulation.** We a explored the propensity of the established *ERK* pathway to present substrate biphasic responses (with total amounts of *ERK*) and enzyme biphasic responses (with total amounts of *MEK*) in the concentration of the maximally modified *ERK* (*pYpTErk*). Studies in the past have focussed on the capacity of the *ERK* regulatory pathway to show substrate biphasic responses computationally [123].

The *ERK* pathway is capable of exhibiting substrate and enzyme biphasic responses in the maximally modified substrate form (see fig. 3.4AB). The system is in fact capable of biphasic dose response with both doses simultaneously, for the same underlying kinetic parameters (see fig. 3.4C). Analytical work further reveals that substrate biphasic responses are guaranteed to exist for some total amounts of Mek and phosphatase, irrespective of intrinsic kinetic parameter values. Analytical work, revealed an explicit condition (dependent on total amount of phosphatase and kinetic parameters) that guarantees the existence of enzyme biphasic response. This condition is however only sufficient and not necessary for observing this (see appendix C, section 5).

These results are summarised and consolidated in fig. 3.4 and section 3.5.

## 3.4 Discussion

Biphasic dose response are ubiquitous in cell signalling and biochemical networks. The overall attenuation with dose suggests in safeguarding mechanisms which limit and set bounds on output and are associated with ”optimal” range of activation. We studied this using a three pronged approach: first focussing of basic biochemical building blocks, then studying their impact in basic network structures and finally investigating concrete systems. Central to our approach is a unified approach considering essentially distinct doses: both total enzyme amounts and total substrate amounts

Our analysis focused on a suite of commonly occurring substrate modification systems and revealed explicitly which systems were capable of generating biphasic responses. For these systems some exhibited biphasic responses only to the enzyme, while others did so only to the variation of the substrate amount. Other systems exhibited biphasic responses to both substrate and enzyme, this representing a robust multi-pronged inbuilt safeguarding mechanism.

We find that enzyme sharing promotes biphasic responses, enabling them when they didn't exist otherwise. Multisite modification, especially with commonality of enzymes, promotes this as well. Interestingly, a commonality of kinase promotes a substrate biphasic response, while a commonality of phosphatase promotes an enzyme biphasic response (also seen in cascades). A commonality of both kinases and phosphatases thus enables both substrate and enzyme biphasic responses. In many instances the behaviour is surprisingly robust, and for any given kinetic parameters, a suitable choice of enzyme and substrate amounts (experimentally accessible factors) guarantees that it will occur.

We then investigated the impact of such biphasic responses in networks by associating certain interactions (which may contain multiple species implicitly) with biphasic responses. We find that biphasic interaction patterns can completely abolish expected behaviour in networks, and generate unexpected new behaviour which may not be characteristic of the network. We also explored the interplay of two sources of biphasic interaction: the network regulation and substrate modification and show that this could lead to reinforcement, cancellation, or even multiphasic or other (eg multistable responses). Consequently the presence of biphasic responses has an even greater import in the behaviour of signalling networks than may be expected from basic considerations.

We now discuss the relevance of our results for both systems and synthetic biology

The first basic observation is that many ubiquitously occurring structures either enzyme biphasic or substrate biphasic behaviour, and in some cases both. An example is enzyme sharing which is common in cellular networks, especially since kinases significantly outnumber phosphatases, leading to phosphatases deployed in different kinase mediated reactions [101]. Interestingly while certain building blocks do not admit

biphasic responses (unless some additional feature such as substrate inhibition occurs), others do, and in certain cases they can even be obtained just by tuning substrate and enzyme total amounts (completely independent of intrinsic kinetics). Since these building blocks are ubiquitous in cells, this already indicates a number of built in sources of safeguarding against overexpression of active entities in signalling networks. Thus a careful inspection of signalling networks and modules is likely to reveal multiple potential intrinsic sources of such safeguarding, in addition to network regulatory structures. As an illustration, we have investigated a single tier of the *ERK* signalling network and demonstrated how this exhibits robust substrate biphasic responses (without restriction on intrinsic kinetic parameters, in contradiction to the picture suggested in [123]) and can also exhibit enzyme biphasic responses in broad parameter ranges.

The fact that there are multiple intrinsic sources of safeguarding and biphasic responses in these networks (i.e. even intrinsic to a node in these network), suggests that there may be a hierarchy of such mechanisms–ranging from network structures (such as incoherent feedforward networks) to additional ingredients in modification systems (such as scaffolding) to the intrinsic modification systems themselves. This is especially relevant in inferring network structures from data and suggests that the inference process must be able to account for different sources of such behaviour. Our analysis suggests that focussed experiments can discriminate between these different sources of biphasic responses and simple controlled test tube experiments can reveal the presence of intrinsic biphasic responses.

It is worth examining the presence or absence of biphasic responses relative to desired functionality (both in natural and engineered biology). To start with, if we focus on enzyme biphasic responses then we have have 4 distinct possible cases depending on the presence or absence of this behavior and when its presence is desired or not. Two of these cases (biphasic present, desired) (biphasic absent, undesired) represent desirable outcomes, and the key insight in the former case is that this may be obtained in broad swathes of parameter space by just varying total amounts of enzyme and substrate. The other two cases (biphasic not present, desired), (biphasic present, undesired) represent a mismatch between desired functionality and capability. In the former case, alternative structures can be engineered or realized in evolution, for instance the creation of incoherent feedforward networks. In the latter case a tuning of total enzyme and substrate amount can be used to prevent biphasic responses in desired ranges of operation, and if this can be done consistent with information processing goals, a desired outcome can be obtained.

The same picture can now be expanded to consider the possibility of both enzyme and substrate biphasic in relation to whether those biphasic responses are desirable or not (see Fig 4). This results in 16 cases, 4 of which correspond to desired outcomes being met, 4 of them corresponding to both outcomes not desired, and 8 outcomes where one outcome is met and one is not. In order to "engineer" the system towards desired goals, different approaches can be taken depending on the specific case and mismatch.

For instance mismatches in enzyme biphasic responses can be addressed as above. For mismatches in substrate biphasic responses, the response can be bypassed by tuning enzyme and substrate amounts. However if its presence is desired then a way of approaching this is by introducing extra structures in the substrate modification system (for instance similar to other building blocks which realize this). Clearly depending on the context the managing of desired outcomes requires a combination of structural and (easy to manipulate) parameter manipulation and this may involve a delicate optimization to keep it consistent with the information processing goals. This provides a basis from which to evaluate what steps may have incorporated in evolution, in particular the extent to which intrinsic capabilities have been used.

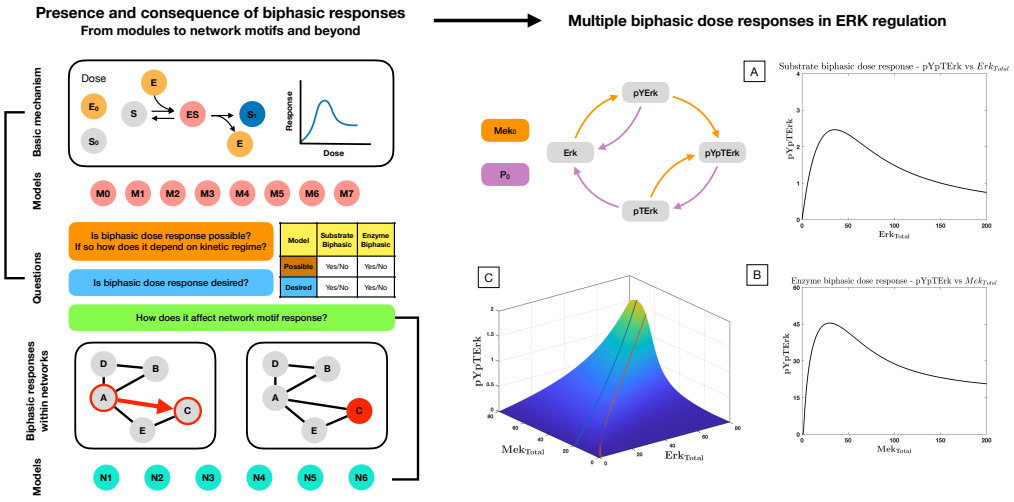
Another aspect is the relevance to network behaviour. We have already demonstrated that the presence of biphasic responses of the kind studied lead to biphasic interactions between network networks and this can significantly alter network behaviour, even resulting in behaviour which might be expected from a very different network structure. The presence of safeguarding mechanisms/biphasic behaviour at different locations in a network have important consequences for the overall robustness of the network as well as the maintenance of behaviour within bounds. The analysis of how distributed sources of such safeguarding (both enzyme and substrate biphasic behaviour) determines the overall safeguarding behaviour in the network and keeps relevant species behaviour within suitable bounds is a topic of future investigation.

Biphasic responses can be used in networks to access specific downstream behaviour in specific ranges of inputs. As such it can combine with complex dynamical behaviour (eg. bistable switching, oscillations) in different ways giving rise to behaviour which may appear unintuitive, if the presence of biphasic behaviour is not recognized. Interestingly, the building blocks we have examined giving biphasic behaviour can intrinsically also generate other complex behaviour such as bistability, which can add an additional layer of complexity to the network of which it is a part.

The results have further specific consequences for synthetic biology. Biphasic responses have been generated from synthetic gene regulatory incoherent feedforward circuits [138]. There have been multiple recent efforts to engineer biochemical circuits to obtain desired behaviour [84, 139, 140] and these include engineering substrate modification as well [65, 83]. Our demonstration of basic building blocks generating biphasic responses is of particular use in engineering biochemical networks: the presence and guarantee of obtaining such behaviour, in some cases the widespread occurrence of such behaviour occurring, which can be accessed by varying simple easy to access factors (substrate and enzyme amounts). Our analysis in totality provides structural criteria which can generate such behaviour and where necessary also provides kinetic parameter ranges where this may be obtained. This provides engineering design principles for creating such behaviour from basic biochemical building blocks. Furthermore, our analysis can be used to engineer both substrate and enzyme biphasic responses, representing a robust safeguarding mechanism for preventing over expression of spe-

cific substrate forms.

Building on the analysis here, it is possible to design biochemical biphasic modules, with a range of desired behaviour along with the shaping of the quantitative characteristics of the biphasic behaviour. This can then be incorporated into chemical reaction network to shape outcomes, regulate behaviour and create different safeguarding mechanisms against over expression in networks. Interestingly some preliminary analysis shows how basic modules studied here when combined with positive and negative feedback can in fact maintain, not just the biphasic response, but specific characteristics such as the peak amplitude, thus testifying to both the robustness of the insights and its value in engineering design.




| Biphasic dose responses in enzymatic modification modules | | | | | | | Consequence of biphasic interactions in network motif response | | |
|---|---|---|---|---|---|---|---|---|---|
| System | | Bistability | Substrate biphasic | | Enzyme biphasic | | Network | Expected output | Perturbation |
| | | | Biphasic | Multistability with Biphasic | Biphasic | Multistability with biphasic | | | |
| Single site modification | Covalent Modification Cycle (CMC) | | No | | No | | Regulation of Covalent modification cycle | • Monotonic response<br>• Switch like response | • Homeostasis<br>• Biphasic response |
| | Protein-Protein Interaction | | No | | Yes | | | | |
| Double Site | Common Kinase Common Phosphatase | Yes | Yes | Yes | Yes $^{\S}$ | Yes | Regulation of Positive Feedback motif | • Multistabilitiy<br>• reinforcement | • Remove bistability and induce mono stability<br>• Biphasic response |
| | Common Kinase Separate Phosphatase | Yes | Yes | Yes | No | | | | |
| | Separate Kinase Common Phosphatase | Yes | No | | Yes $^{\S}$ | Yes | | | |
| | Separate Kinase Separate Phosphatase | | No | | No | | Regulation of Negative Feedback motif | • Homeostasis | • Homeostasis<br>• Homeostasis with bi-stability<br>• Monotonic response |
| Beyond Multisite | Coupled CMC (Common Enzymes) | | Yes | | No | | | | |
| | Coupled CMC (Different Kinases) | | No | | No | | | | |
| | Two Tier Cascaded system (Common Phosphatase) | Yes | Yes $^{\S}$ | Yes | No | | Regulation of Incoherent Feedforward | • Biphasic response | • Homeostatic response<br>• Multiphase response |
| | Two Tier Cascaded system (Different Phosphatase) | | No | | No | | | | |

Note: Unless otherwise mentioned behavior can be observed in all kinetic regimes
$^{\S}$ - Can be observed in certain well defined kinetic regimes

**Figure 3.4: Presence and consequence of biphasic dose responses in cellular signaling arising from enzyme mediated substrate modification.** *Left hand panel*: Schematic representation of the study. The basic mechanism of enzyme-mediated substrate modification allows for simple biochemical systems to present biphasic dose-response behavior. This is analyzed in a series of commonly encountered biochemical building blocks of cell signalling (M0-M8, see fig. 3.9 for detailed schematic). In each case, we assess whether the system exhibits biphasic behaviour with respect to variation of total amounts of enzyme and substrate as doses, and if is does so robustly. The table below, provides a classification system to determine the different possibilities of biphasic responses (enzyme/substrate) vis-a-vis the desirability/undesirability of biphasic responses. This provides a framework (in synthetic biology) for determining what changes need to be made to a system, to obtain desirable characteristics. We then extend the study to consider the consequence of such biphasic dose responses in network motif structures and study how biphasic interactions can alter expected outcomes from these systems (N1-N5). *Right hand panel*: Case study: Extracellular signal-regulated kinase (Erk) Regulation. The ERK regulation model is capable of both exhibiting substrate and enzyme biphasic response in the double phosphorylated Erk (pYpTErk) in the steady state dose response with $Mek_{Total}$ (Enzyme biphasic) and $Erk_{Total}$ (Substrate biphasic). A-B shows an instance of enzyme and substrate biphasic. C shows how the steady state concentration of pYpTErk changes with $Mek_{Total}$ and $Erk_{Total}$ (for a fixed kinetic parameter set). This indicates the presence of simultaneous substrate and enzyme biphasic response in the system (red and blue lines indicate the peak concentration of pYpTErk achieved for a given amount of total substrate and total enzyme amounts respectively). *Table*: Detailed summary of the various results discussed in the manuscript, including the possibility and impossibility of specific dose-response behaviors in a model, and the different ways in which expected outcome of a network motif structure is undercut by biphasic responses within interactions. *Note*: These results discuss analytical results which can be found in appendix C

Our examination of different building block structures in comparison with one an-

other provides further insights. For instance the presence of a "dead-end" complex (protein-protein interaction system) reveals a robust presence of enzyme biphasic responses, while double site substrate modification with either common kinase or different kinases but common phosphatases provides enzyme biphasic responses in a substantial region of kinetic parameter space (the region in fact being exactly the same). This suggests that the separate kinase common phosphatase system provides additional tuneability, along with access to other behaviour, with no loss of capability in terms of biphasic responses. Interestingly, a number of substrate modification systems (double site modification common kinase common phosphatase, common kinase separate phosphatase, as well as coupled modification cycles sharing a common kinase and phosphatase all yield robust substrate biphasic responses across kinetic parameter space. In the first two cases, the possibility of having different phosphatases may again provide additional tuneability. In the third case, one of the cycles is guaranteed to give substrate biphasic responses, the identity of which is determined by an explicit analytic expression involving kinetic constants. Taken together, this illustrates how our analytical work reveals tangible structural and organizational insights directly relevant to design.

All in all, our unified exploration of enzyme and substrate biphasic responses through the analysis of ubiquitous building blocks in networks opens up new vistas for exploration in natural biology, engineered biology and the emerging area straddling the two.

## 3.5 Appendix

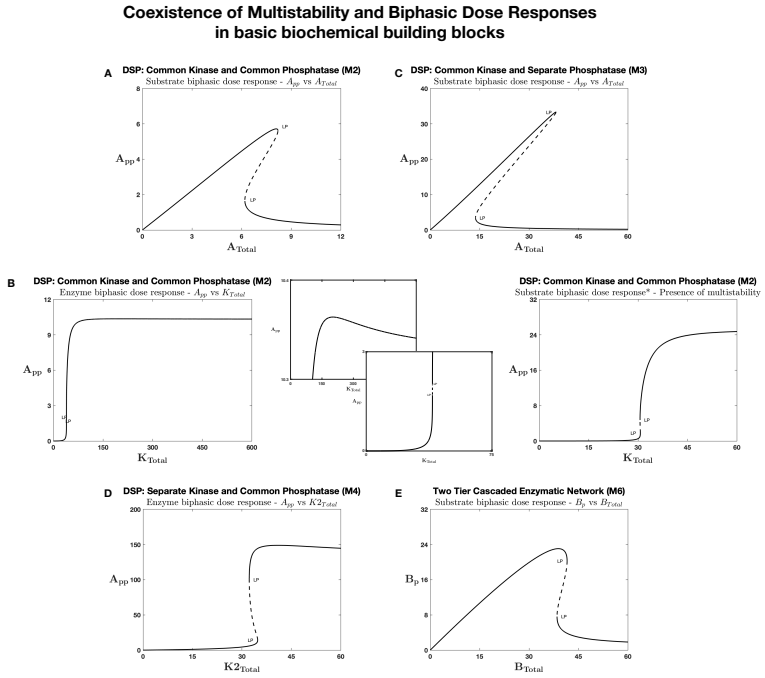



**Figure 3.5: Coexistence of multi-stability and biphasic dose response in various enzymatic models for the same underlying kinetic regime**. The above panels illustrate computational evidence of the possibility of coexistence of biphasic response with multi-stability, in a 1D bifurcation along the respective dose (for the same/similar underlying kinetics), for the different networks considered. The results are shown for all models capable of biphasic response (from within our study) and multistability. Note that coupled covalent system (with common kinase) while capable of substrate biphasic is incapable of multistabilty. A. DSP with common kinase and phosphatase: Substrate biphasic response and multistability for the same underlying kinetics. B. DSP with common kinase and common phosphatase: Enzyme biphasic response and multistability - for different (but close) parameter regimes. Inset plots show magnified regions from the main figure. C. DSP with common kinase and separate phoshpatase. Substrate biphasic response and multistability for the same underlying kinetics. D. DSP with separate kinase and common phosphatase. Enzyme biphasic response and multistability for the same underlying kinetics. E. Two tier cascaded enzymatic network. Substrate biphasic response and multistability for the same underlying kinetics. This shows how the enzymatic modification mechanism (and the associated sequestration and non-linearity in interaction) within these networks allows for not just the capacity to present biphasic dose response behavior but also to be present in conjunction with other qualitative signaling behaviors. [LP: saddle node bifurcation, solid and dotted lines denote stable and unstable steady states respectively]

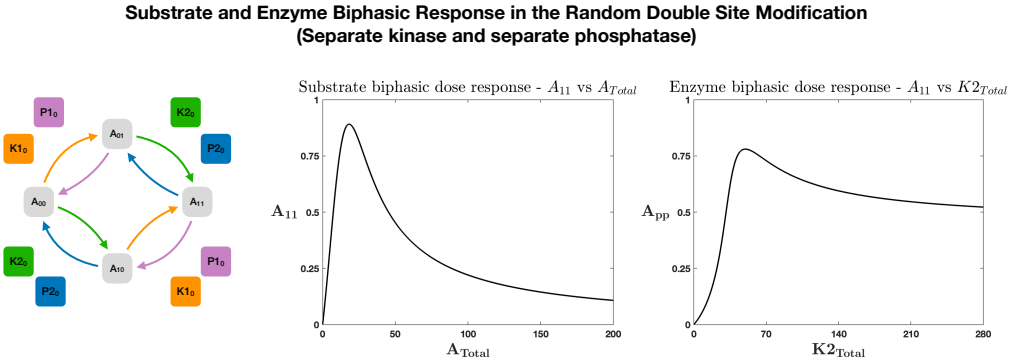



**Figure 3.6: Biphasic responses in random double site modification network with separate enzymes effecting each modification.** While the ordered DSP with separate enzymes acting on each modification site is incapable of exhibiting any biphasic dose response, random ordered (de)modification exhibits both substrate (*left*) and enzyme (*right*) biphasic dose response. This illustrates how additional complexity (inherently) associated with enzyme mediated substrate modification (random) can further expand propensity for such networks containing them to generate biphasic dose response.

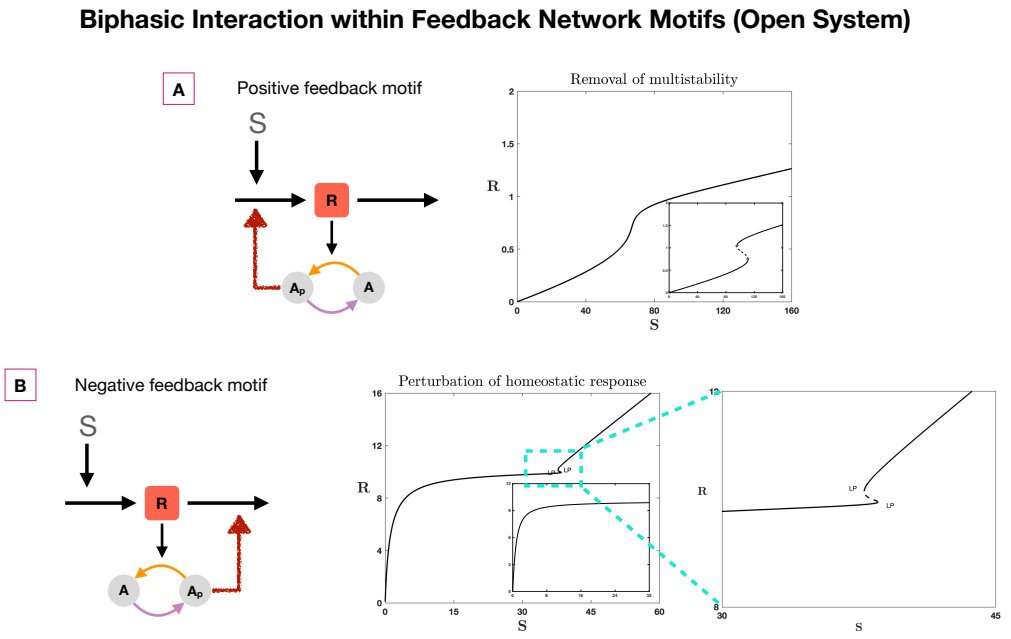



**Figure 3.7: Effect of biphasic response in interactions within feedback network motifs (open systems)**: The figure shows introduction of novel behavior (bistability in NFB, see *B*), and removal of expected behavior (bistability in PFB, see *A*) even when the biphasic responses in interactions are present within open system feedback network motifs. This complements our analysis in the paper where similar results where shown with motifs constructed with closed system (see **??**). The inset plots show basal behavior without the biphasic interaction. *B* shows expanded region in the aside main plot. [Brown arrows in the network schematics represent biphasic interactions. LP: saddle node bifurcation. Solid lines and dotted lines denote steady and unstable steady states respectively.]

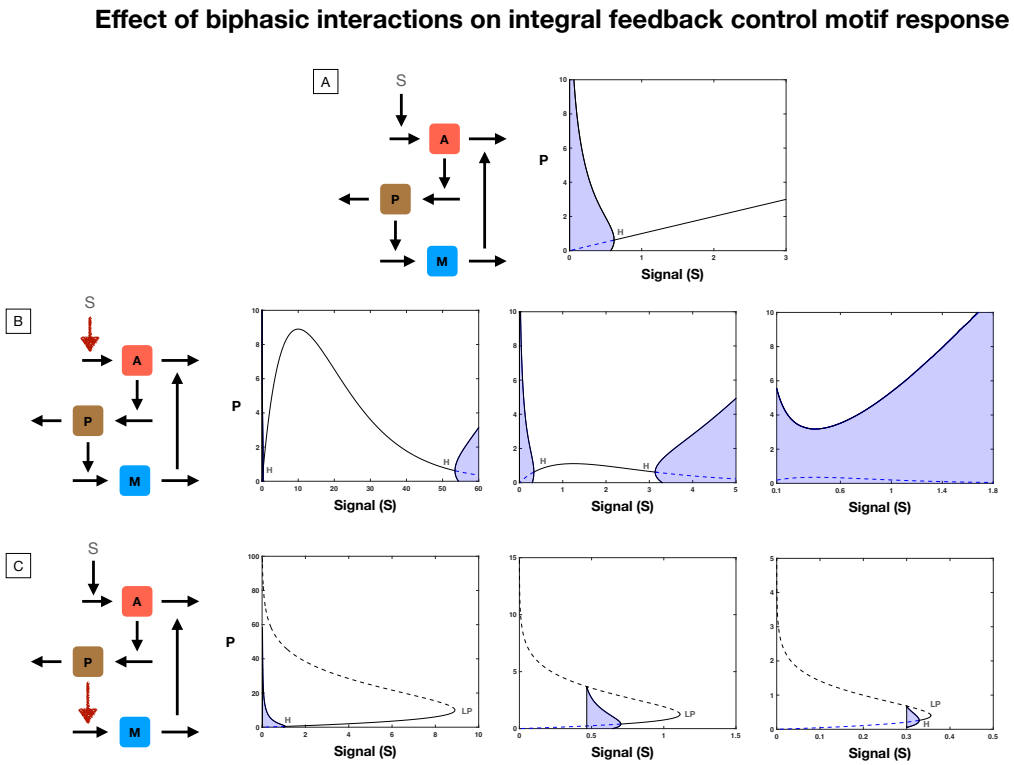



**Figure 3.8: Effect of biphasic response in interactions within integral feedback control motif:** perturbation of expected homeostatic response. The simple integral feedback control motif is capable of exact homeostasis in the substrate form *M*, as signal (S) and substrate forms *P* and *A* change (see plot *A*, *M* not shown). The lower regions of signal activation are unstable for the system and oscillate. There exists a threshold of signal above which the single steady state is stable and there is steady exact homeostasis in M. In plots *B* and *C* we show how expected response from the motif can be affected by 1. biphasic signal regulation (*B*), 2. biphasic response in an interaction within the motif (*C*)). *B* shows how upstream biphasic signal regulation can diminish the range over which stable steady state homeostatic response is seen (in concentration of M - not shown), by making the system unstable through a hopf bifurcation leading to oscillations, for both low signal and high signal ranges. Depending on the nature and strength of the biphasic signal regulation (manipulated here by changing parameter values associated with the signal regulation), the stable steady state regime in signal can diminish and for a sufficiently strong biphasic signal regulation, the system can be completely unstable, completely undermining the exact homeostasis expected in *M* from the motif. *C* shows how biphasic responses in an interaction within the motif (indicated in the adjacent schematic with a brown arrow) can further arrest the expected biphasic response by 'capping' the signal range over which the motif can function (due to the saturating effect from the biphasic response - see text). Further, the steady state of the system can loose stability even for lower signal ranges (on the existing stable branch), due to oscillations. Thus, biphasic interactions within even a robust homeostatic motif such as integral feedback control can fundamentally subvert expected motif response. [Brown arrows in the network schematics represent biphasic interactions. LP: saddle node bifurcation. H: Hopf point bifurcation. Solid lines and dotted lines denote steady and unstable steady states respectively. Regions shaded in blue represent oscillations.]

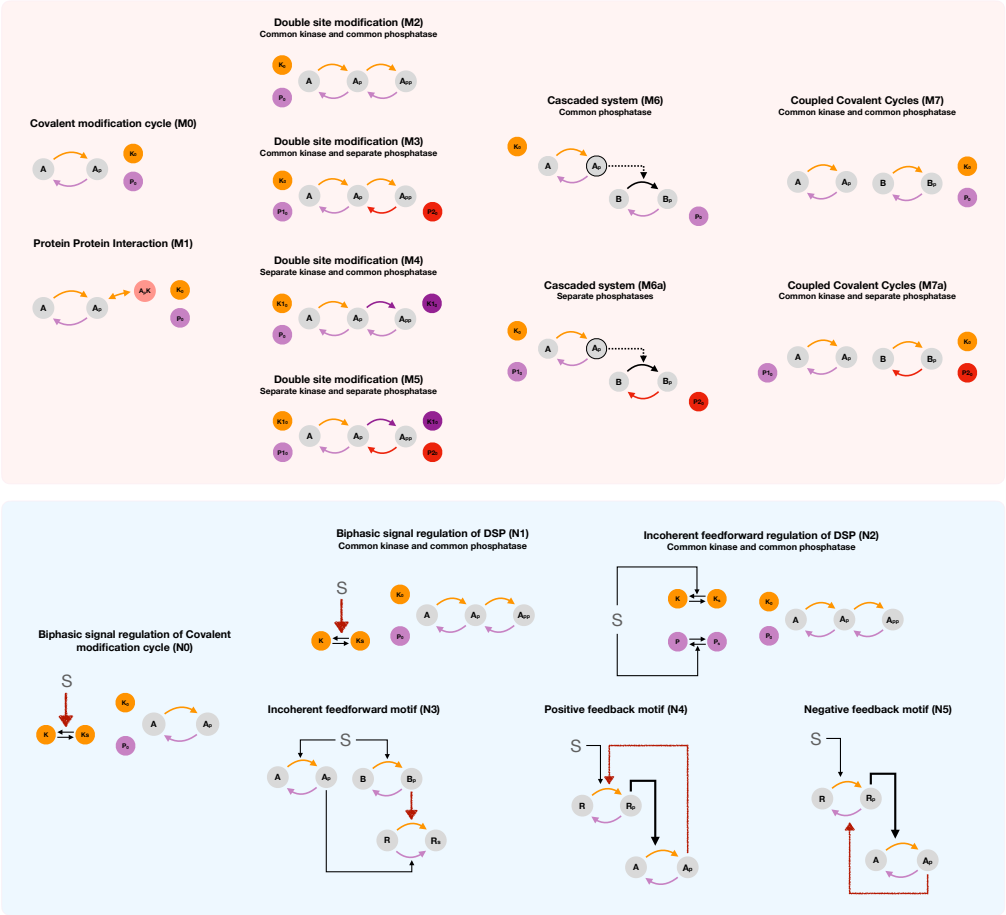



**Figure 3.9: Schematic representation of the various models considered in the manuscript.** Models M0-M1 represent the basic unit of (de)modification, the covalent modification cycle and its simple extension with the substrate forming an additional complex with the enzyme. Models M2-M5 represent different systems of ordered distributive double site modification of a substrate (by different or common enzymes acting on each modification site). Models M6-M7 represent other architectures of modification systems in chemical reaction networks, such as cascaded system (with different and common enzymes affecting (de)modification) and coupled covalent modification networks (coupled through enzyme sharing in (de)modification). Models N0-N5, represent different networks and network motifs with biphasic dose-response subsumed in an interaction.

## Executive summary of analytical results and resources

**Table 3.1:** Executive Summary - Chapter 3: Biphasic responses in cell signalling: A unified approach

| Model | Biphasic in | Brief description of analytical result | Anal. | Comp. |
|---|---|---|---|---|
| | | **Single site modification system** | | |
| Covalent modification system | Substrate | Absence | C(1.1a) | - |
| | Enzyme | Absence | C(1.1b) | - |
| Protein Protein Interaction | Substrate | Absence | C(1.2a) | - |
| | Enzyme | Present for all kinetic regimes | C(1.2b | fig. 4.2A |
| | | **Double site modification system** | | |
| Common kinase and phosphatase | Substrate | Presence for all kinetic regimes | C(2.1a) | fig. 4.2A |
| | Enzyme | Presence when kinetic constraint (involving catalytic constants) is satisfied | C(2.1b) | fig. 4.2B |
| Common kinase and separate phosphatase | Substrate | Presence when kinetic constraint (involving catalytic constants) is satisfied | C(3.1a) | fig. 4.2C |
| | Enzyme | Presence for all kinetic regimes | C(3.1b) | fig. 4.2C |
| Separate kinase and common phosphatase | Substrate | Absence | C(3.2a) | - |
| | Enzyme (K1) | Absence | C(3.2b) | - |
| | Enzyme (K2) | Present | - | fig. 4.2C |
| Separate kinase and phosphatase | Substrate | Absence | C(3.3a) | - |
| | Enzyme | Absence | C(3.3b) | - |

Continued on next page

Table 3.1 – continued from previous page

| Model | Biphasic in | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| | | **Beyond multisite systems** | | |
| Two tier cascaded (Common phosphatase) | Substrate (B) | Presence for all kinetic regimes | C(4.2a) | fig. 4.2E |
| | Substrate (A) | Absence | C(4.2a) | - |
| | Enzyme | Absence | C(4.2b) | - |
| Two tier cascaded (Separate phosphatase) | Substrate (both) | Absence | C(4.2b) | - |
| | Enzyme | Absence | C(4.2b) | - |
| Coupled Covalent modification (Common kinase and phosphatase) | Substrate | Present for all kinetic regimes | C(4.3a) | fig. 4.2F |
| | Enzyme | Absence | C(4.3a) | - |
| Coupled Covalent modification (Separate kinase and common phosphatase) | Substrate | Absence | C(4.3b | - |
| | Enzyme | Absence | C(4.3b) | - |
| Coupled Covalent modification (Common kinase and separate phosphatase) | Substrate | Absence | C(4.3c | - |
| | Enzyme | Absence | C(4.3c) | - |
| | | **Biphasic in interaction within network motifs** | | |
| Negative feedback motif (Open system) | Neg. feedback | Absence of multistability | C(5.1a) | - |
| Negative feedback motif (Closed system) | None | Absence of multistability | C(5.1a) | - |
| | Neg. feedback | Presence of multistability | - | fig. 4.3 |

Continued on next page

Table 3.1 – continued from previous page

| Model | Biphasic in | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| | | **Concrete pathways** | | |
| ERK regulation | Substrate | Presence for all kinetic regimes | C(6a) | fig. 4.4 |
| | Enzyme | Presence when kinetic constraint (involving catalytic constants) is satisfied | C(6a) | fig. 4.4 |

# Chapter 4

# Network regulation meets substrate modification chemistry

Note: Work presented in this chapter builds on work by ex-research group member Dr. *Thapanar Suwanmajo* in his research thesis [1]. Dr. *Thapanar Suwanmajo* performed the initial computational analysis exploring the interplay between the network motifs and multisite modification in his thesis chapter. This work was built on by Dr. *J. Krishnan* and myself. I expanded on his computational work; including recreating the models, confirming and generating the results for different parameter regimes, as well as identifying new results computationally (particularly in the context of isolas in Inc. FF, tristability from NFB, and the entirety of the analysis on feedback with $A_p$ - discussed in detail within the chapter). I also initiated and carried out in its entirety the analytical and semi-analytical work described in this paper. Synthesis of the results and the methodological approaches (presented as a toolkit for similar systems) were carried out by Dr. *J. Krishnan* and myself. All figures presented here in were generated by myself from scratch.

## 4.1 Introduction

Networks represent a pervasive theme across the natural sciences, engineering and the social sciences. They have become a basic way of representing and analyzing complex interacting systems/agents/entities in these various disciplines. One of the disciplines where the reality of networks is inescapable is the life sciences. The fundamental unit of life, the cell, relies on a massive, highly sophisticated information processing operation for maintenance and progression of life, and a flexible response to both the external and internal environment. Along with genetic networks, biochemical networks (both signalling and metabolic) represent the heart of information processing in cells.

On the face of it, biochemical networks are networks where the nodes and interactions represent biochemical reactions, and cellular biochemical networks (seemingly) represent an unstructured agglomerate, shaped by evolution to perform specific tasks.

However a closer look reveals that this is not the case. There are recurring network structures and motifs - for instance different types of feedback and feedforward regulation which are repeatedly encountered [127, 141, 52, 126, 133]. These motifs simply encode key regulatory logic structures in these circuits and are often key drivers for commonly observed behaviours such as switching, oscillations, multistability, homeostatic and biphasic responses [52, 142, 143, 144].

The central component of these networks are proteins. Proteins themselves perform multiple functions and a key way of establishing functionality of different proteins is by post-translational modifications. The post-translational modifications, involves a variety of basic covalent modifications (eg. phosphorylation, ubiquitation) along with more elaborate multi-site modification mechanisms, and represents a distinct axis which is a part of signalling networks. In addition, there are also other regulatory factors at the level of substrate modification, including scaffolding and enzyme-substrate feedback [100, 145, 146, 83]. Interestingly, it has been found that post-translational modification of substrates themselves may be capable of generating non-trivial qualitative information-processing [77, 26, 78, 74, 29, 147, 148, 123, 149].

It is striking to note that while both the network level and the substrate modification level are focal points of many investigations (including, increasingly, data driven and data-rich approaches [150, 151]), they are often studied separately. There is very little systematic systems understanding of how these two levels combine and interact, even though they both are part of the same cellular biochemical networks. This is surprising as their interplay is central to a systems understanding of cellular information processing, being relevant to many aspects of cellular functioning: basic information processing characteristics, behaviour of concrete pathways, the impact of evolution on pathways/networks, the reverse engineering of networks, and network engineering in synthetic biology.

An example of the complexity of the substrate modification level is provided by multisite substrate modification (wherein a substrate is reversibly modified at multiple sites by different or common enzymes). Multisite modification, is a key constituent of many biochemical networks present alongside network regulation. This includes degradation of cell-cycle protein *Sic1* [131, 152], degradation of the *p53* protein (which itself undergoes numerous modifications [153]) mediated by multisite phosphorylation of *MDM2* [154, 155], the regulation of insulin (via phosphorylation of *SER* 312 protein) and *IP3* (implicated in Huntington's disease) via *IPMK*, both involving potential negative feedback [156, 157, 40]. The combination of multisite modification and network regulation underpins important cellular processes, including cell-cycle regulation [158], circadian rhythms [120], inflammation [109], and different diseases [23].

There are many instances in such cases where the interplay of multisite modification and network regulation is a direct focal point of interest. An increasing number of studies focus on both the inference and study the role of the multiple independent feed-

back regulation in *p53* [49, 159, 160] and *Sic1* proteins [131]. Of special interest is how the network regulation is affected by the modification chemistry of these proteins and how their combination provides functionality to the proteins in the respective pathways. Such interplay is also seen to underscore functioning of key cellular processes such as Circadian rhythms (transcription-translational feedback loops [161] and in the *MAPK/ERK* pathway [51, 146]). The complexity in these networks (both the number and nature of modifications, and the different network regulation present) requires a systematic dissection to understand the contribution of each ingredient to the emergent systems behavior.

In this paper, we develop a systems framework to investigate the interplay between network regulation (and network structure/topology) and substrate modification biochemistry. We use multisite substrate modification as an exemplar case for substrate modification. At the network level, we examine canonical network structures of feedforward and feedback regulation, which are broadly encountered.

The basic question we aim to address is: how do the network level and the substrate modification level combine? What does substrate modification contribute to network behavior and what does the ambient network contribute to substrate modification? We deploy a consolidated approach combining computation, analytical work and semi-analytical approaches to address this and reveal the systems landscape of multi-level interaction.

Addressing this question provides important insights into how the two levels combine in various concrete biochemical networks and allows us to trace the actual origins of a given behaviour (whether from the network, the substrate modification layer or a combination). In so doing, we reveal fundamental constraints associated with working at one level, and how distinct emergent behaviour may arise through a non-trivial confluence of levels. This also has implications for the conceptualization and modelling of cellular networks, the inference of networks from data and assessing the role of evolution. Finally it provides a framework for multilevel engineering of such networks in synthetic biology.

Our analysis provides a number of key insights and testable predictions regarding the impact of feedback and feedforward regulation on multisite modification. In the context of double-site modification this reveals key aspects of this interplay, which may be encountered in specific systems, as referenced above, In addition, by focussing on key recurring structures and building blocks at both the network and substrate modification levels and performing a consolidated analysis, we reveal principles and obtain transparent generalizable insights into multi-level interaction, relevant to multiple different contexts. This then provides a distinct and fruitful (and in light of this study, a necessary) vantage point and methodology from which to study specific cases, by engaging with further systems-specific details (see Discussion). In fact specific contexts may be best approached by starting with foundational platforms such as these.

We point out that focussing on basic building blocks and network modules/motifs has been a very fruitful approach in cell signalling and information processing generally.

All in all, our analysis creates a bridge between studies (including many experimental studies) which focus on the network level alone, and studies which focus on substrate modification alone.

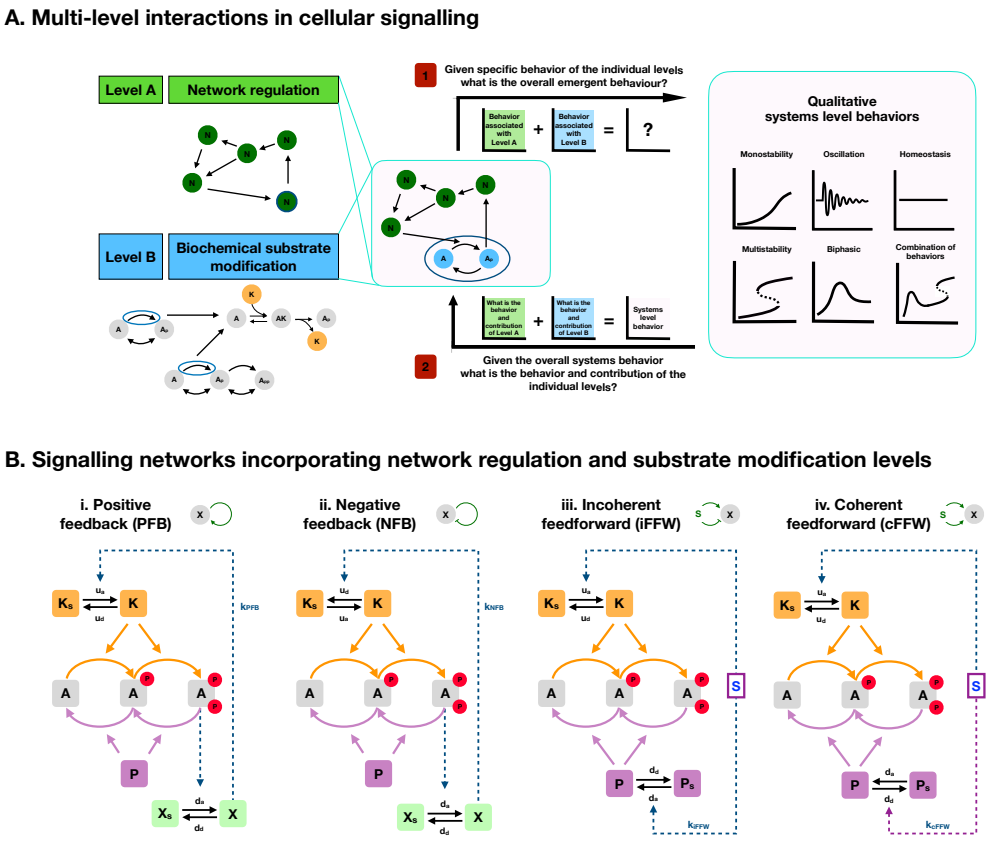



**Figure 4.1: Schematic description of signaling networks involving two levels, the network regulation level and the substrate modification level.** (A) The left panel depicts each of these levels individually; a regulatory network on one hand and a substrate modification level on the other. Each substrate modification reaction involves enzyme reversibly binding to substrates, creating complexes which results in the modification. Also depicted is a network incorporating both these levels. How these two levels interact is an open question. This can be viewed from two different vantage points 1. Given specific behavior of the individual levels, what is the overall emergent behavior? 2. Given the overall systems behavior, what is the behavior and contribution of the individual levels? The right panel depicts different types of representative behaviors of interest, either for the individual level or the complete system. (B) depicts the different signaling circuits studied in this paper, involving canonical network regulation (positive and negative feedback, and incoherent and coherent feedforward) and ordered double site modification, mediated by a common kinase and common phosphatase, as an exemplar of substrate modification.

## 4.2 Models and Methods

Our assessment of the interplay of network characteristics and substrate modification biochemistry is studied as follows. Our focus on the substrate modification level is on multisite modification. Multisite modification is capable of distinct information processing characteristics intrinsically, and this has been studied in detail [80, 31, 162, 122, 74, 26, 99, 4, 163]. For the substrate modification biochemistry, we focus on a fairly well-characterized substrate modification system: ordered distributive double site modification of a protein by a common kinase and common phosphatase (see appendix C). This is because, this system is directly encountered in multiple signaling contexts and because it is a building block for substrate modification systems. Focusing on multisite modification/demodification by a common kinase/phosphatase allows us to explore the effect of substrate modification complexity without having to consider the effects of multiple enzymes (and potentially multiple regulatory networks) at this stage (discussed later). We point out that our methodology and specific insights are also relevant to other multisite substrate modification systems (eg. random double site modification with separate/common kinases and separate/common phosphatases).

The model for the double-site modification system is a widely employed kinetic model (see appendix C) describing individual modifications with the relevant enzymes binding to the substrate, forming a complex, and then dissociating, leaving the modified substrate. The models for the individual modification use standard mass-action kinetic descriptions. This model has been widely studied in the literature.

With regards to the network level, we consider canonical network architectures: feedforward regulation (both coherent and incoherent) and feedback regulation (both positive and negative). The models of these networks again employ descriptions which are widely used in the literature (see appendix C). Each network structure contains different nodes, with inter converting active and inactive forms. A kinetic description of the inter conversions in each node as well as the network (node to node) interactions determines the model of the network The behaviour of these networks is also well characterized.

We analyze systematically the interplay between the network and the substrate modification complexity. We focus on key qualitative information processing characteristics (eg. biphasic dose responses, multistability-given its importance for switching, oscillations, homeostasis) as these are some of the most basic focal points in modelling, and which promise to provide the sharpest non-trivial insights in relation to experiments.

We consider network regulation overlaid on multisite modification (both realized through generic models: see appendix C) as follows. We regard the entire multisite modification system, along with its associated active and inactive enzymes as a node in the network. Regulation acts by altering the balance between active and inactive enzymes, either via feedforward regulation (purely by signals) or or via feedback regulation (involving a

modified substrate): see fig. 4.1 and appendix C. Further details on the computational tools and methodologies are also presented here.

Ordered double-site modification is known to intrinsically be capable of a range of behaviour such as hyperbolic dose-response characteristics, biphasic responses, bistability and oscillatory behaviour [77, 26, 74, 29]. This rich repertoire of behavior can be traced to the sequestration of enzymes and substrate in multiple complexes. With regard to network behaviour, positive feedback is associated with self-reinforcement, elevation of output, and in the presence of co-operativity, can give rise to bistability [164]. Negative feedback, by contrast has an attenuative effect, and is implicated in many homeostatic structures, including adaptation as well as (in conjunction with delay) oscillations [50]. Incoherent feedforward behaviour is a basic way of realizing adaptation while coherent feedforward networks represent a reinforcement from both regulatory legs.

Understanding the interplay of the network and the substrate modification amounts to addressing the question: what does the substrate modification contribute to network behaviour and what does the network contribute to substrate modification? In each instance, we aim to monitor transitions in qualitative behaviour: for instance, a distinct type of behaviour which is seen in the network as a consequence of substrate modification biochemistry, and transitions in behaviour of multisite modification owing to the network.

We note at the outset that elucidating the interplay of substrate modification and network regulation involves effectively engaging with the associated sets of parameters. This is achieved by a multi-pronged exploration of the underlying models using computational, analytical and semi-analytical approaches, specifically tailored to the questions of interest. Each of these approaches brings complementary insights and new perspectives and taken together, they constitute a consolidated global approach to answer these questions and reveal different facets of the interplay of the levels and their dependence on parameters. This is crucial for revealing the multi-level systems landscape transparently.

Computational results (combining simulations and bifurcation analysis), directly reveal transitions in behaviour and distinct new facets which emerge. In most instances (unless otherwise mentioned) the results obtained are seen widely in parameter space, and computations in multiple parameter regimes reveals this. In some cases, special/unexpected behaviour which is less widespread may be observed and this is noted as such. Analytical work on the other hand is focussed on different goals: (i) Explaining unexpected new behaviour observed or key transitions in behaviour (ii) Establishing unambiguously, the impossibility of certain behaviour/transitions in behaviour. Taken together, this establishes strict boundaries between the presence and absence of certain behaviour. This allows us to make sharp conclusions regarding the origins and key drivers of overall systems behaviour. Such analytical approaches have been applied in

different ways: (i) To the full combined system (requiring no assumptions/restrictions on any parameters) (ii) With some restrictions on the parameter regime of substrate modification (but none on the network). The restrictions could either be on relative enzyme and substrate total amounts or requiring certain enzymatic modifications act in the unsaturated regime. In each case, results demonstrating impossibility of behaviour are obtained independent of all other parameters (see appendix C). Semi-analytical approaches bridge computational and analytical approaches, and in so doing (i) provides a tighter and more transparent elucidation of computational observations and (ii) expands the applicability of analytical approaches by combining it with computation in different ways.

## 4.3 Results

Computational results are presented, followed by a concise summary of analytical/semi-analytical approaches.

### 4.3.1 Computational results.

We consider different possibilities of the behaviour of the multisite modification module in isolation as reference points for evaluating the impact of the network. This includes monotonic dose responses, bistability, biphasic responses and oscillations. Each of these reference behaviours is represented with a representative parameter set. For each of these reference behaviours, we evaluate the effect of each of the networks, varying key network parameters as needed. Such an approach exploring different combinations of multisite modification behavior and different network characteristics yields an array of results revealing the landscape of network substrate modification interaction with multiple insights. The results are presented for each of these networks, in turn.

The computational results are underpinned by a broad computational exploration. In this regard, we note: (i) The transitions we report are representative, in the sense that the transitions are seen for different parameter sets representing the basal behaviour of the multisite module (ii) In a few cases, there is a special behaviour which is less widespread and in those cases we note it as such. (iii) We present the noteworthy transitions in behaviour observed computationally. (iv) Many aspects of the computational results, especially the steady state behaviour, can be explained analytically and rationalized semi-analytically as well and our complementary explorations from these viewpoints provide associated results and approaches.

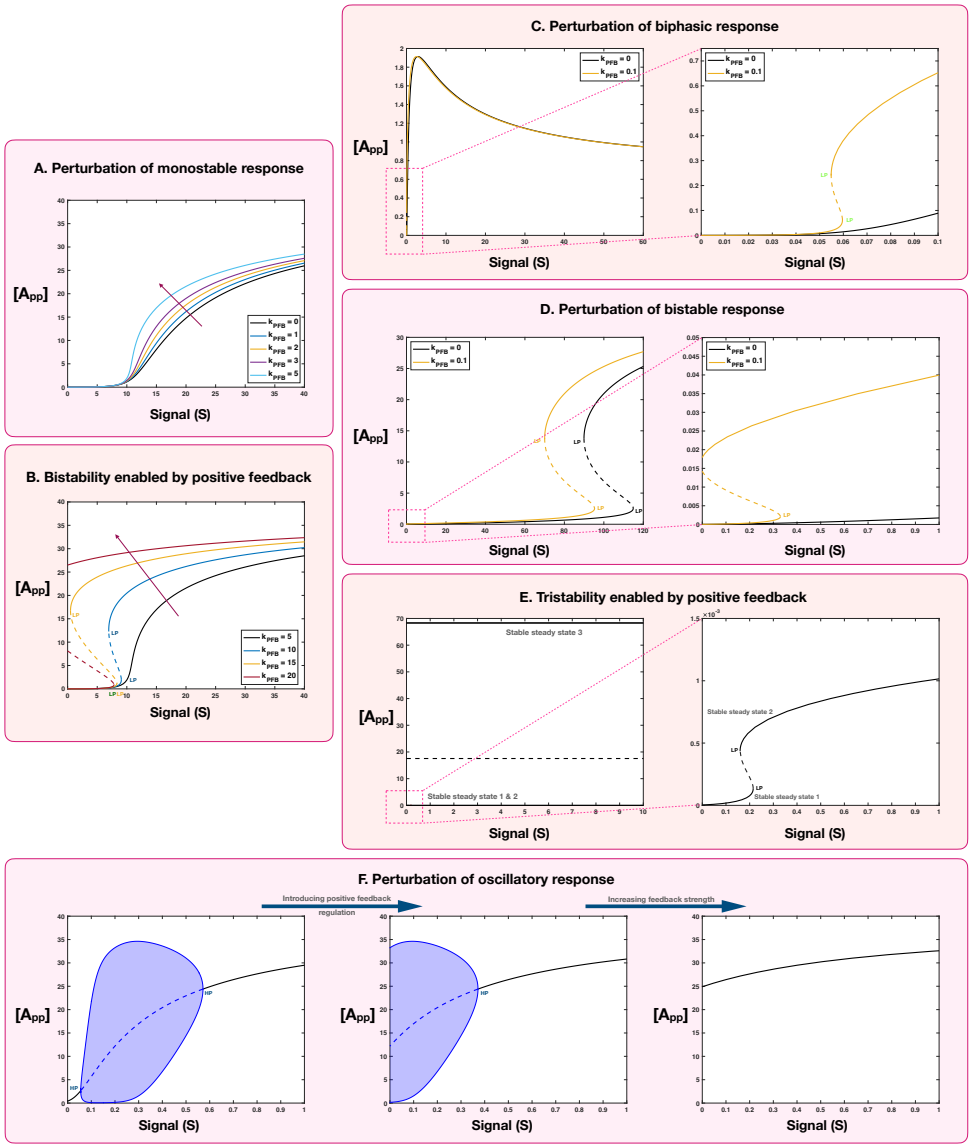



**Figure 4.2: Impact of positive feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system.** The results shown below demonstrate the effect of positive feedback on different basal characteristics of the double site modification system. (A,B). Perturbation of monostable basal response. (A) shows increasing positive feedback resulting in more switch like behavior. (B) shows how increasing feedback can induce multistability in the system (reversible and irreversible).) (C). Perturbation of biphasic basal response; increasing feedback strength can lead to a combination of multistability and biphasic behavior. (D). Perturbation of basal bistable response; increasing feedback strength can lead to distinct regions of multistability emerging. (E). Feedback can also induce distinct behavior such as tristability, that both the network regulation and the substrate modification levels are otherwise intrinsically incapable of. (F). Increasing feedback strength can shift and reduce the range of oscillations leading to its eventual abolishment. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during oscillations. HP: Hopf bifurcation, LP: Saddle node bifurcation.]

**Positive feedback.** Figure 4.2 demonstrates different aspects of positive feedback regulation associated with the fully modified phosphoform.

**Positive feedback can significantly impact switching behavior and enable the creation of new zones of dose-response characteristics.** Figure 4.2(A,B) demonstrates the impact on monostable threshold behaviour. Increasing the feedback can result in a more pronounced (monostable) switch-like behaviour and even result in bistability, which still maintains the threshold-like characteristic. The capacity of positive feedback inducing bistability is also seen in how positive feedback affects multisite modification in a bistable regime (Figure 4.2(D,E)). One manifestation of this is the introduction of an additional zone (in the dose-response curve) of bistability, resulting in multiple regimes of bistability. Here one can associate each distinct bistable zone with the impact of a different factor (intrinsic bistability, feedback). This demonstrates how different zones of dose-response can be engineered and tuned.

**Positive feedback can enable behaviour otherwise inaccessible.** In contrast to generating distinct zones of bistability, positive feedback can also enable tristability (See Figure 4.2E and analytical results in appendix C). This is noteworthy because it is impossible to obtain tristability both in the multisite modification module in isolation, and in the positive feedback network acting on a single site modification. Thus, this is an instance of emergent behaviour arising from the combination of multisite modification and positive feedback.

Positive feedback with co-operativity giving bistability is a well established concept. Co-operativity emerges from basic considerations in such multisite modification, and has been invoked (in concert with positive feedback) to obtain bistability in cell cycle dynamics [136]. The above results indicate other essentially different ways in which multisite modification can reinforce positive feedback.

**Positive feedback as a tool for creating, tuning and combining characteristics.** Positive feedback impacts biphasic behaviour (Figure 4.2C) in a specific way: the peak of the biphasic response is maintained, as is the asymptotic response for large signal. The signal value at which the peak is attained is however shifted towards lower signal values and eventually the biphasic response is eliminated. These results are also established analytically (refer appendix C, section 2.4). Positive feedback can also result in biphasic behaviour being combined with bistability. Positive feedback can result in the shifting of the range of oscillations, and a progressive reduction and eventual elimination of the oscillatory range (Figure 4.2F).

**Negative feedback.** We examine the case of negative feedback through the fully modified phosphoform (Figure 4.3).

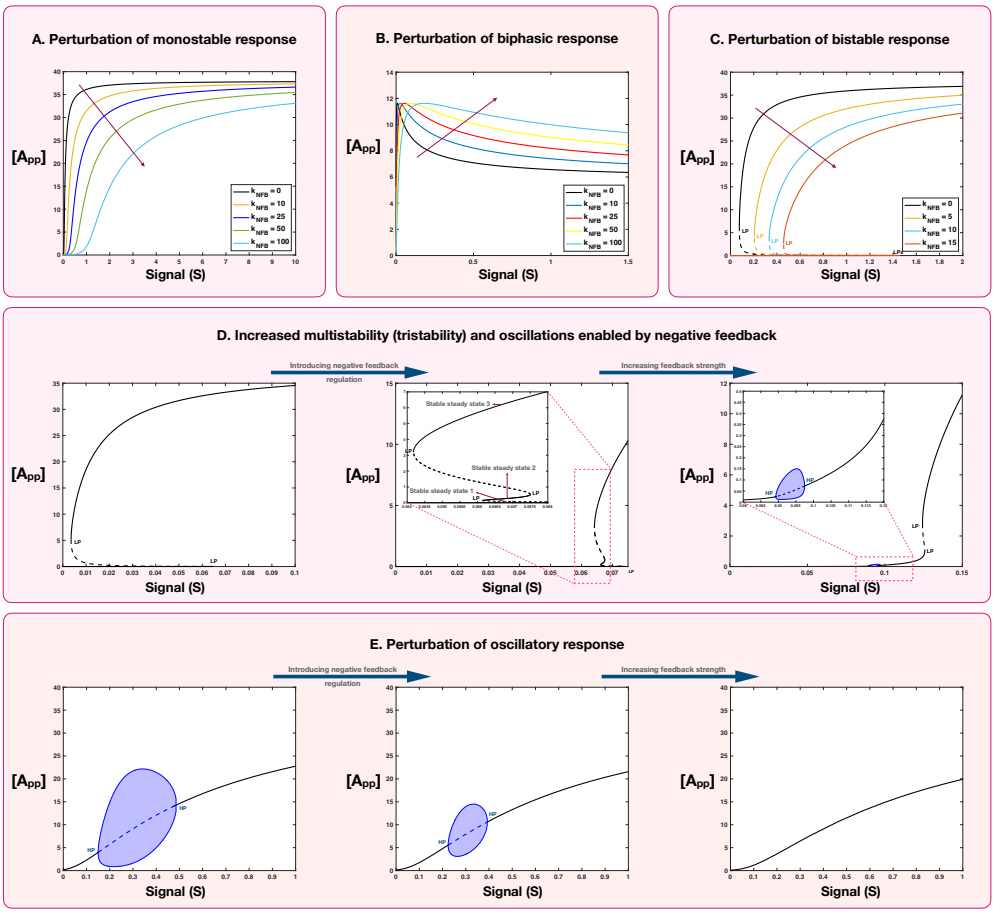



**Figure 4.3: Impact of negative feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system.** The results shown below demonstrate the effect of negative feedback on different basal characteristics of the double site modification system. (A). Perturbation of monostable basal response; eroding switch like response. (B). Perturbation of biphasic response by increasing feedback strength. Results suggest that the peak concentration of $A_{pp}$ remains the same as feedback is changed, while the location of the peak shifts towards higher signal values. (C). Negative feedback can impact bistable responses by reducing the range over which they occur. (D). Other ways in which negative feedback can impact basal bistable behavior is by the creation of tristability (a behavior impossible in the basal substrate modification system) as well as the creation of combined oscillatory and multistable responses. (E). Negative feedback can shift and reduce the range of oscillations leading to its eventual abolishment. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during oscillations. HP: Hopf bifurcation, LP: Saddle node bifurcation.]

**Negative feedback can result in an increase in the maximal number of steady states possible.** Negative feedback can accentuate threshold-like behaviour while diminishing switching behaviour (Figure 4.3A). Negative feedback also has the effect of shrinking the range of bistability (Figure 4.3C), in the case when the multisite modification system is in the bistable regime. Unexpectedly however, a negative feedback can also result in tristability as seen in Figure 4.3D. In fact it is possible to observe tristability

along with zones of oscillations both alongside and in the absence of multistability (see fig. 4.7). This is particularly noteworthy since (i) the multisite modification system in isolation is bistable, and cannot exhibit tristability (ii) the negative feedback has resulted in an expansion in the (maximal) number of steady states attainable. This feature encountered here is not seen if the modification system has only a single site. Analytical results provide further insights into when this occurs.

The general expectation of feedback (based on many studies of networks) is that positive feedback can increase the (maximal) number of steady states of a system, while negative feedback plays a dampening/homeostatic role. The ability of negative feedback to increase the maximal number of steady states suggests the need to carefully reassess the role of negative feedback. Associated experimental observations may be incorrectly inferred as a positive feedback, while in fact being a confluence of negative feedback and modification complexity.

**Feedback regulation can tune biphasic responses while maintaining their amplitude.** The perturbation of the biphasic response (Figure 4.3B) shows that the biphasic behaviour appears to be eroded by negative feedback. However a careful computational examination complemented by analytical work reveals that the biphasic behavior exists as before, and that the amplitude of the peak of the response is fixed, independent of the feedback, as is the asymptotic value for large signal values. However the value of signal at which the peak is attained is progressively pushed to higher signal values. Thus if we focus on a fixed range of signal, negative feedback can eventually eliminate the biphasic response.

Observations from both positive and negative feedback indicate that certain key features (peak and saturation concentration expressed by maximally modified substrate) of biphasic response are robustly maintained, while other features can be varied. This suggests that such feedback could be a valuable tuner of biphasic responses.

**Reassessing the role of negative feedback in generating oscillatory behaviour in networks.** Negative feedback can shrink the range of oscillations, eventually destroying them. This is an instance of how negative feedback which may be seen as helpful in generating oscillations actually works in the exact opposite way. Thus ascribing oscillations observed in data (eg Figure 4.3E (middle panel)) to the underlying negative feedback would be fundamentally misleading. It reinforces the point that certain emergent effects can result from unintuitive combinations at two distinct levels. This means it is insufficient to study features at only one level (network level say), with the expectation that details at another level (modification level say) could be incorporated after the fact.

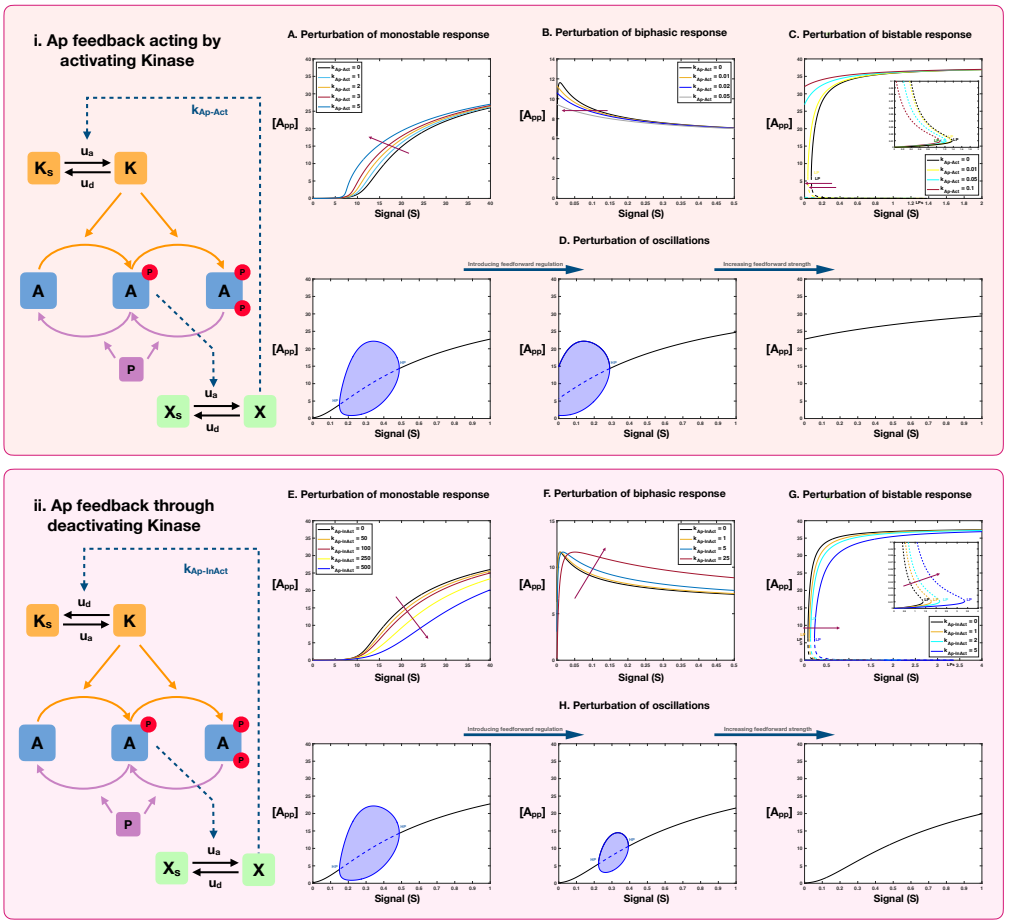



**Figure 4.4: Impact of feedback mediated by the partially modified substrate $A_p$ (through activation or inactivation of kinase) on the behavior of the double site modification system.** In both these cases, the feedback regulation which ensues is a mixed feedback (combination of positive and negative feedback). (A-D) Impact of $A_p$ feedback acting by activating kinase on (A) monostable response (showing strengthening of switch-like behavior), (B) biphasic response (showing biphasic peak moving to lower signal ranges and the eventual erosion of the behavior), (C) bistable response (introducing irreversible switching) and (D) oscillatory response (showing abolishment of oscillations with increasing feedback). (E-G) Impact of $A_p$ feedback acting by deactivating kinase on (E) monostable response (showing reduced switch-like behavior), (F) biphasic response (the biphasic peak moving towards higher signal ranges), (G) bistable response (increasing the range but reducing the switching amplitude) and (H) oscillatory response (increasing feedback reduces range of oscillations and leads to its eventual abolishment). The two sets of results above indicate the contrasting effects of mixed feedback regulation by the partially modified substrate and indicates that the specific way in which the mixed feedback is realized has a strong impact on resulting behavior. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during oscillations. BP: Pitchfork bifurcation, HP: Hopf bifurcation, LP: Saddle node bifurcation.]

**Substrate modification chemistry can enable the creation of mixed feedback regulation.** One of the basic characteristics of a multisite substrate modification system is that there are multiple phosphoforms possible, each capable of being a mediator of feedback. A basic examination of this reveals the distinct feature which emerges. A

feedback mediated by the partial phosphoform $A_p$ acts as a mixed positive and negative feedback. This is true whether the nature of the feedback was a mirror of the positive feedback or that of the negative feedback considered above. This simply relies on the fact that the same kinase both produces $A_p$ (from $A$) and converts it (to $A_{pp}$). $A_p$ activating the kinase results in a mixed feedback, wherein it activates both its production (from A) and conversion to $A_{pp}$. $A_p$ inactivating the kinase results in a different mixed feedback, wherein it inhibits both its production (from A) and conversion to $A_{pp}$.

The feedback from $A_p$ activating the kinase exhibits basic features of a positive feedback (see Figure 4.4i); for instance the transition of a basic threshold response to a monostable switch-like behaviour, perturbing bistability to generate irreversible switching, the shifting of the peak of a biphasic response towards low signal levels and the shifting of the oscillatory range towards low signal values and eventually destroying it. We have not however observed the creation of tristability with such a feedback. In a similar way, the feedback from $A_p$ deactivating the kinase (see Figure 4.4ii) erodes potential monostable switch like behaviour, reduces the amplitude of bistable switching, and shrinks and eventually destroys the range of oscillations. In this case, analytical results reveal that tristability cannot be obtained under certain conditions. Analytical work pinpoints further clear-cut differences between feedback mediated by the fully-modified phosphoform and the partially modified phosphoform,

A greater number of modification sites allows ample scope for such mixed feedback to occur in networks. Mixed feedback (resulting from feedback mediated by partial phosphoforms through the enzymes) can however be eliminated by having distinct enzymes for each modification step. Thus an apparently minor detail (intrinsic to a node at the network level) of modification biochemistry can have a strong impact on the network regulation.

**Feedforward regulation.** We now consider feedforward regulation wherein, a signal regulates the activation/deactivation of two enzymes simultaneously. For any fixed signal value, at steady state, the system exhibits a response corresponding to the behavior of the system in isolation (perhaps for a different parameter value, see [78]). Thus the different types of feedforward regulation in many respects (especially steady state behaviour) amount to the signal regulating the system, traversing specific paths in the space of isolated system behaviours.

**Incoherent feedforward regulation.** An incoherent feedforward regulation applied to a simple covalent modification cycle results in different types of adaptive behaviour - exact adaptation, under specific conditions, and inexact adaptation which can manifest as over-adaptation and under-adaptation. What impact does it have on multisite modification?

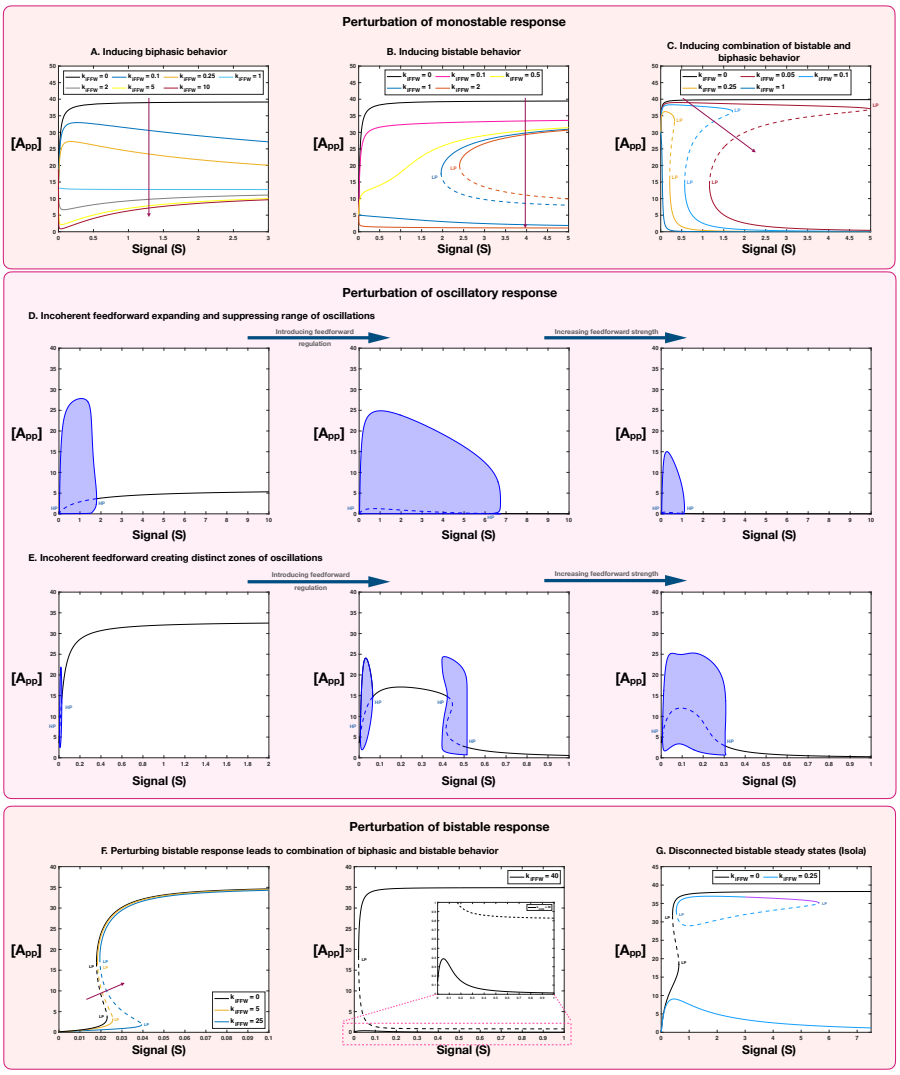



**Figure 4.5: Impact of incoherent feedforward network regulation on the double site modification system.** In all cases we vary the kinetic parameter associated with the signal regulation of the phosphatase as a way of probing the impact of incoherent feedforward regulation and assess its impact on the dose-response curve. Note that if this parameter is 0, there is no incoherent feedforward regulation. (A-C) Perturbation of basal monostable behavior. (A) shows how perturbation of a monostable response by feedforward regulation can generate a sequence of transitions involving generation of biphasic response and its subsequent inversion. (B) shows how feedforward regulation can result in bistable responses. Interestingly feedforward network regulation is not associated with multistability and this behavior demonstrates how the substrate modification system can introduce new behavior to the network. (C) shows how feedforward regulation can lead to a combination of bistable and biphasic response. (D-E). Perturbation of oscillatory response by feedforward network regulation. (D) shows how feedforward regulation can both strengthen oscillations (increasing range over which oscillations are observed) and can destroy them, while (E) shows how feedforward regulation can lead to distinct regions of oscillations emerging, which can coalesce into a single oscillatory region. (F-G). Perturbation of basal bistable behavior. (F) shows how basal bistable response can be strengthened by increasing feedforward strength and can also lead to a combination of biphasic and bistable response. (G) shows how basal bistable response can be perturbed to give disconnected closed loop steady steady state branches (Isola - refer to main text for further discussion). [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during such oscillations. BP: Pitchfork bifurcation, HP: Hopf bifurcation, LP: Saddle node bifurcation.]

**Incoherent feedforward regulation can introduce both biphasic and bistable dose response characteristics.** Incoherent feedforward regulation results in a number of non-trivial transitions (Figure 4.5). If the multisite modification system in isolation exhibits behavioural characteristics of a monostable switch, or even a simple hyperbolic response, then the incoherent feedforward regulation can result in biphasic behaviour. Furthermore, tuning the relative strengths of the arms of the feedforward regulation can even reverse a biphasic response with a subsequent increase and plateauing (Figure 4.5A). The incoherent feedforward regulation can also result in the introduction of bistability, which would not have occurred without the signal regulating both kinase and phosphatase. This also shows how a signal regulating an antagonistic enzyme can promote not only biphasic behaviour but bistability as well (Figure 4.5B). In fact it can allow for the creation of a combination of both behaviour together in the dose response curve (Figure 4.5C).

An incoherent feedforward network is associated with in-built competition, and is, as such, hardly a candidate associated with bistability (which involves reinforcement and co-operative effects). The above results show that (a) this is possible and (b) characterizing the network and the modification system (for the same parameters) in isolation provides no indication of how such behavior may be generated by incoherent feedforward regulation.

**Incoherent feedforward regulation can create, merge and separate zones of oscillation.** We found that feedforward regulation could impact oscillations by increasing the range of oscillations, and then dramatically reducing it and even eliminating oscillations (Figure 4.5D). Another way in which the feedforward regulation may impact oscillatory behaviour is by creating distinct regions of oscillations in the dose response curve which eventually merge (Figure 4.5E). This again highlights how feedforward regulation can enhance robustness of oscillations.

**The creation of isolas.** Incoherent feedforward regulation can significantly shape and regulate bistable behaviour. This includes an expansion of the range of bistability and the introduction of simultaneous bistable and biphasic behaviour (See Figure 4.5F). Another distinct outcome, is the creation of an isola (an isolated closed curve in the dose response plot), which is a different manifestation of multi-stationarity, which has not been observed (to our knowledge) thus far in double site modification (See Figure 4.5G). Isolas have been observed in different scientific contexts including complicated multisite modification networks [165, 166]. It could play a role in biological settings, by creating distinct new states which may not be accessible by conventional switching, but may be realized via strong perturbations of the system.

**Robustness and homeostasis.** Incoherent feedforward networks are associated with homeostatic responses to signal. If we start with a basal symmetric multisite modification, incoherent feedforward regulation can be used to generate symmetry-breaking and the creation of absolute concentration robustness (ACR) thus allowing for the net-

work to present a suite of homeostatic responses including both signal homeostasis (itself expanded) and ACR (see fig. 4.8 and discussion in section 4.5).

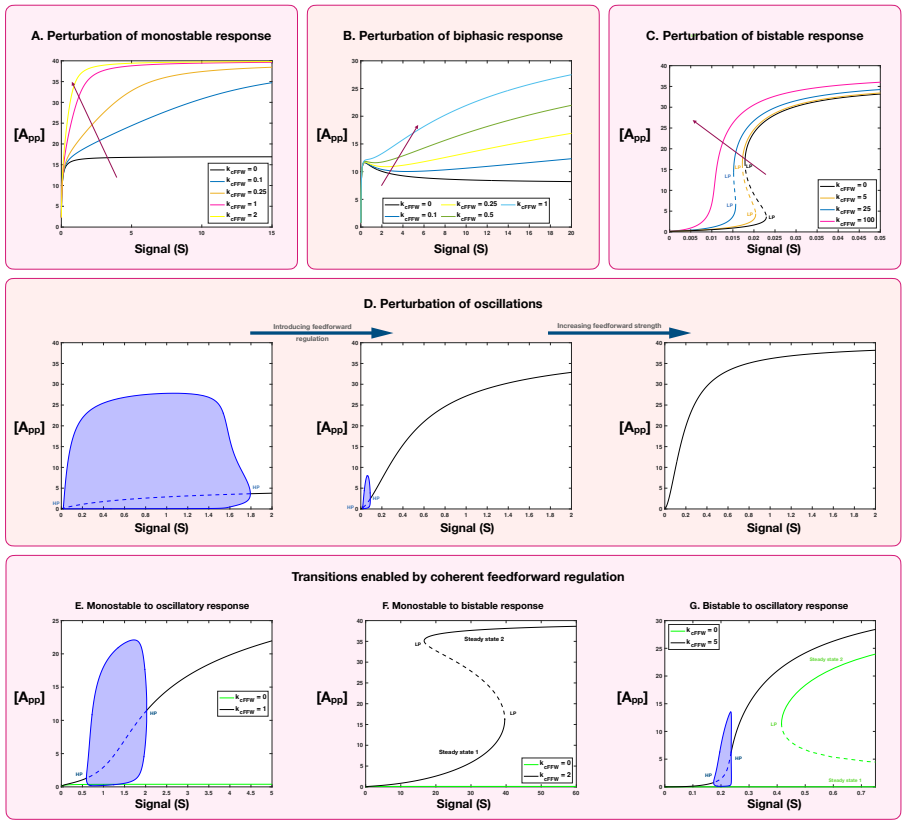



**Figure 4.6: Impact of coherent feedforward network regulation on the double site modification system.** In all cases we vary the kinetic parameter associated with the signal regulation of the phosphatase as a way of probing the impact of coherent feedforward regulation and assess its impact on the dose-response curve. Note that if this parameter is 0, there is no coherent feedforward regulation. (A). Perturbation of basal monostable response by coherent feedforward regulation (indicating an elevated response - as may be expected from the reinforcement effects of coherent feedforward regulation). (B). Perturbation of basal biphasic response by coherent feedforward regulation, demonstrating an erosion of biphasic response and a transition via multiphasic responses to monotonic dose-response. (C). Increasing feedforward regulation can lead to a suppression of bistability. (D). Coherent feedforward regulation can lead to a significant reduction in the range of oscillations and its eventual suppression. (E-G). Transitions enabled by coherent feedforward regulation. (E) shows how coherent feedforward regulation can enable a transition from monostable response (green) to an oscillatory response (black). (F) shows how monostable response (green) can transition to multistability (black), and (G) shows how a basal bistable response (green) can be eliminated with the creation of oscillations (black). (E-G) demonstrates how the simple reinforcement of kinase and phosphatase can result in dramatically different behavior from the basal dose-response (wherein only the kinase is regulated). This indicates that the effect of coherent feedforward regulation is not one of simple reinforcement. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during such oscillations. BP: Pitchfork bifurcation, HP: Hopf bifurcation, LP: Saddle node bifurcation.]

**Coherent Feedforward regulation.**

**Coherent Feedforward regulation is much more than a simple push-pull network.** The simultaneous regulation of kinase and phosphatase to reinforce one other is suggestive, at the outset, of a simple outcome, corresponding to more efficient kinase regulation. However a closer look reveals that such a picture is simplistic, to say the least. While this simple insight can explain the strengthening of the amplitude of a hyperbolic response (and a monostable switch) as seen in Figure 4.6A, we also find the erosion of a biphasic response over the range of input where it occurred (while also presenting multiphasic response in the transition, see Figure 4.6B) and the erosion and eventual destruction of bistability (Figure 4.6C). As shown in Figure 4.6D the feedforward regulation can significantly affect the range of oscillations, reducing it and even eliminating it.

**A reinforcement of modification directions is capable of sophisticated control regulation.** The versatility of even the simple coherent regulation is seen in Figure 4.6(E,F), where we see that it can orchestrate multiple types of transitions, from monostable behaviour to both oscillatory and bistable behaviour and also perturb oscillatory behaviour to create bistability; the introduction of bistable behaviour here is in contrast to its tendency to shrink and destroy bistability, seen above. The creation of multiphasic responses from biphasic response is another instance of this (fig. 4.6B). This demonstrates the unexpected impact that simple co-regulation of enzymes reinforcing one and other can have.

Taken together, both coherent and incoherent feedforward regulation afford simple ways of exploring the landscape of the modification chemistry and accessing a range of behaviour, even by varying one (experimentally accessible) parameter.

### 4.3.2 Analytical results.

Analytical results complement the computational results above, providing binary answers to questions underpinning the interplay of the two levels (see appendix C further details). This section can be skipped by readers not interested in the details.

**Summary of analytical results.** Analytical work focusses on steady state behavior and involves systematically reducing the set of model equations at steady state, eliminating most variables, and studying the reduced equations. Analytical work either reveals/explains the presence of a behaviour, or proves its absence. In the latter instance this is established for a broad a range of system parameters as possible:the only restrictions (all in the multisite modification) being either kinetic restrictions in certain steps of enzymatic modification (assumed to be in the unsaturated limit) or total enzyme amounts being much less than total substrate amounts

A common thread in all the analytical work is the establishing of boundaries between the possibility and impossibility of a given behaviour in either the full system or a

simplification/restricted regime thereof. The sharp delineation of structural features which enable or prevent a given behaviour reveal minimal structural ingredients for the manifestation of the behaviour.

**Feedback networks and multistability.** We address the fundamental question of how feedback impacts the number of steady states, using two approaches: (a) focussing on enzymatic regimes of specific modification steps being in the unsaturated limit which correspond to the absence of corresponding enzyme-substrate complexes (b) answering associated questions without making any such simplifying assumptions.

As an example of the first approach, we focus on a series of cases, starting with all enzymatic modification steps in the unsaturated limit (i.e. no complexes) where the behaviour is simple, and then progressively relaxing this assumption. Doing this allows us to isolate key structural requirements for different behaviour (here for simplicity, we assume feedback is directly mediated by the relevant substrate, without an intermediate stage): 1. If all complexes are removed, the modification system is strictly monostable. Here, a positive feedback mediated by $A_{pp}$ can generate bistability, but a negative feedback cannot. Furthermore feedback mediated by $A_p$, precludes bistability, whether the feedback activates or inactivates kinase (analogues of positive and negative feedback from $A_{pp}$). 2. Now we consider the case where only one enzyme-substrate complex is present in the substrate modification system. If this involves $A_p$, then bistability is precluded in the double site modification system (this continues to be true even if both $A_p$ complexes are present;on the other hand complexes associated with substrates A and $A_{pp}$ readily allow for the possibility of bistability). In this instance (a) Positive feedback from $A_{pp}$ can result in bistability as expected, but negative feedback is still incapable of doing so. (b) Feedback from $A_p$ activating the kinase is capable of inducing bistability if the only complex present is $A_pP$, but not if the only complex present is $A_pK$. This highlights how the presence of specific complexes in conjunction with feedback can create conditions for bistability, and also highlights the contrast between feedback through $A_p$ and $A_{pp}$. Feedback mediated by $A_p$ deactivating the kinase will not induce bistability. Both instances of feedback mediated by $A_p$ constitute mixed feedback, but a clear delineation between these two cases is seen in this simplified setting. (iii) In the case where both $A_p$ complexes are present, negative feedback (involving deactivating kinase, mediated either by $A_{pp}$ or $A_p$) is incapable of inducing bistability. A detailed synthesis of the impact of feedback on these systems where certain complexes are (essentially) absent is discussed in appendix C (and summarised in section 4.5).

The second approach is used to examine how feedback can generate tristability, a behaviour not possible in the double site modification system. In all cases, unless otherwise mentioned, we examine feedback mediated via an extra covalent modification step (see Models and Methods). 1. We show that a positive feedback mediated by $A_{pp}$ to the kinase activation can enable tristability, and this can happen even without the extra covalent modification step in the feedback loop. 2. Analytical work shows how

tristability may be generated by negative feedback. In both the positive and negative feedback cases, tristability is possible even when total enzyme amounts are much less than substrate amounts. 3. We analyze negative feedback mediated by either $A_p$ or $A_{pp}$, when total enzyme amounts are much less than total substrate amounts and the extra covalent modification step is absent, and show that tristability is precluded irrespective of systems parameters.

**Regulation via feedforward networks.** To what extent can feedforward networks achieve transitions in behavior between monostability and bistability in multisite modification? 1. At steady state for any fixed signal, the steady state(s) of the system with coherent/incoherent feedforward regulation can be mapped to the steady state of the isolated double-site modification module, with altered parameters. 2. If the multisite modification system in isolation satisfies an equation (involving catalytic constants) which enables multistability (meaning it can be attained for some values of total concentrations of enzymes and substrates), then the introduction of feedforward regulation (whether coherent or incoherent) will still enable this. 3. There exist criteria for the preclusion of multistability for the multisite modification system in isolation. Now using this, we can establish conditions under which a system which was in the "exclusion region" for multistability, remains so when the signal is varied. This provides an explicit analytic criterion guaranteeing monostability (and prevention of transitions to multistability) as the signal is varied. 4. In the same way, we can also establish conditions under which the system can exit the "exclusion region" and provide a basis for determining the minimum signal required to do so. 5. Using analysis in the low enzyme limit, we demonstrate that the incoherent feedforward network case can exhibit isolas (explaining computational results), while the coherent feedforward network will not (and neither will the system in isolation).

**Biphasic responses.** Multisite modification systems in isolation can generate biphasic responses (in the maximally modified phosphoform) to varying total kinase amount or signal. We examine how feedback perturbs this. For biphasic responses to change in signal, there is a range of feedback strength over which the biphasic response is maintained (with the peak and saturation value of $A_{pp}$ remaining fixed). However, a sufficiently strong positive feedback will move the biphasic location to negative signal values, implying an abolishment of the biphasic response. Negative feedback, by contrast, pushes the location of the biphasic response to higher signal values.

While analytical approaches (especially those which establish results independent of most parameters) provide a number of insights, they are limited by their applicability to certain model simplifications or kinetic regimes. Semi-analytical approaches, can expand the applicability of analytical approaches on one hand, while bringing nontrivial analysis to purely computational explorations.

**Semi-analytical approaches.** Reduction of the steady state equations results in either single or coupled polynomials, whose coefficients contain all system parameters. While

in some cases this may be amenable to analytical work directly, simplifications are often needed. However with semi-analytical work, it is possible to combine computation and analytical work to focus on the effect of key network features, which are present as a small number of parameters. The computational component arises here either in the specification of some of the systems parameters (for eg. user supplied) and their basal exploration (which is then built upon by analytical work), or in a fuller exploration of the analytical equations obtained, using computation. Either via a study of the polynomial(s) directly or the steady state Jacobian, it is possible to infer when additional steady states may be created via network regulation, and also when they are guaranteed not to do so. From this, multiple conclusions can be drawn, for instance by computationally exploring a basal parameter set and finding it to be monostable (a) we can obtain explicit bounds in the network parameter, which ensures that the system remains monostable (b) we can find ranges of network parameters where additional feasible steady states may be created, and computationally verify when (and if) those may be obtained. (c) Similar conclusions can be drawn if the initial state is multistable. These insights can be obtained for both feedback and feedforward networks (see appendix C for exemplar cases).

## 4.4 Discussion

Multi-levelness is a defining characteristic of living systems, generally, and a hallmark of cell signalling networks. At one level, information is processed through pathways of biochemical reactions, with recurring regulatory structures (eg feedforward/feedback regulation), which (along with associated non-linearities) contribute to the qualitative complexity. At another level, the functionality of proteins is regulated by protein modifications, and the protein modification status (analogous to a molecular bar-code) determines its function [21]. Qualitative complexity here emerges from ingredients such as sequestration, shared enzymes, scaffolding, multisite modification. The interplay of these levels is thus a theme of broad interest which stems from the fact proteins which are key transmitters of information, are often modified in different ways to achieve different functionalities. We investigate this theme in the instance of multisite modification and feedforward/feedback regulation.

Our systems analysis probed the origins and parametric dependence of emergent behaviour/behavioural transitions which resulted from the interplay of the two levels. By employing a combination of computational approaches (directly revealing transitions), analytical work (establishing when transitions and specific behaviour may and may not occur) and semi-analytical work bridging the two approaches, we reveal the multi-level systems landscape with a variety of associated insights. The broader approaches we employ may be fruitfully used for systems dissection of other multilevel systems.

**Transitions between different qualitative behaviour.** Our results show that the am-

bient network can facilitate transitions to different types of qualitative behaviour in multisite modification. Generally speaking, we find that such transitions are feasible across many parameter sets of the multisite modification module, and different types of behaviour may be obtained depending on the nature (and parameters) of the network. We reveal structural restrictions which may prevent the possibility of transitions from occurring in certain parameter regimes. The network can also induce behaviour not intrinsically observed in substrate modification (eg tristability in feedback networks, isolas in incoherent feedforward networks). Similarly, multisite modification introduces behaviours to the network, which would otherwise not be intrinsically observed (irrespective of parameters), for instance bistability in feedforward (both coherent and incoherent) and negative feedback networks. Since, this is due to sequestration effects in the multisite module, rather than any explicit positive feedback, we see that bistability emerges in the overall network without any obvious positive feedback present.

**System behaviour and its origins.** Reversing the perspective, and starting from the outcome end, and tracing the ingredients responsible, yields further insights. If a behaviour is intrinsically observed in the multisite modification module, the network can either make it more or less robust. If a behaviour is enabled by both multisite modification and the network, the combination is not guaranteed to make the behaviour more robust: in fact it can lead to the behaviour being abolished (seen in the case of biphasic behaviour and the incoherent feedforward network, or oscillatory behaviour and negative feedback for instance). This "cancellation effect" implies that simple behaviour of the overall system may not necessarily be associated with simple behaviour of the network and the multisite module (a plausible assumption based on the overall system behaviour). In other cases, the behaviour needs a combination of contributions from both multisite modification and the network, which may however be unintuitive (eg. tristability arising from negative feedback and multisite modification with intrinsic bistability). These results reveal pitfalls in ascribing experimentally observed behaviour in signalling networks to certain factors without a careful dissection of both substrate modification and network levels.

**Guarantees for maintenance of behaviour.** A complementary aspect of behavioural transitions is establishing guarantees that certain transitions in behaviour do not occur, perhaps for certain ranges of network parameters. Analytical and semi-analytical approaches can provide stronger complementary results to direct computation. For instance in exploring potential transitions from single to multiple steady states, it is possible to analytically establish a range of network parameters where a change in the number of steady states is guaranteed not to happen, as an explicit function of the multisite modification parameters. Semi-analytical approaches expand the applicability of such approaches. A similar approach is applied to biphasic dose responses.

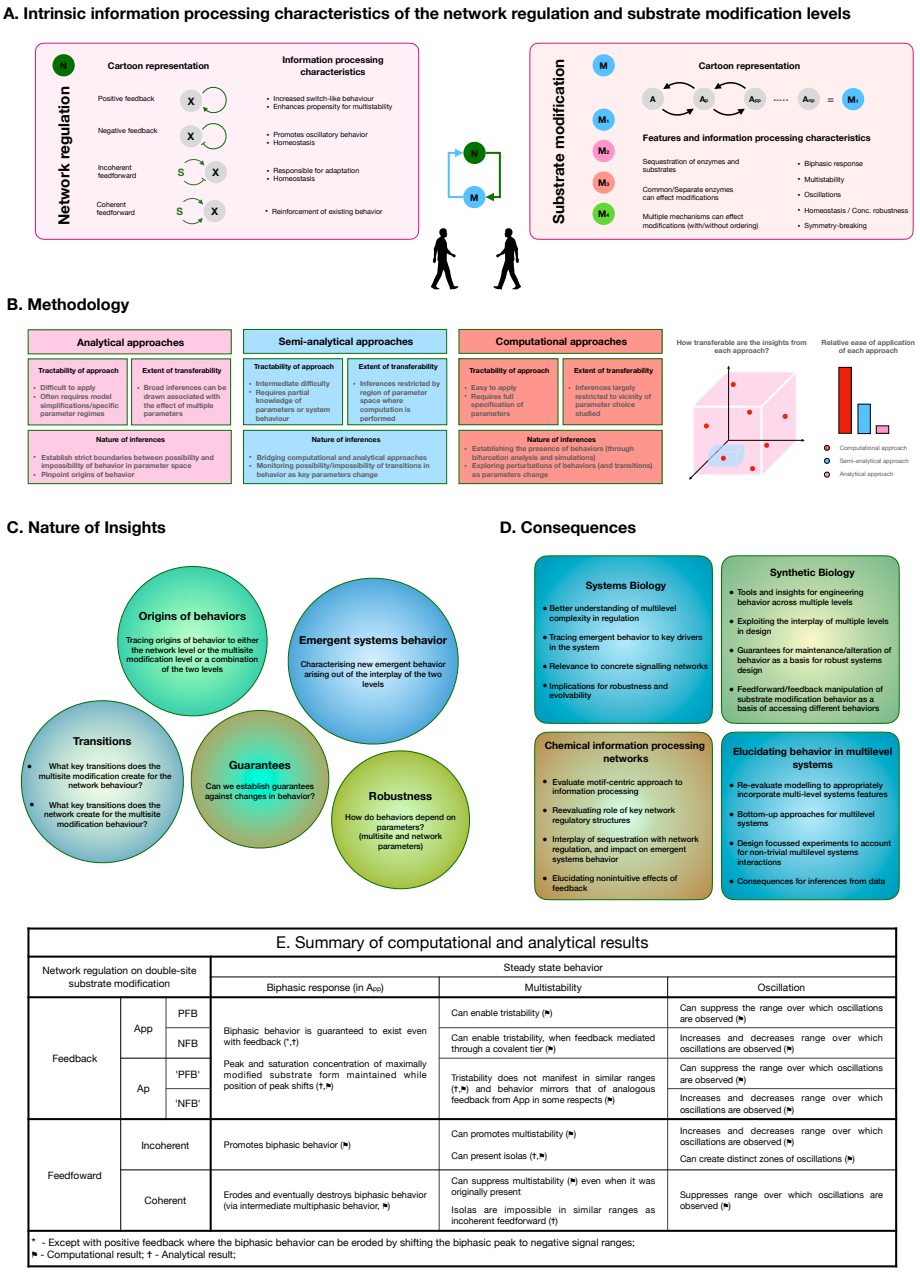
**Chemical modification complexity meets network complexity.** What do the network and substrate modification levels contribute to each other? The network can enable fundamentally new behaviour (eg. tristability) and establish distinct types of transi-

tions between behaviour intrinsically observed in multisite modification. Feedback regulation can critically depend on the specific phosphoform mediating the feedback. Feedforward regulation provides an appealing way of making transitions between different behaviours. This includes accessing a range of behaviour with just one or two experimentally-manipulatable signals, something beneficial in synthetic biology. Multisite modification can confer distinct new behaviour to the network including bistability and oscillations in feedforward networks (otherwise absent). The presence of multisite modification can allow negative feedback regulation to increase the number of possible steady states, something which fundamentally goes against the intuition of negative feedback. More generally the multisite module can cause a particular feedback network to behave like a feedback network of a different type. The root cause for this is sequestration effects in substrate modification, which is (a) internal to a node in the network description and (b) cannot be simply represented as an appropriate subnetwork in the network description. The creation of distinctly non-adaptive behaviour in incoherent feedforward networks is another example of the impact of multisite modification. Overall networks and multisite modification can combine in distinctive and unexpected ways.

**Consequences.** These insights have distinct implications for the "modularity" of such networks and the extent to which they could be understood, studying one level at a time, whether experimentally or theoretically. On one hand, it indicates the need for focussed experiments designed to explicitly probe the interaction of levels (something accessible, since there are plenty of experiments probing each level separately). It also indicates the need to carefully manage and combine data from studies of each of these levels (also see [167]). There are consequences for modelling and theoretical approaches as well. While a motif is a suitable vantage point for investigating core building blocks of networks in general, it will be necessary to investigate motifs which incorporate multi-level information (eg. core network structures with key building blocks of substrate modification, say). Our studies use models which represent multi-level building blocks, since they explicitly incorporate building blocks at two levels. Thus they represent a bottom-up approach to dissecting the multi-level interaction between substrate modification and network regulation. Such combined theoretical and experimental analysis can provide important insights into when substrate modification may or may not need to be described explicitly in complex networks.

We now discuss the consequences of our results for systems biology, synthetic biology and chemical information processing.

**Systems Biology.** Multisite modification is a key constituent of signalling networks, and understanding how such networks function involves understanding the contribution of multisite modification. There are multiple instances of this, for example *ERK/MAPK* signalling with feedback [107], circadian oscillators [120, 161], Sic1 regulation in the cell cycle [131]. There are already multiple studies of *MAPK* signalling with feedback; multisite modification is believed to be a key driver in the generation of




E. Summary of computational and analytical results

| Network regulation on double-site substrate modification | | | Steady state behavior | | |
|---|---|---|---|---|---|
| | | | Biphasic response (in $A_{PP}$) | Multistability | Oscillation |
| Feedback | App | PFB | Biphasic behavior is guaranteed to exist even with feedback (*,†) | Can enable tristability (⚑) | Can suppress the range over which oscillations are observed (⚑) |
| | | NFB | | Can enable tristability, when feedback mediated through a covalent tier (⚑) | Increases and decreases range over which oscillations are observed (⚑) |
| | Ap | 'PFB' | Peak and saturation concentration of maximally modified substrate form maintained while position of peak shifts (†,⚑) | Tristability does not manifest in similar ranges (†,⚑) and behavior mirrors that of analogous feedback from App in some respects (⚑) | Can suppress the range over which oscillations are observed (⚑) |
| | | 'NFB' | | | Increases and decreases range over which oscillations are observed (⚑) |
| Feedfoward | Incoherent | | Promotes biphasic behavior (⚑) | Can promotes multistability (⚑)<br>Can present isolas (†,⚑) | Increases and decreases range over which oscillations are observed (⚑)<br>Can create distinct zones of oscillations (⚑) |
| | Coherent | | Erodes and eventually destroys biphasic behavior (via intermediate multiphasic behavior, ⚑) | Can suppress multistability (⚑) even when it was originally present<br>Isolas are impossible in similar ranges as incoherent feedforward (†) | Suppresses range over which oscillations are observed (⚑) |

* - Except with positive feedback where the biphasic behavior can be eroded by shifting the biphasic peak to negative signal ranges;
⚑ - Computational result; † - Analytical result;

**Figure 4.7: Conceptual summary of the multi-level interaction between network regulation and substrate modification levels studied, the methodology used, and a synthesis of results and insights.** A. **Network regulation meets multisite substrate modification.** The non-intuitive interplay of these two levels is studied here. Network regulation is an integral component of cellular signaling and is responsible for enabling key qualitative behaviors such as homeostasis, multistability and oscillations (network topology and non-linearities are key ingredients here). Multisite substrate modification on the other hand provides proteins specific functionality and is itself capable of complex signal processing capabilities (arising primarily from sequestration effects). B. **Methodology.** A three-pronged methodology is adopted to tackle the complexity of multi-level interactions. Each approach comes with internal trade-offs between easy applicability and transferability of results across parameter space, and is best suited to provide insights of a specific nature.

C. **Nature of insights.** Our study provides a number of insights ranging from exploring origins of emergent systems behavior to studying transitions, robustness and guarantees of behavior exhibited by the overall system. D. **Consequences.** Our results have consequences in a number of specific contexts. We show non-trivial emergent systems behavior and non-intuitive roles of network regulation of relevance in systems biology and chemical information processing. By delineating origins of behaviors and studying key drivers of behavior across multiple levels our study provides results relevant to systems biology. Its relevance to synthetic biology is borne out by the insights it provides to the engineering of multi-level coupled network interactions while guaranteeing robust designs for specific outcomes. The multi-pronged analytical, semi-analytical and computational approach utilized provides insights to modelling, elucidating and designing of experiments for probing multi-level interactions, and in formulating appropriate data inference methodologies. E. **Summary of computational and analytical results.**

circadian oscillations [168, 161]; there are models of specific systems such as *Sic1* signalling or *Tau* protein regulation [24, 131]. In such cases, our study provides a systematic platform from which to examine the behaviour of the system, and how it emerges from the contribution of multisite modification. In these cases, there are other aspects of the networks involved as well (eg. *MAPK* has an additional layer with double-site modification, circadian oscillators have additional network layers). It is therefore especially important to isolate how different overall behaviour can emerge, and which components/interactions may be key contributors. Our results provide insights and testable predictions into stripped down versions of these systems, and can be used as a template with further augmentation to systematically dissect these systems.

How does our approach apply to other systems, in natural or engineered biology, which may involve the interplay of different kinds of networks and multisite (or other complex) modification systems? Our analysis of ordered double-site modification contributes a number of direct insights into the behaviour in networks, including emergent behaviour, impossibility of behaviour, effects of parameters, essential trends and underlying principles associated with interplay of levels, as well as conceptual and methodological insights. Firstly with regard to other networks (eg composite feedback systems, or combinations of feedforward and feedback regulation), these can be analyzed in a similar way building on the results/approach here. With regards to substrate modification, for investigations of ordered double site modification our results provide direct insights and a direct platform for expansion and use in other contexts (eg *p53*, *MAPK*) which allows us to systematically assess when the additional factors present make an essential difference. In other cases with a modest number of modifications sites (ordered or random mechanisms), a similar approach can be used: while some of the insights translate to those systems as well, further exploration of the multisite modification, along with the methodology and approach here (including the use of semi-analytical work, complementing computational and analytical work), will be very useful. In all these cases, inputs from experiments for either constraining parameters or testing behaviour are especially useful: in this regard our approach allows us to develop specific experiments to discriminate between different possibilities (of modification system and network) which give rise to the same systems behaviour. Finally

when the number of sites becomes large, an approach which employs simplifications, informed by the experiments/known biochemistry will be the most useful approach (as it will be for dissecting the modification system in isolation itself). which can then be combined with the approach used here.

A common methodology in systems biology is to infer the underlying networks from data [37, 169, 167]. Our study (for eg. results on multistability, oscillations and biphasic behaviour) indicates that there are multiple routes to such behaviour which can involve different contributions from multisite modification and the network. In some cases the behaviour emerges from a complex interplay of the two and may not be adequately captured by a description focussing on the network level alone: in fact a purely network level description may result in incorrect inferences drawn about network interactions (eg. an increase in number of steady states attributed to a positive feedback may in fact be due to a negative feedback, which needs a proper description of the substrate modification layer; oscillations may be attributed to a negative feedback which is actually acting to eliminate oscillations). An appropriate network model needs to account for both the substrate modification chemistry and network interactions, and require experiments which discriminate between different ways in which the behaviour could be obtained. More generally, our study indicates limitations of using simple networks/motifs and the need for employing the flexibility of multi-level descriptions in the inference process.

Tracing the origins of behaviour to key contributing components is also directly relevant to network robustness and evolvability. Networks in systems biology are products of evolution. How a given network performing an information-processing task emerged is often studied by mimicking the evolutionary process computationally / experimentally starting with simpler structures [170, 171]. These studies are done at a fixed level (either network level or chemical modification level), but both are potential sources for evolutionary tinkering. Our analysis suggests a broader question: for specific functions (especially encoded by non-trivial systems behaviour) what are the relative contributions of the network and substrate modification? While the answer could vary from system to system, it might suggest broader strategies used by biology: for example tinkering with a network as a primary determinant with the substrate modification layer being a modulator. If this is the case, this would be suggestive of a broad design principle. On the other hand, depending on the current stage of evolution, the relative roles may be reversed. Similar issues arise while evaluating robustness of behaviour of these systems: key contributors to robustness could come from either the network, the substrate modification layer or their interplay. Investigating these issues involves first recognizing the possibility of such multi-level contributions and then developing experimental and theoretical methodologies to evaluate different possibilities.

**Synthetic Biology.** There are multiple studies which have employed multisite modification in synthetic biology [65, 83]. In parallel, manipulation of chemical information

processing is being investigated in a variety of cellular and cell-free settings [172, 173]. Synthetic biology approaches combining both these elements could provide a richer repertoire for engineering. Our study provides a basis for investigating possibilities of systems behaviour and different ways of exploiting these different aspects in tandem, It could be a starting point for providing a tool-kit for developing designs for chemical modification networks, with substrate modification capability and flexibility which could be especially advantageous. It highlights engineering features of feedback regulation from different phosphoforms–for instance the natural creation of combined positive and negative feedback, and the tuning of biphasic dose responses keeping key features fixed. The versatility of feedforward regulation is demonstrated by showing how (i) it can flexibly access and switch between a multiplicity of different behaviour (using factors which are easy to manipulate experimentally), an appealing candidate for design. (ii) it can enable a suite of homeostatic responses, including both signal homeostasis and Absolute Concentration Robustness. Finally, our study of different theoretical guarantees to ensure that certain transitions in behaviour do not occur (for certain ranges of network parameters or signals) could be directly relevant in establishing robustness in the design process.

**Chemical Information Processing.** The networks we examine are examples of complex multi-level information processing networks. Their defining characteristic is that, in a network description, the nodes represent complex mass-conserved systems which are capable of rich behaviour themselves (not easily encapsulable in input-output form). Our study provides multiple instances of the subtlety and complexity of such multi-level systems and challenges in dissecting them. For instance, our analysis reveals new dimensions to basic regulatory structures like negative feedback, and indicate that actual behaviour of regulatory structures may crucially depend on apparently minor components of the regulatory pathway. Future work building on the results presented here will consider multisite modification with different enzymes, which brings in the complexity of different enzymes, different associated networks, and cross-talk due to shared substrates. Our studies in this paper (not requiring any essential sequestration effects of the network component in the substrate modification layer, or vice versa) already demonstrate the systems complexity of multi-level interactions. Extending this to examine the effects of such sequestration can build on the current study.

Our study has focussed on different aspects of the "forward problem" of how network regulation and multisite modification combine, and the emergent behaviour therefrom. The "inverse problem" of determining the behaviour(s) of the network and multisite modification which enable a particular emergent behaviour is significantly more challenging needing a broad parametric exploration, and to start with, a characterization of different types of behaviour of multisite modification in parameter space. This is an avenue for future work, and can use the present study as a platform.

All in all, the fact that cellular biochemical networks can present complex information processing both at network levels and at the (substrate) modification chemistry level

has a range of important and non-trivial consequences. These are critical for understanding different aspects cellular networks, provide new possibilities for engineering biochemical networks in synthetic biology/chemistry and provide new vistas for investigating complex chemical information processing systems. This paper providing a combination of conceptual, methodological and direct insights provides a platform for exploration in all these contexts.

## 4.5 Appendix

### Negative feedback regulation of multisite modification

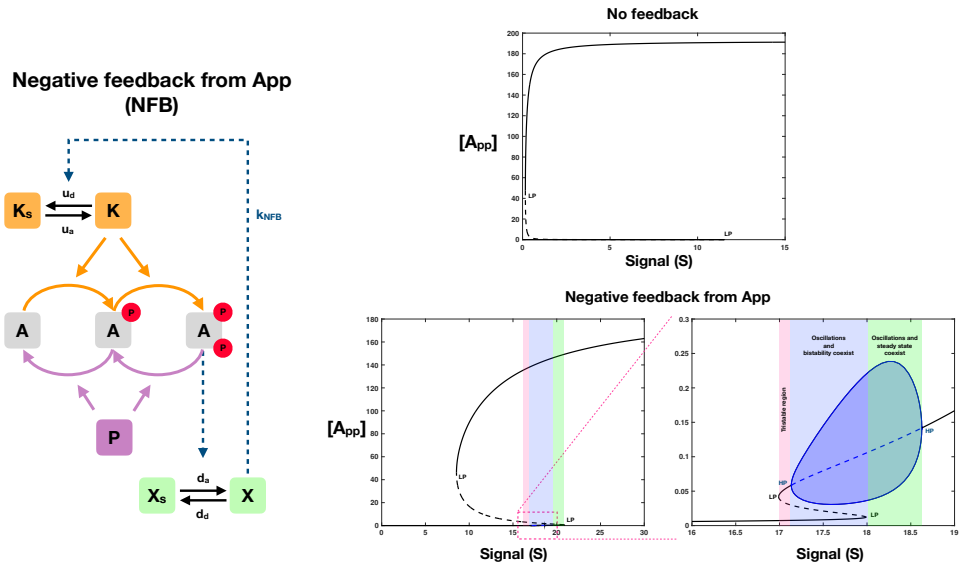



**Figure 4.7: Negative feedback from $A_{pp}$ can enable a combination of multistability and oscillations.** Negative feedback can, in addition to enabling tristability and oscillations, generate combinations of such behaviors to coexist for certain ranges of signal. Here we see negative feedback enabling the coexistence of tristability, oscillations with bistability, and oscillations with monostable steady state together in ranges of signal concentrations. [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during oscillations. BP: Pitchfork bifurcation, HP: Hopf bifurcation, LP: Saddle node bifurcation.]

## Incoherent feedforward regulation of multisite modification

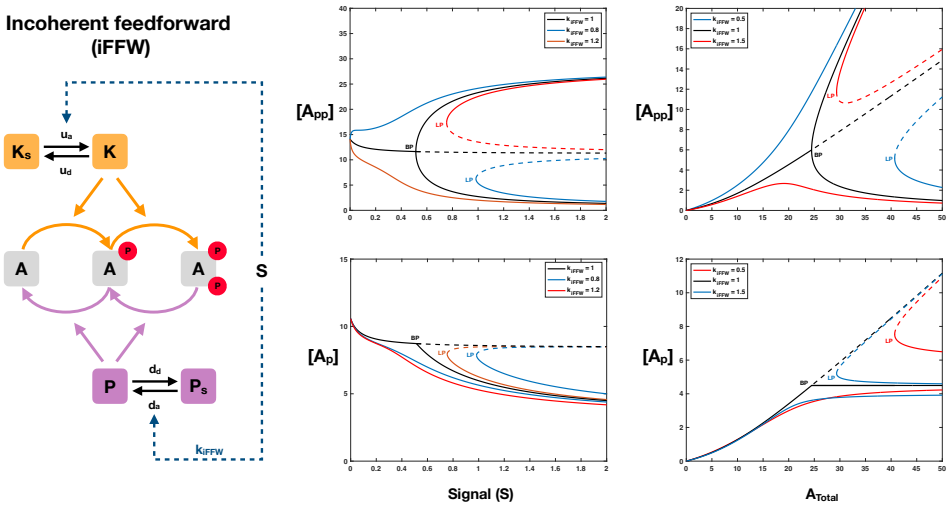



**Figure 4.8: Incoherent feedforward regulation can generate multiple homeostasis capabilities in substrate through both simple network regulation and by exploiting and tuning intrinsic substrate modification capabilities such as symmetry breaking.** We consider the case of a symmetric multisite modification system (symmetry between $A_{pp}$ and A and associated reactions) and examine the effect of incoherent feedforward regulation. Column 1 shows how incoherent feedforward regulation (itself symmetric - black curve) can generate an extended region of signal homeostasis (which expands the simple homeostatic capability of incoherent regulation - see text). However increasing the signal can lead to symmetry breaking which in turn confers new homeostatic capabilities. Column 2 shows absolute concentration robustness in $A_p$ as the total substrate concentration is varied (a distinct form of homeostasis) which emerges as a consequence of symmetry breaking (see black curve in right lower panel). [Dashed lines indicate unstable steady states, while solid lines represent stable steady states in the bifurcation diagram. Shaded regions in the bifurcation diagram indicate regions of oscillations and the blue lines indicate bounds on concentrations during oscillations. BP: Pitchfork bifurcation, HP: Hopf bifurcation, LP: Saddle node bifurcation.]

Incoherent feedforward network regulation is often associated with homeostatic responses. Incoherent feedforward network regulation acting on a single site substrate modification is capable of exhibiting exact adaptation in weak signal activation ranges (assuming high catalytic constants and thus negligible enzyme substrate complexes). In different parameter regimes, the incoherent feedforward network regulation can result in over or under adaptation.

Interestingly, this feature carries through for multisite modification. Assuming sufficiently high catalytic constants of modification, leading to negligible enzyme substrate complexes, we find that when the signal is sufficiently low (just as for the case of an incoherent feedforward regulation acting on a single covalent modification cycle) $K \propto S$, $P \propto S$ , and thus $K/P$ is a constant. Now examining the substrate modification cycles in tandem leads to the following conclusions $A_{pp}/A_p \propto K/P$, $A_p/A \propto K/P$, where the proportionality constants are independent of signal. Now we use this along with the substrate conservation condition $A + A_p + A_{pp} = A_{tot}$ to establish that the individual

substrate concentrations are independent of signal as well. In other words, in this limit, the homeostatic capability carries through to the case of multisite modification.

Elsewhere [148], we have examined an ordered double site modification system (of the type considered in the paper) which is symmetric: the symmetry is between the fully modified and the unmodified phosphoforms and the associated reactions. It was demonstrated computationally and analytically that the system can break this symmetry and lead to absolute concentration robustness in the partially modified substrate. The concentration of the partially modified substrate ($A_p$) remains fixed as the total substrate concentration is changed. This phenomenon is associated with multistability achieved through a pitchfork bifurcation.

We show here (see Figure 2), an instance of the impact of incoherent feedforward regulation on symmetric or near-symmetric multisite modification systems. The incoherent feedforward regulation can provide a homeostatic response of the substrates to varying signal, when the signal is relatively weak consistent with the behaviour of incoherent feedforward regulation as a core module of adaptation as discussed above, when all catalytic constants are sufficiently high. We now focus on the instance of symmetric incoherent feedforward regulation (symmetry between the legs of the incoherent network). Here we find an expanded capacity for homeostatic response to signal (discussed further below, see Figure 2 column 1, black curve). Here the feedforward regulation maintains the symmetry of the network. Interestingly, variation of the signal can result in the symmetry being broken. This in turn ushers in another regime of homeostasis, involving an absolute concentration robustness of $A_p$ to changes in total substrate concentration (see black curve, lower right panel).Thus feedforward regulation acting on (symmetric) multisite modification can result in the accessing of two distinct homeostatic responses.

**Expanded homeostatic capability in symmetric incoherent regulation.** We now demonstrate how the action of a symmetric incoherent regulation acting on a symmetric multisite modification network, as discussed above, can result in an expanded (and less restrictive) homeostatic response to signal. Specifically, we demonstrate that in this instance, when the total enzyme amounts are much less that the total substrate amounts, then homeostasis to signal follows. This relaxes two restrictions (i) it does not require negligible enzyme-substrate complexes, and in fact allows for sequestration of enzymes in complexes (ii) It does not require signal levels to be small. We sketch out the mathematical argument for this below.

We first note that even with sequestration of enzymes in complexes, we have $A_{pp}/A_p \propto K/P$, $A_p/A \propto K/P$. Again, the proportionality constants are independent of signal. Note that K and P denote the concentrations of active (free) kinase and active (free) phosphatase respectively. Since the total substrate concentration is much larger than the total enzyme concentration, the substrate conservation equation reads $A + A_p + A_{pp} = A_{tot}$. Thus the entire focus now is on investigating $K/P$.

The conservation equation for K reads $K + AK + A_pK + K_s = K_{tot}$. We write this equation in terms of K, noting (i) concentration of enzyme substrate complexes is proportional to the product of concentrations of the constituent enzymes and substrates. (ii) $K/K_s \propto S$. This arises from the steady state of the inactive form of the enzyme $K_s$.

Incorporating this results in an equation of the form $K(1 + \alpha A + \beta A_p + \gamma/S) = K_{tot}$. where $\alpha, \beta, \gamma$ are constants which depend on kinetic parameters.

In an exactly analogous way, the phosphatase conservation $P+A_pP+A_{pp}P+P_s = P_{tot}$ simplifies to an equation of the form $P(1+\alpha_1 A_{pp}+\beta_1 A_p+\gamma_1/S) = P_{tot}$, where $\alpha_1, \beta_1, \gamma_1$ are constants which depend on kinetic parameters.

Now symmetry of the incoherent regulation implies $\gamma = \gamma_1$, while symmetry of the multisite modification network implies $\alpha = \alpha_1$, $\beta = \beta_1$ and $K_{tot} = P_{tot}$

Now examining both expressions together indicates that if the symmetry of the system is maintained (i.e $A_{pp} = A$, then the expression for K and P are identical, in which case $K/P = 1$. In other words, $K/P$ is a constant independent of signal (and substrate amounts). Using this information in the steady state substrate equations and substrate conservation establishes that all substrate forms exhibit homeostasis independent of signal.

Note that this depends not only on the symmetry of the associated parameters but also on the symmetry of the network (i.e. $A_{pp} = A$). Interestingly we show that this condition does not always hold good–and the violation occurs due to symmetry breaking. That is however the ingredient which allows for the other homeostatic response (this time with respect to total substrate concentration )- Absolute Concentration Robustness.

## Executive summary of analytical results and resources

**Table 4.1:** Executive Summary - Chapter 4: Network regulation meets substrate modification chemistry

| Model | Network | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| | | **Feedback - Reduced models** | | |
| *DSP (no complex)* | *None* | *Absence of multistationarity* | C(2.4.1) | - |
| DSP (no complex) | $A_{pp}$ PFB (direct) | Presence of multi-stationarity | C(2.4.1a) | - |
| | $A_{pp}$ NFB (direct) | Presence of multi-stationarity | C(2.4.1b) | |
| | $A_p$ PFB (direct) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ NFB (direct) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_{pp}$ PFB (cascaded) | Presence of multi-stationarity | C(2.4.1a) | |
| | $A_{pp}$ NFB (cascaded) | Presence of multi-stationarity | C(2.4.1b) | |
| | $A_p$ PFB (cascaded) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ NFB (cascaded) | Absence of multi-stationarity | C(2.4.1c) | |
| DSP (no complex) | $A_{pp}$ PFB (direct) | Presence of multi-stationarity | C(2.4.1a) | - |
| | $A_{pp}$ NFB (direct) | Presence of multi-stationarity | C(2.4.1b) | |
| | $A_p$ PFB (direct) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ NFB (direct) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_{pp}$ PFB (cascaded) | Presence of multi-stationarity | C(2.4.1a) | |
| | $A_{pp}$ NFB (cascaded) | Presence of multi-stationarity | C(2.4.1b) | |
| | $A_p$ PFB (cascaded) | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ NFB (cascaded) | Absence of multi-stationarity | C(2.4.1c) | |

Continued on next page

Table 4.1 – continued from previous page

| Model | Network | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| *DSP (w. $AK$ complex)* | *None* | *Presence of multistationarity* | C(2.4.2a) | - |
| DSP (w. $AK$ complex) | All types | Presence of multistationarity | C(2.4.1b-e) | |
| *DSP (w. $A_pK$ complex)* | *None* | *Absence of multistationarity* | C(2.4.2a) | - |
| DSP (w. $A_pK$ complex) | $A_{pp}$ PFB (direct) | Presence of multi-stationarity | C(2.4.2b) | |
| | $A_{pp}$ NFB | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ PFB | Absence of multi-stationarity | C(2.4.1d) | |
| | $A_p$ NFB | Absence of multi-stationarity | C(2.4.1e) | |
| *DSP (w. $A_{pp}P$ complex)* | *None* | *Presence of multistationarity* | C(2.4.2a) | - |
| DSP (w. $A_{pp}P$ complex) | All types | Presence of multistationarity | C(2.4.1b-e) | |
| *DSP (w. $A_pP$ complex)* | *None* | *Absence of multistationarity* | C(2.4.2a) | - |
| DSP (w. $A_pP$ complex) | $A_{pp}$ PFB (direct) | Presence of multi-stationarity | C(2.4.2b) | |
| | $A_{pp}$ NFB | Absence of multi-stationarity | C(2.4.1c) | |
| | $A_p$ PFB | Presence of multi-stationarity | C(2.4.1d) | |
| | $A_p$ NFB | Absence of multi-stationarity | C(2.4.1e) | |
| *DSP (w. both $A_p$ complex)* | *None* | *Presence of multistationarity* | C(2.4.3a) | - |
| DSP (w. both $A_p$ complex) | $A_{pp}$ PFB (direct) | Presence of multi-stationarity | C(2.4.3b) | |
| | $A_{pp}$ PFB (direct) | Presence of multi-stationarity (in the enzyme limiting regime) | C(2.4.3b) | |
| | $A_{pp}$ NFB | Absence of multi-stationarity | C(2.4.3c) | |
| | $A_p$ PFB | Presence of multi-stationarity | C(2.4.3d) | |
| | $A_p$ PFB | Presence of multi-stationarity (in the enzyme limiting regime) | C(2.4.3d) | |
| | $A_p$ NFB | Absence of multi-stationarity | C(2.4.3e) | |

Continued on next page

Table 4.1 – continued from previous page

| Model | Network | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| | | **Feedback - Full models with direct feedback** | | |
| *DSP* | *None* | *Absence of tristability* | [74] | |
| DSP (full system) | $A_{pp}$ PFB (direct) | Presence of increased multi-stationarity (tristability) | C(2.3.1a) | - |
| | $A_{pp}$ PFB (direct) | Presence of increased multi-stationarity (tristability in the enzyme limiting regime) | C(2.3.1a) | |
| | $A_{pp}$ NFB (direct) | Presence of increased multi-stationarity (tristability) | C(2.3.2a) | |
| | $A_{pp}$ NFB (direct) | Absence of increased multi-stationarity (enzyme limiting regime) | C(2.3.2b) | |
| | | **Feedback - Full models with cascaded feedback** | | |
| DSP (full system) | $A_{pp}$ PFB (cascaded) | Presence of increased multi-stationarity (tristability) | C(2.2.1a) | fig. 4.1 |
| | $A_{pp}$ PFB (cascaded) | Presence of increased multi-stationarity (tristability in the enzyme limiting regime) | C(2.2.1b) | |
| | $A_{pp}$ NFB (cascaded) | Presence of increased multi-stationarity (tristability) | C(2.2.1b) | fig. 4.2 |
| | $A_{pp}$ NFB (cascaded) | Presence of increased multi-stationarity (tristability in the enzyme limiting regime) | C(2.2.1b) | |

Continued on next page

Table 4.1 – continued from previous page

| Model | Network | Brief description of analytical result | Ex. App. | Comp. |
|---|---|---|---|---|
| | | **Feedback - Full models with direct feedback** | | |
| | | **Effect of feedback on biphasic dose response with signal** | | |
| *DSP (full system)* | *None* | *Guaranteed biphasic response when kinetic constraints (involving catalytic constants) are satisfied* | C(2.5.1) | 4.2 |
| DSP (full system) | Any except App PFB | If the isolated system presents with biphasic dose response, the system with FB will too | C(2.5.1b) | 4.2 to 4.4 |
| | Any | Features of biphasic response such as peak $A_{pp}$ concentration, and saturation value are preserved with feedback | C(2.2.1b) | |
| | Any | The signal concentration when biphasic behavior manifests (peak) shifts | C(2.2.1c) | |
| | $A_{pp}$ PFB | Peak signal concentration shifts to lower ranges | C(2.2.1c) | 4.2 and 4.4 |
| | $A_{pp}$ NFB | Peak signal concentration shifts to higher ranges | C(2.2.1c) | |
| | | **Feedforward - Full models with kinase and phosphatase activation** | | |
| DSP | Incoherent FF | Mapped behavioral landscape of the system to that of the isolated system with no FF | C(3.2.1) | 4.5 |
| | | Derive modified kinetic constraint guaranteeing and promoting multistability | C(3.2.1a) | |
| | | Presence of isolas (full system) | C(3.2.1c) | |
| | | Presence of isolas (enzyme limiting regime) | C(3.2.1d) | |
| | Coherent FF | Mapped behavioral landscape of the system to that of the isolated system with no FF | C(3.3.1) | 4.6 |
| | | Derive modified kinetic constraint guaranteeing and promoting multistability | C(3.3.1a) | |
| | | Absence of isolas (full system) | C(3.3.1c) | |

# Chapter 5

# Multi-level complexity of cells and populations

## 5.1 Introduction

Cells communicate and respond to stimuli and information all the time. This happens both internally within the cell (intra-cellular) and externally as the cell interacts with its enviornment (extra-cellular). In the previous chapters we explored the intra-cellular signalling capacity of cells by studying how the substrate modification mechanism, in particular multisite protein modification, can enable complex and robust information processing with minimal ingredients. But just as how proteins are the powerhouses of intra-cellular signaling within a cell, cells themselves act as the basic (yet complex) information processing unit at the population level, communicating effectively with each other, other cell types, and the population, through a number of communication channels, to enable robust and complex information processing at the population level. Cells are also continuously subject to stimuli from the environment (physical, chemical, and otherwise), and they constantly react and adapt to this, while sustaining robust communication between each other.

The sum of these interactions is a collective population level behavior and response, with the cell, population of cells and the environment all contributing to it. Central to this however is the cell. This collective systems behavior is highly robust to the environment, to the stochasticity of internal expression and reactions, and other interactions. However, while such population level response is well understood, in many contexts, we are yet to have a clear understanding of how cells are able to effectively interact with one and other and enable this collective multi-level behavior. This can be attributed to the complexity of the system. For example, even within a single population of cells, at any given time, a cell is communicating with other cells (through

multiple pathways), with the environment (both reacting to passive factors, and actively interacting with substrates and agents at the environmental level), and with the population level (either directly or through the enviornment). Each of these classes of interactions can be achieved in a number of ways and can be of differing complexity. The resulting behavior at the population level is thus a systems phenomenon, a sum of the parts of the many simple one-one communications coming together to result in non-trivial behavior at the population level. Cell-cell communication is broadly characterized into four categories; autocrine, paracrine, juxtacrine, and endocrine signaling. Autocrine signaling is when the cell secretes 'ligands' that induce a cellular response through cognate receptors for the same cell (or cell type). Paracrine signaling is when a cell communicates with those in its vicinity through extra-cellular diffusion of signaling molecules. Juxtacrine signaling refers broadly to contact dependent signaling. The latter, endocrine, refers to signalling through extra-cellular diffusion of molecules, but over long distances, typically through extra-cellular fluids [174].

While there is significant interest in inferring expression and behavior at the population level, the population level behavior can be so different from that of the single cell. Recent advances in single cell studies have shown how extrapolating behavior from one level to another can be quite misleading, further underscoring the importance of unpacking the systems behavior at play. Increasing number of studies are exploring the interactions between intra & inter-cellular mechanism within cells and populations [175]. These studies have allowed for an improved understanding of key cellular processes such as homeostasis of cellular response at the population level, apoptosis, quorum sensing, and contact dependent communication. Synthetic biology has also leveraged intra and inter-cellular communication to engineer effective population level control [56]. These studies include engineered multi-population behavior with sender-receiver populations and metabolic interactions between 2 populations [176, 177, 178, 179, 180, 181, 54, 182, 183]. These efforts involve communication through a range of channels, including contact dependent communication and diffusing chemicals and metabolites, and have allowed for a range of robust non-trivial information processing responses at the population level [138, 184]. A concise summary of the progress in synthetic efforts to engineer population level response is provided by Bacchus and Fussenegger [177]. A summary of the different types of one-one interactions that exist in enabling population level behavior and/or between cells is documented below in table 5.1. This list is not exhaustive, but seeks to highlight the ubiquity and multiplicity of different communication channels between cells (of the same type and different types) and how they come together to enable population level response.

**Related studies**

| | Interaction type | References |
|---|---|---|
| Cell-Cell | Diffusible ligands | [185] |
| | Notch-Delta | [186] |
| | Micro-vesicles | [187] |
| | Diffusing cytokine communication | [188, 189, 36, 190] |
| | Tumor micro-environment | [191] |
| | Adherens and tight junctions | [192] |
| | Gap Junctions | [193] |

**Table 5.1:** Examples of communication and cellular information processing at the cellular level

In this chapter, we try to unpack the complex machinery (rightly called 'networks of networks' [194]) through which cells communicate with one and other and are able to effectively establish a collective population level behavior and response.

## 5.2 Methodology

The focus in this chapter is in trying to unpack the different interactions that exist between a single cell and populations and the environment, and how these enable a collective population level response. As seen in the table above (table 5.1), a typical population of cells (or populations of different cells) is capable of a myriad of such interactions.

With a plethora of such interactions being possible, it is helpful to define the three agents or levels at play,

- Single cell
- Population of cells
- Environmental factors

and classify interactions into the five types,

- Cell $\Longleftrightarrow$ Cell
- Cell $\Longleftrightarrow$ Population
- Cell $\Longleftrightarrow$ Environment
- Population $\Longleftrightarrow$ Population
- Population $\Longleftrightarrow$ Environment

These above categories consolidate the different ways in which one of these levels could communicate or interact with the others. A cell for example through its intra-cellular circuitry, can interact with other cells' intra-cellular circuitry (cell $\Longleftrightarrow$ cell), while also modulating the population level growth death dynamics (cell $\Longleftrightarrow$ population). These interactions themselves can be realized in different ways and in many combinations.

For example, a populations of cells can modulate a circuit response in other cells, either by secreting an biochemical species that diffuses in to the extra-cellular environment, and elicits a response, or diffuses into the other cells and elicits a response or this modulation could be contact mediated, where in it's the population strength that determines the strength of the interaction. However, this presents a complex combinatorial problem to study, as the number of interactions and the possible models that can be constructed by choosing a combination of these levels and interactions is extremely high. These models are nonetheless relevant (seen in population experiments) and thus require a systematic and dedicated study.

We employ a systems framework to tackle this complexity. The framework is grounded in two separate studies, focussing on 1) inter-cellular and population level behavior in a single populations of cells, and 2) communication between two populations of cells.

In each study, we begin by employing generic models of populations; with specific definition of the cellular, population, and environmental levels. Within our study of two populations of cells, we also explore interactive complexity at the population level by considering models that can capture multi-population behavior.

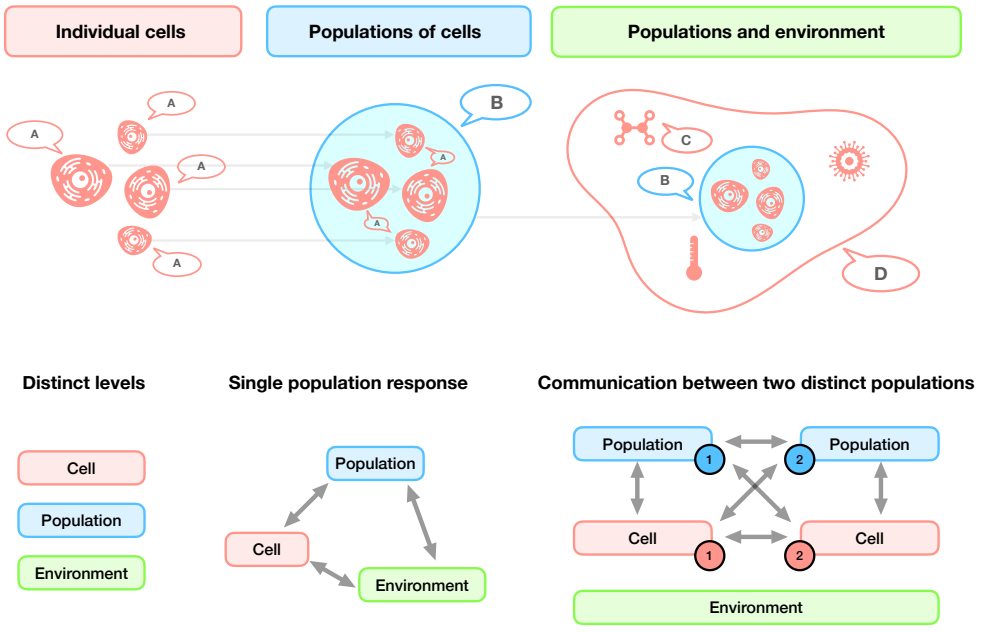



**Figure 5.1: Schematic representation of multi-level interaction and emergent collective behavior at the level of cells and populations of cells.** *Top Row:* a single cells' response (say, *A*) can be different from the response from the population (say, *B*). This can fundamentally be modulated further by factors in the environment, including physical factors such as temperature, stress, and chemical stimuli, finally resulting in a response (say, *D*), that is truly reflective of the collective behavior emerging. *Bottom row:* we focus on studying this emergent collective behavior by grounding our study on communication between and amoung three distinct levels, 1. cell, 2. population, and 3. environment; We study how these combine to enable the collective population level response in, a. single population and b. across two populations of cells.

The population level captures the growth-death dynamics of the aggregate of cells. The cellular level is designed to represent intra-cellular circuitry and its responses, and is modelled using canonical signaling motifs. The environmental factors are modelled using a single variable which represents or accounts for the sum of interactions present between the enviornment and either of the above levels.

This chapter, grounded in this higher level framework is structured as follows. We first present the single population study, and leverage the insights to focus on the two population study next. Within each study, we build a generic mathematical (master) model of the three levels and the various interactions present between them. This master model is built in such a manner that individual interactions can be turned on and off using the parameters (labelled switches within the model). This allows us to carefully explore the entire combination of cellular populations that can exist with specific interactions. A detailed description of the interactions is then listed, along with the motifs considered for the cellular level circuitry. The assumptions used in the modelling approach are then described. Within the two population study we also explore in a sub-study, multi-population level behavior arising from non-linear interactions (Lotka-Voltera) between two populations of cells, and how they affect circuit response.

Following this, we begin our analysis of the various systems by studying scenarios, each representing the presence of specific combinations of interactions between the three levels (see table 5.2 and table 5.3). We build these scenarios in clusters of increasing complexity (combinatorial and systems), finally culminating in scenarios with all interactions present. These are built to provide a clear and unambiguous insight into the role of key components and interactions in enabling the rich collective behavior of cells. This represents a bottom up approach to unravelling the complexity of a system that contains potentially hundreds of parameters and interactions. We then articulate the questions of interest that these clusters of scenarios will aim to address (and provide a suitable rationale for the construction of these clusters in this order and manner). The results are then presented.

**Figure presentation.** The title of the plots presented here follow a standard description template to denote the system studied, the variables plotted, and scenario simulated. For example, if the figure plots '$x$' and '$y$' vs basal signal '$S$', for a system containing motif '$A$' in population 1 and motif '$B$' in population 2, and the scenario simulated is '$z$', then the title of the plot is presented as follows,

$$\textit{Circuit: A - B} \mid \textit{x \& y vs S} \mid \textit{Sc. z}$$

All figures (where relevant) come with legends. Some figures have multiple y-axes that each plot different variables. The different y-axis are colour coded differently to distinguish this. Any other colour coding denotes changes to the stability of the steady state (see captions of the figures for more details). There is a broad nomenclature section (see preamble) that details the symbols used in the chapter.

### 5.2.1 Computational methodology

The system of cells and its interactions (both for the one population study and the two population study) are described within master models as a system of ODEs. These models are capable of representing the numerous combinations of interactions and population setups that we then consider in our studies, by a suitable choice of parameters associated with regulation (see equations in the relevant sections). Through such a parametric choice, the model can consider regulations that are absent, present & activating, and present & inhibitory. This system of ODEs are numerically simulated in *MATLAB*. Specific models have been sampled and these models have been cross validated with *COPASI* to ensure accuracy of the master models. Extensive bifurcation analysis is carried out using the *MATLAB* package *MatCont* [72].

**Choice of bifurcation parameter**: The 1D bifurcation is predominantly carried out by varying the basal intra-cellular circuit signal (signal to the motif $S_{basal}$). The rational for this choice is that, increasing the strength of the circuit signal, activates the cellular level response enabling a strengthening (within the relevant models) of the further interactions between the three levels. However in limited cases, when trying to explore the effect of specific parameters, including relative strength of key interactions, population growth rate, diffusion strength; we perform bifurcations along the other relevant parameters.

**Parameters**: The parameters for the models are chosen reasonably within the bounds of experimentally observed kinetic values. We also take into the account that the model incorporates levels (population dynamics and cellular dynamics for example) that often have kinetics over different timescales. Reflective of this, for much of our analysis, we ensure that the parameters chosen for the population level (both growth and death dynamics) and the environmental factors are two order of magnitude lower than those chosen for the cellular level. However, the kinetics of the two levels are not always distinct in this manner, and in some studies, especially when each level is capable of emergent behavior, we relax this constraint and consider the effect and role of dynamics of the population level and the cellular level occurring on similar timescales. We assume in the relevant models with extra-cellular diffusion of species, that the volume of the cell is considered to be a significantly smaller than the volume of the environment (2000 times smaller - see appendix A for more details).

**Computational approach:** The number of scenarios to consider are, as explained earlier, extremely high (both in the single population study and in the two population study). But a systematic analysis of them is none the less necessary to provide insight into the complicated origins of population level response.

In order to address this in an efficient manner, we built an automated pipeline in *MATLAB* that generates the required 1D bifurcation plots for all possible manifestations of the master models (including combinations of circuit level motifs, population level and environment level interference and interactions), that only requires definition of the

scenarios by mention of the switch values from table 5.2 and table 5.3. This automation pipeline performs the bifurcation analysis by calling the subroutines within *MatCont*, and plots the required bifurcation plots with clear labelling of multiple critical points such as limit points (saddle node bifurcation), Hopf bifurcation point, Branch point bifurcation etc. In instances were the system presents oscillation, the pipeline carries out a limit cycle continuation from an observed Hopf bifurcation point, thereby allowing us to better explore the dynamics of the emergent oscillations. These plots are generated for all variables of the system (across the three levels).




**Figure 5.2: Schematic representation of the computational pipeline used to explore the multiplicity of scenarios of interactions and cellular circuit combinations in a robust scalable manner.** From a modelling perspective, in order to confront the combinatorial complexity of the number of interactions possible (and present) between cells within a population, and the different cellular circuits that they can contain, we implemented an automation pipeline in *MATLAB* to perform bifurcation analysis for different scenarios of cellular systems (see tables 5.2 and 5.3) and circuit motif choices we considered. The automation pipeline was built using the sub-routines of the *MatCont* package. It follows a eight part process where in, 1. We initialize the master model (described in eqs. (5.1) and (5.2)), 2. Impose the scenario conditions (i.e., the interaction choices), 3. Select/sample the choice of parameters for the kinetics and interactions, 4. Initiate the system with these conditions in *MATLAB* and *MatCont*, 5. Run a *MATLAB* simulation, 6. Run a *MatCont* bifurcation along chosen parameter (by default $S_{basal}$), 7. Save figures of all relevant species concentrations and behaviors, 8. Synthesis and conclusions based on cluster of scenarios and plots. This automation pipeline generates a landscape of results with all possible combinations of scenarios and circuit motifs of interest (for both the single population study and the two population study).

A schematic representation of the 8 step computational pipeline is depicted in fig. 5.2. This pipeline thus allows us to explore the number of combinations of interactions present that might otherwise be impossible to reasonably tackle through manual bifurcation analysis.

## 5.3 Single population study

### 5.3.1 Mathematical model

Our analysis of the single population system composed of a single cell type with the same internal circuitry (motif), begins with the following mathematical model description. As mentioned earlier, the populations are modelled as a system of ODEs. The higher level description of the model is provided below followed by the assumptions made, a description of the various intra-cellular circuits studied, and an elaborate description of the interactions modelled (and their biophysical meaning).

The model (and the system) can itself be interpreted as the interplay between five classes of variables

1. Population level $n$
2. Intra-cellular circuit $c$
3. Intra-cellular circuit response (diffusible across intra & extra-cellular spaces) $e_1$
4. Extra-cellular circuit response (diffusible across intra & extra-cellular spaces) $e_2$
5. Environmental factor $\alpha$

These are modelled using the following five ODEs, where except for $\frac{dc}{dt}$ which represents the class of variables making up the intra-cellular circuit, each other ODE directly describes the dynamics of the relevant systems. The circuit level ($\frac{dc}{dt}$) is defined as a function here ($f(S)$ of a signal $S$), as depending on the motif considered there are additional circuit level variables to be defined. For example, if a covalent modification system is considered as the intra-cellular circuit, two circuit level variables ($\frac{dc}{dt}$ and $\frac{dR_p}{dt}$) are defined. In this master model these additional ODEs are considered subsumed into $f(S)$. For a detailed description of the different motifs employed and their description, see appendix B.2. This ODE level description which represents the relevant circuit motif being regulated by a upstream signal ($S$), can be perturbed/regulated as shown in the model above by the population level and the environmental level (see interactions 5 and 6).

$$
\begin{aligned}
\frac{dn}{dt} &= k_{gn}\left[1+\text{Sw1}\frac{k_\alpha \alpha}{K_\alpha+\alpha}\right]\left[1+\text{Sw3}\left(\frac{k_{e1_3}e_1}{K_{e1_3}+e_1}+\frac{k_{e2_3}e_2}{K_{e2_3}+e_2}\right)\right]\ldots \\
&\quad -k_{dn}n\left[1+\text{Sw2}\frac{k_{\alpha_2}}{K_{\alpha_2}+\alpha}\right]\left[1+\text{Sw4}\left(\frac{k_{e1_4}e_1}{K_{e1_4}+e_1}+\frac{k_{e2_4}e_2}{K_{e2_4}+e_2}\right)\right] \\
\frac{dc}{dt} &= f(S) \quad \text{where} \quad S = S_{basal}\left[1+\text{Sw5}\frac{k_{n_5}n^{h_5}}{K_{n_5}+n^{h_5}}+\text{Sw6}\frac{k_{\alpha_6}\alpha}{K_{\alpha_6}+\alpha}\right] \\
\frac{de_1}{dt} &= k_{a_{e1}}c-k_{de_1}e_1+D\left[e_2-e_1\right] \\
\frac{de_2}{dt} &= k_{e2_{pr}}+D\left(\frac{nV_1}{V_2}\right)\left[e_1-e_2\right]-k_{d_{e_2}}e_2 \\
\frac{d\alpha}{dt} &= k_{\alpha-pr}-kd_\alpha\alpha\left[1+\text{Sw7}\frac{k_{n_7}n^{h_7}}{K_{n_7}+n^{h_7}}+\text{Sw8}\frac{k_{e2_8}e_2}{K_{e2_8}+e_2}\right]
\end{aligned}
\tag{5.1}
$$

*The population level ($n$)* has a basal growth rate ($k_{gn}$) and a basal death rate ($k_{dn}$) that gets affected/modulated by two different levels; either this basal growth rate gets strengthened by the environment ($\alpha$) or the intra-cellular circuit (both from within the cell ($e_1$), and from the extra-cellular environment ($e_2$). Similarly the basal death rate of the population can get strengthened through the environment and the cellular level. However these are assumed to be from different processes of the cell, and hence their effect is captured within our model in a multiplicative manner rather than additive. This allows us to explore better the scenarios where the environment is strengthening the population, but the circuit undermines it and masks the environmental effect (and vice versa).

*The intra-cellular circuit* is considered to be modular as explained above. The basal signal of the circuit can be strengthened however by the population level, and the environmental level. An additive structure is assumed here for these effects.

The circuit motif response $c$ (or $cR$) however in turn produces an *intra-cellular circuit response* ($e_1$), which has its own basal linear degradation within the cell $k_{d_{e1}}$. This intra-cellular circuit response $e_1$ can diffuse out of the cell as extra-cellular circuit response $e_2$. We assume that there is a basal production of this response in our model in the extra-cellular environment (to account for such production from the intra-cellular response of another cell type for example). *The extra-cellular response* $e_2$ is assumed to also degrade (linearly) in the extra-cellular environment.

**The environmental level** $\alpha$ is an aggregate of all the environmental factors that can be modulated or otherwise modulate other levels. The environmental factors are assumed to have a basal production term ($k_{\alpha-pr}$), and basal degradation term ($kd_\alpha$). This degradation can be affected by the population level $n$ and by the extra-cellular circuit motif response $e_2$.

### Interactions

The model described above as mentioned briefly, has 8 interactions between the different levels. These interactions are all contained within the master model and can be

switched on and off with a suitable choice of parameters for the model variables (Sw1) to (Sw8). Below we list them again and expand on the modelling choice for the interaction term.

Each interaction is modelled with a Michaelis–Menten description. This choice is made to account for a saturating effect within the interaction and to ensure there is no run away response from the system. Such a description also helps to keep mathematical complexity to a minimum without the need for any complicated assumptions regarding specific interactions. The interaction for the population response however is modelled as a hill function (with hill coefficient $h_5$ & $h_7$). This is assumed as as to better represent the cooperativity and increased switch-like response that population level responses can potentially deploy. Such co-operativity can underpin qualitatively new types of behavior and hence is worth consideration.

**Sw1** Environmental factor promotes population growth rate directly

**Sw2** Environmental factor promotes population decay rate directly

**Sw3** Intra-cellular circuit response promotes growth rate directly (through two channels, 1. from within the cell $e_1$, and 2. from the extra-cellular enviornment)

**Sw4** Intra-cellular circuit response promotes death rate directly (through two channels, 1. from within the cell $e_1$, and 2. from the extra-cellular enviornment)

**Sw5** Population strength strengthens intra-cellular circuit activation

**Sw6** Environmental factor strengthens intra-cellular circuit activation

**Sw7** Population strength promotes environmental factor degradation

**Sw8** Intra-cellular circuit response promotes environmental factor degradation (through the extra-cellular environment)

We analyze this model (with various intra-cellular circuit motifs) under a multitude of scenarios where the individual interactions (and their combinations) are all either present or absent. Each of these scenarios represent a realistic system with these different interactions coming together in enabling a population level response. These various deployed at the intra-cellular level are listed below.

#### Circuits and motifs

We consider five intra-cellular circuits in our analysis (as representative examples of intra-cellular information processing circuits). These models and their characteristic behavioral features are listed below.

| | |
|---|---|
| 1. Covalent Modification Cycle | Mono-stability |
| 2. Positive feedback | Multi-stability |
| 3. Negative feedback | Homeostasis |
| 4. Negative feedback | Oscillation |
| 5. Substrate depletion oscillator | Relaxation oscillation |

These motifs are adapted from Tyson et al., [52], and the mathematical ODE level description of these circuits is provided in appendix B.2. The circuit level response from

these intra-cellular motifs is given in fig. 5.3.

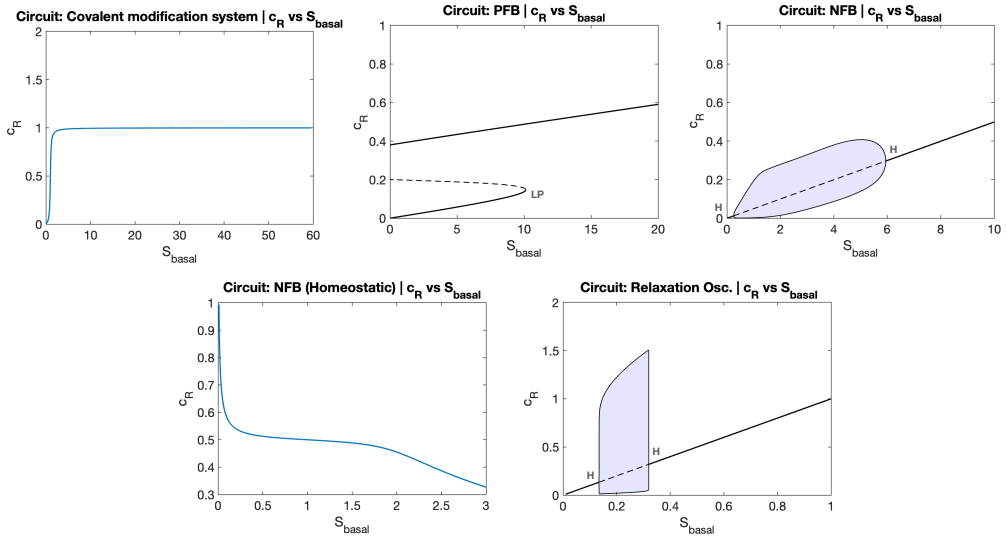



**Figure 5.3: Expected circuit response from the intra-cellular circuit motifs.** The intra-cellular circuits are modular in our mathematical model description, and the signal response behavior is as shown above. The motifs and the parameter values for their operation are adapted from Tyson et al., [52]. The simple covalent modification system, PFB, NFB, NFB (homeostatic) and Relaxation oscillator are capable of switch-like response, multistability, oscillations, homeostasis, and relaxation oscillations in the resulting response of the motif ($c_R$).
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

### Assumptions

The following assumptions are made in the single population model described above.

1. The **population** grows and decays at a basal level which can both be affected by circuit response ($e_1$ or $e_2$) and the environmental factors ($\alpha$).

2. The processes and factors through which the environment ($\alpha$) is able to affect the **population growth-death dynamics** ($n$) are considered different from that due to the cellular levels' effect ($e_1$ & $e_2$). To account for this we have a multiplicative structure for the effect of interactions from the circuit level and the environmental level onto the population[1].

3. The **circuit** motif is always modular. The circuit response is $c$ and leads to the production of an intra-cellular species ($e_1$). This is decoupled to allow for diffusion and sequestration in extra and intra-cellular without affecting the modularity of the intra-cellular circuit.

[1] Another plausible modelling approach considering these different interactions would be to allow for an additive structure between their effects. In our study, the role of each interaction is independently studied and thus insights that are made are ensured to be broadly relevant to such a setup as well.

4. The **intra-cellular species** ($e_1$) can diffuse out of the cell and the dilution is accounted for by taking the volumes (the cell and the extra-cellular environment) in to account. The intra-cellular volume is assumed to be negligible in comparison to extra-cellular volume (which is considered to be fixed). See parameters (appendix A) for more details.
5. The **extra-cellular species** ($e_2$) has a linear rate degradation to account for any environmental degradation.
6. The **environmental factor** ($\alpha$) has a linear rate degradation to account for any environmental degradation.

### 5.3.2 Scenarios

We study the role of various interactions and communications in a single population of cells and how it affects and enables a collective behavior at the population level by studying different scenarios with select combinations of interactions present. These scenarios are constructed by toggling the switches on and off in the master model described above in eq. (5.1).

The scenarios are built into clusters, starting with the systems with the least interactions and complexity, and building on them; at each stage carrying forward the insights from the previous cluster of scenarios to better inform further simulations. These clusters and scenarios are listed in the table below.

Unless explicitly mentioned below the interaction is absent in the model (i.e, every switch (Sw) is equal to 0)

**Table 5.2:** Scenarios used to analyze role of interactions in enabling population level response

| Sc. | Active levels | | | Interactions | Descriptions | Reference |
|---|---|---|---|---|---|---|
| | Cir. | Pop. | Env. | | | |
| 1 | × | × | × | None | Establishes reference | 5.3 and 5.4 |
| | **No environmental interference** | | | | | |
| 2 | ✓ | ✓ | × | Sw5 | Effect of pairwise cell-pop. interaction in isolation (with feedback) | 5.5 |
| 3 | ✓ | ✓ | × | Sw3 | | |
| 4 | ✓ | ✓ | × | Sw4 | | |
| 5 | ✓ | ✓ | × | Sw3, Sw5 | | |
| 6 | ✓ | ✓ | × | Sw4, Sw5 | | |
| | **No circuit interaction** | | | | | |
| 7 | × | ✓ | ✓ | Sw1 | Effect of interactions between population and environment (with feedback | 5.6 to 5.8 |
| 8 | × | ✓ | ✓ | Sw2 | | |
| 9 | × | ✓ | ✓ | Sw1, Sw7 | | |

Continued on next page

Table 5.2 – continued from previous page

| Sc. | Active levels Circ. | Pop. | Env. | Interactions | Descriptions | Reference |
|---|---|---|---|---|---|---|
| 10 | × | ✓ | ✓ | Sw2, Sw7 | | |
| | **No population interactions** | | | | | |
| 11 | ✓ | × | ✓ | Sw8 | Effect of interactions between circuit and enviornment (with feedback) | 5.9 |
| 12 | ✓ | × | ✓ | Sw6 | | |
| 13 | ✓ | × | ✓ | Sw8, Sw6 | | |
| | **All levels active** | | | | | |
| *Assumption: Population and circuit always have bidirectional interaction* [ $c \leftrightarrow n$ ] | | | | | | |
| | | | | $c \leftrightarrow n \rightarrow a \rightarrow c$ | | |
| 14 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw7, Sw6 | Study the role of secondary interaction between pop. and circ. through the environment | 5.10 |
| 15 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw7, Sw6 | | |
| | | | | $c \leftrightarrow n \leftarrow a \leftarrow c$ | | |
| 16 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw8, Sw1 | Study the role of secondary interaction between circ. and pop. through the environment | 5.10 |
| 17 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw8, Sw2 | | |
| 18 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw8, Sw1 | | |
| 19 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw8, Sw2 | | |
| | | | | $c \leftrightarrow n \leftrightarrow a \rightarrow c$ | | |
| 20 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw7, Sw1, Sw6 | Study the role of direct but secondary interaction between env. and circ. | 5.11 |
| 21 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw7, Sw2, Sw6 | | |
| 22 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw7, Sw1, Sw6 | | |
| 23 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw7, Sw2, Sw6 | | |
| | | | | $c \leftrightarrow n \leftrightarrow a \leftarrow c$ | | |
| 24 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw7, Sw1, Sw8 | Study the role of direct but secondary interaction between circ. and env. | 5.11 |
| 25 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw7, Sw2, Sw8 | | |
| 26 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw7, Sw1, Sw8 | | |
| 27 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw7, Sw2, Sw8 | | |
| | | | | $c \leftrightarrow n \leftarrow a \leftrightarrow c$ | | |
| 28 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw6, Sw8, Sw1 | Study the role of direct but secondary interaction between env. and pop. | 5.11 |
| 29 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw6, Sw8, Sw2 | | |
| 30 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw6, Sw8, Sw1 | | |
| 31 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw6, Sw8, Sw2 | | |
| | | | | $c \leftrightarrow n \rightarrow a \leftrightarrow c$ | | |
| 32 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw6, Sw8, Sw7 | Study the role of direct but secondary interaction between pop. and env. | 5.11 |
| 33 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw6, Sw8, Sw7 | | |
| | | | | $c \leftrightarrow n \leftrightarrow a \leftrightarrow c$ | | |

Continued on next page

Table 5.2 – continued from previous page

| Sc. | Active levels | | | Interactions | Descriptions | Reference |
|---|---|---|---|---|---|---|
| | Circ. | Pop. | Env. | | | |
| 34 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw6, Sw8, Sw7, Sw1 | Study the role of multiple interactions existing between each level | 5.11 |
| 35 | ✓ | ✓ | ✓ | Sw3, Sw5, Sw6, Sw8, Sw7, Sw2 | | |
| 36 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw6, Sw8, Sw7, Sw1 | | |
| 37 | ✓ | ✓ | ✓ | Sw4, Sw5, Sw6, Sw8, Sw7, Sw2 | | |
| | **Exemplar scenarios** | | | | | |
| | | | | $c \leftarrow n \leftarrow a \leftarrow c$ | Study the unidirectional communication between all three levels | |
| 38 | ✓ | ✓ | ✓ | Sw5, Sw8, Sw1 | | 5.12 |
| 39 | ✓ | ✓ | ✓ | Sw5, Sw8, Sw2 | | |
| | | | | $c \rightarrow n \rightarrow a \rightarrow c$ | | |
| 40 | ✓ | ✓ | ✓ | Sw3, Sw7, Sw6 | | 5.12 |
| 41 | ✓ | ✓ | ✓ | Sw4, Sw7, Sw6 | | |

**Reasoning**

The clusters of scenarios considered above are built to help untangle the different complexities that are present with having to consider the multiplicities of interactions and intra-cellular circuit motifs.

This approach begins with trying to unpack pair-wise communication between the different levels before focussing on the more complicated setup with each level communicating with the other. Here, we give precedence to the interactions between cells and populations by making these interactions present for all scenarios, as this is the primary focus of the study. The environmental level, apart from being a major factor and modulator of the expected response (both at the population and circuit level) is used to study the effect of secondary communication channels between cells and the population level.

All in all, this set up allows us to answer key questions that lie at the heart of cellular information processing and population level response without getting muddled in the complexity (of both a combinatorial and parametric nature arising from the model).

### 5.3.3 Questions

The key questions we aim to answer are listed below. While these don't represent the full scope of the insights obtained from the study (see section 5.3.4 for a detailed synthesis), these offer a purview of the initial focus behind the study.

- How does extra-cellular response differ from circuit level response (for different intra-cellular circuit motif)?
- How does the population level affect circuit level response and extra-cellular response (for different intra-cellular circuit motif)?
- How does the enviornment modulate or affect the population level and cellular response?
- How does the environment get affected by the population level and the extra-cellular circuit response?
- How do different cell-population interactions (contact, extra-cellular, metabolite, intra-cellular) compare in effecting robust extra-cellular response?
- How does dilution and diffusion affect the extra-cellular circuit response?

### 5.3.4 Results and Discussion

The simulations were run in the order of the scenarios listed in the table above (table 5.2), using the automation pipeline described in the computational methodology (see section 5.2.1). The results below are organized in the same spirit and style as the scenarios.

#### Results

It is worth noting at this stage how the different levels are related in the absence of any interaction (scenario 1). The population level ($n$) has a basal growth-death dynamics, that is unaffected by any of the other two levels. The intra-cellular circuit similarly has a basal signal (along which the bifurcations are carried out), which results in a cellular response ($c$, identical to the expected circuit motif response). This response then produces a intra-cellular species ($e_1$) that diffuses out of the cell ($e_2$). The environmental factor ($\alpha$) has a basal growth death dynamics again, which is unaffected by any level.

Thus with this representation, the population level and the environmental factors are decoupled, while the circuit response and its associated intra & extra-cellular species concentrations, closely follow the circuit motif response. This is seen in the figure above (see fig. 5.4), and represents the benchmark against which we compare impact of any interactions between the levels. The intra & extra-cellular species concentrations ($e_1$ and $e_2$) however do depend on the population level strength implicitly due to dilution effects arising from the diffusion process. This however is only captured when the population level itself changes (as we will see in the subsequent scenarios).

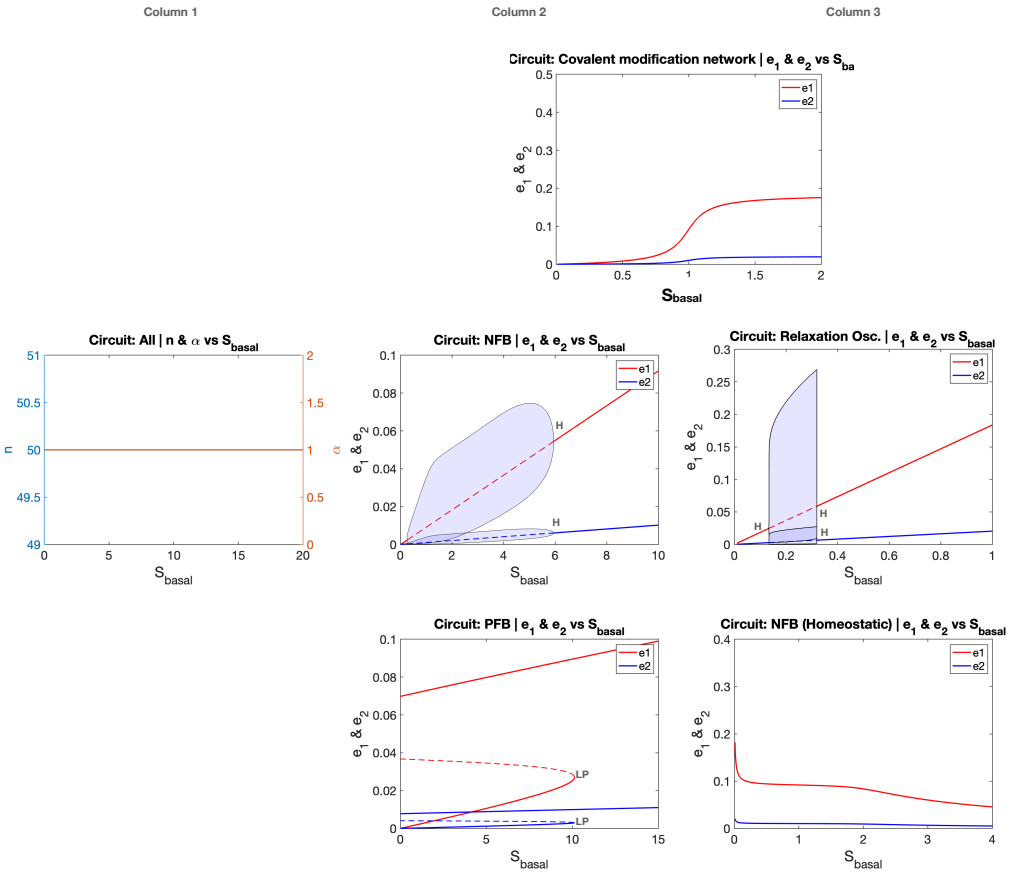



**Figure 5.4: Single population study: Population strength, environmental factors and intra & extra-cellular circuit response in the absence of any interactions between the cell, population and the environment.** *Column 1* shows the population strength and environmental factor as $S_{basal}$ is changed. *Column 2-3* shows the intra & extra-cellular circuit response for the different motifs in the system as $S_{basal}$ is changed. Clockwise from top; covalent modification system, relaxation oscillator, NFB (homeostatic), PFB, NFB (oscillatory). This behavior seen with no interactions serves as a benchmark against which the effect of interactions can be assessed. We see the extra-cellular response and environmental level steady as $S_{basal}$ is changed (as these levels are uncoupled now), and the intra & extra-cellular response ($e_1$ & $e_2$) from the circuit trailing the circuit response closely. This is intuitive as the circuit response directly leads to the production of $e_1$ within the cell which then diffuses out of the cell as $e_2$. Thus intra-cellular response can directly show at the extra-cellular level in the absence of any other interactions or modulating factors. However the extra-cellular response is necessarily subdued (as is the case with oscillations - as we can see for the NFB plot), due to the dilution effect of the environment.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

### *Effect of cell-population interactions on population level response (without any environmental interference)*

Simple interactions between the population level variable and the cellular circuit level can allow information processing to transcend levels and allow for otherwise complicated information processing across multiple levels. This is seen in our simulations of

scenarios 2 to 6 (see fig. 5.5). For example, covalent modification system, NFB (Homeostatic motif), NFB (Oscillations motif), PFB and relaxation oscillator motifs at the intracellular level can afford switch like response, homeostasis, oscillations and even multistability at the population level (seen in the population strength $n$). Figure 5.5 shows two examples of such behavior manifesting from unidirectional interaction between the population level and the circuit level (circ. $\rightarrow$ pop.) resulting in switch-like behavior (*column 1, row 1*) and homeostatic behavior (*column 1, row 2*) in the population strength $n$.

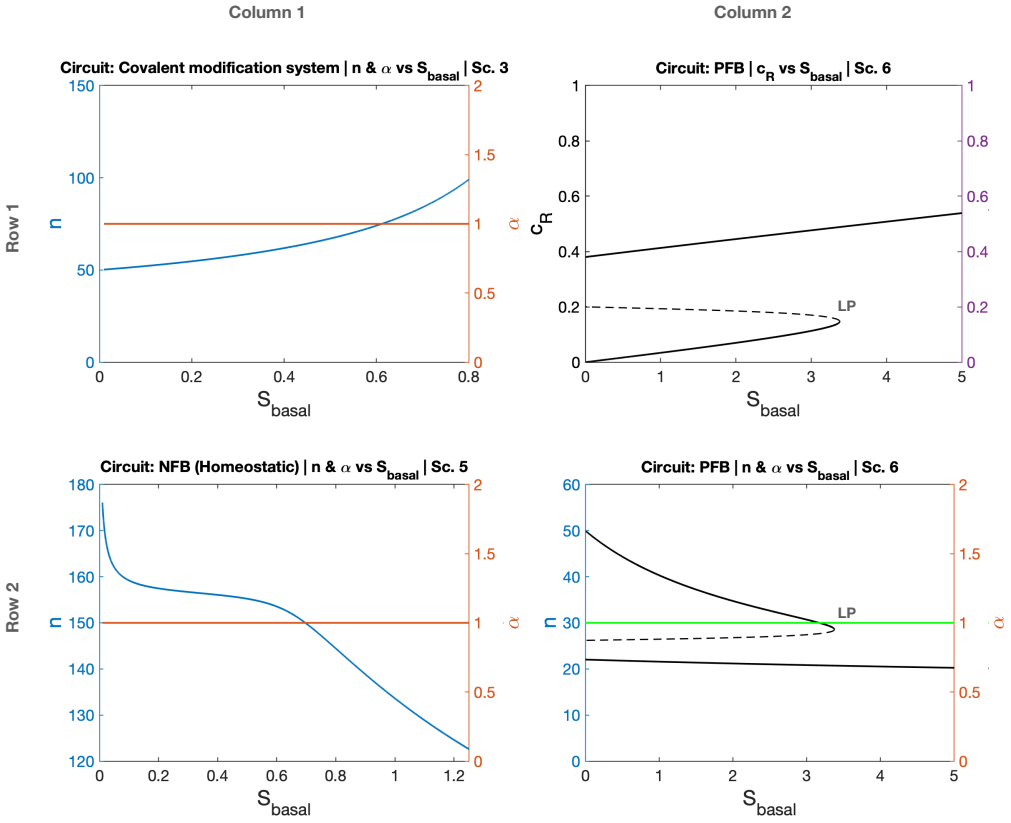



**Figure 5.5: Simple interactions between population level and cellular level allows for information processing behavior to transcend levels.** *Column 1* shows unidirectional interactions between the circuit and the population level (scenarios 3 and 5) lending qualitative circuit behavior to the population level variable (population strength $n$). *Column 2* shows bidirectional interaction between the population level and the cellular level leading to bistability at the population level ($n$) and at the circuit level ($cR$). These results show how simple unidirectional interactions between the population and the cellular level can allow for borrowed behavior to manifest at the other level and across the system. With bidirectional communication the effective systems behavior seen is a modulation of the circuit motif response manifesting at the systems level. This is seen in *column 2* where the range over which the system exhibits multistability is significantly reduced in basal signal regulation vis-a-vis when there is no such interaction (see fig. 5.3). Exemplars are shown from scenarios 3,5, and 6, and results representative of scenarios 2-6. Note the multiple y-axes in the plots with population strength ($n$) and environmental factors ($\alpha$).
[LP - Limit Point (saddle node bifurcation); Dotted lines represent unstable steady states]

Bidirectional communication between these two levels can further perturb responses from each level. This also includes a system level response with diminished range (along $S_{basal}$) over which multi-stability is seen in the population level variables (and the overall system). This is shown in *column 2* fig. 5.5. With bidirectional communication (scenario 6), we can that the range over which bistability was observed has shrunk for the circuit motif, and the overall system; Originally without interactions multistability was observed for $S_{basal}$ ranges between $\sim 0$ to $\sim 10$ - see fig. 3.3, while with bidirectional interaction between population level variables and the cellular circuit, this range has shrunk to between $\sim 0$ to $\sim 6$ - as shown in *column 2* of fig. 5.5. These results underscore the importance of cross-level communication (both unidirectional and bidirectional) that allows for complicated circuitry at one level to both enable collective behavior and sharing such response structures to other levels with simpler dynamics. Together these interactions could form the basis by which levels can utilise, borrow and leverage features and behaviors from one and another.

*Effect of population and enviornment interacting with each other*

Interactions between the environmental factors and the population level variable can modulate the response arising from each other. fig. 5.6 shows different examples of how the population level can effectively modulate the environmental level. While on one hand, these relationships are intuitive, it is worthy to bear in mind how the environmental factors are seldom modular themselves. In experiments they are easily accessible proxies for responses at the extra-cellular environment and can themselves modulate/affect/perturb other systems and cells downstream. Thus the impact that population strengths can have on these factors is significant. Figure 5.6 shows how increasing population growth, affect the environmental factors in a number of ways, and in particular (bottom row) shows how when there is bidirectional modulation, the population level also gets subtly perturbed leading to a dampened and steady response. In particular, scenario 10 *(bottom right plot in fig. 5.6)* shows how when the environmental factor and the population strength mutually inhibit each other, the behavior of each is modulated by the other, and there is a highly suppressed population strength (in comparison to other scenarios with no such inhibition). We have chosen to present this with a bifurcation along growth rate of the population ($k_{gn}$), but such an interaction can be leveraged in many ways, and can lead to complicated modulation of the environmental factors and the system, when the population level dynamics are more complicated. For example, when the population level dynamics is capable of emergent behavior (as we consider later in this study - see section 5.5), such behavior can be seen in the environmental factor and such behavior can be further shaped by bidirectional interactions with the latter.

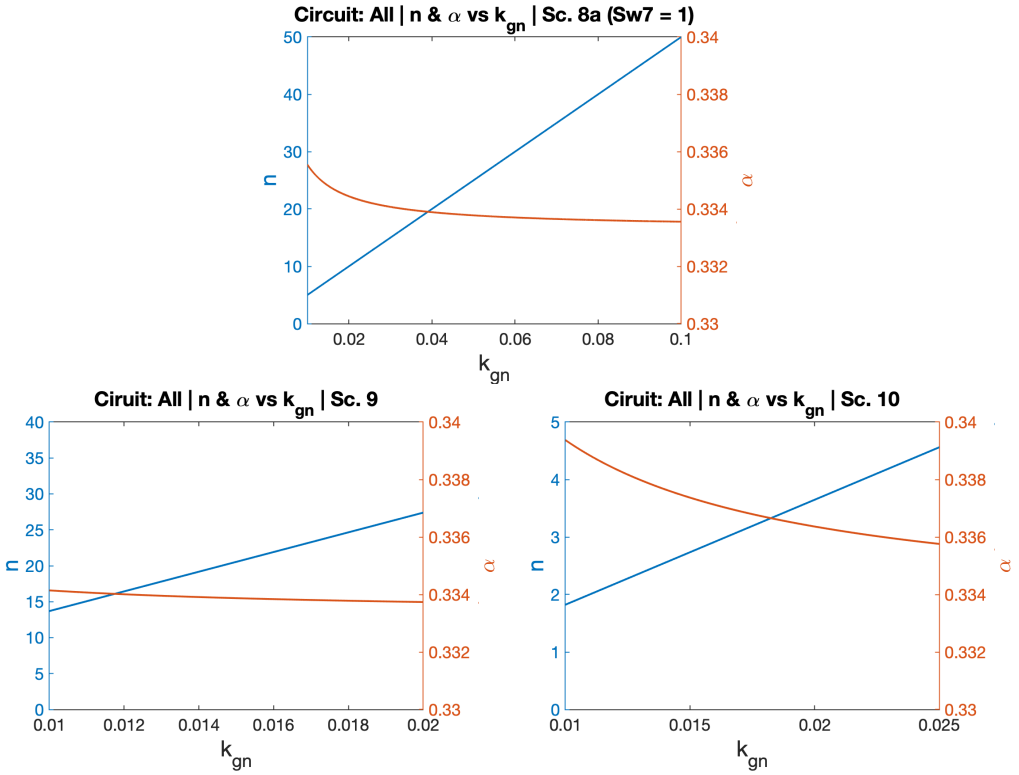



**Figure 5.6: Interactions between the population and environmental factors can allow for the levels to modulate each other.** Figure shows population strength and environmental factors varying as growth rate of the population is changed $k_{gn}$. *Clockwise from top:* scenario 8a - population strength degrades environmental factors (unidirectional), scenario 10 - population level and environmental level bidirectional-ly interact with each other (mutual inhibition), scenario 9 - population level and environmental level bidirectional-ly interact with each other (population strength degrades environmental factors, while environmental factors promote population strength. The above results (*top row*) show how changing growth rate of a population can lead to changes in the environmental factor (scenario 8a, (Sw7) = 1). When bidirectional communication between these are established we can see a modulated response in both layers (*bottom row*). Exemplars shown from scenario 8a, 9 and 10, results representative of scenarios 7 to 10.

While the population level directly doesn't interact with the cellular response in these scenarios, the effect of interaction between the population and the environment, can percolate and can be observed in the extra-cellular circuit motif response due to the continuous dynamic shuttling of the species outside and inside the cells (which depends on the population strength). This behavior is shown in the fig. 5.7. Figure 5.7 shows how introduction of interactions between the environmental factors and the population strength can lead to a diluted concentration of intra & extra-cellular circuit motif response. The figure focuses on the extra-cellular circuit response $e_2$ to illustrate this point. The comparison is made between scenario 1 (with no interactions between the levels) and scenario 10 (with the population and the environmental level interacting bidirectionally). This implies that even when there is no direct interaction between the circuit motif or cellular circuits and the other levels, or factors; interactions between

population level and the environmental level can fundamentally alter the observed strength or concentration of the circuit motif response (both intra & extra-cellular concentration). This insight bears great relevance to inference studies that rely on extra-cellular proxies to determine circuit motif response and behavior. The same circuit motif response can register at different concentrations internally and externally due to interactions between the population and the environment.

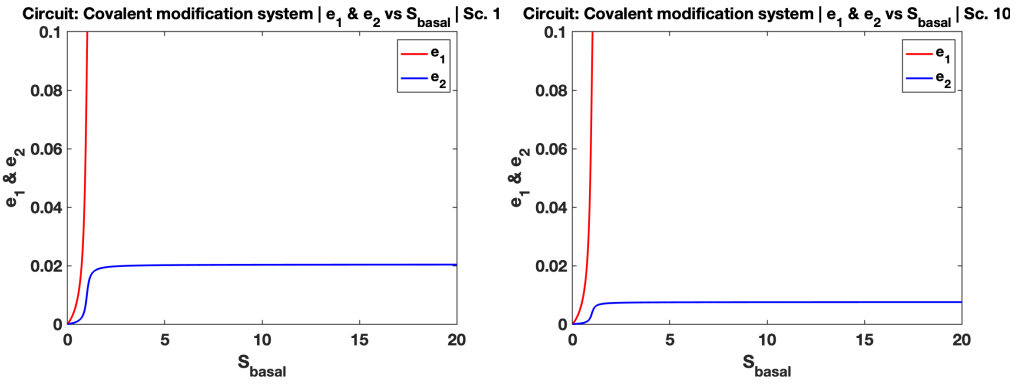



**Figure 5.7: Interactions between population and enviornment factors can also modulate extra-cellular circuit response indirectly (fixed population growth rate $kg_n$ - changing circuit strength $S_{basal}$).** *Left:* Intra & extra cellular circuit response as basal signal of the circuit changes (no interactions in the system). *Right:* Intra & extra cellular circuit response as basal signal of the circuit changes (population strength and the environmental factors mutually inhibit each other). In scenario 10 (*right*), the population strength and environmental factors are can perturb each other as shown in fig. 5.6. However, while the circuit level is uncoupled in this scenario, the extra-cellular circuit response can get modulated by these interactions. As the plots above illustrate, the variation in the population strength ($n$), can lead to dilution effects that manifest in the concentration of the circuit response (due to the extra-cellular response shuttling in and out of cells). The plots above illustrate this dilution effect by contrasting the extra-cellular concentration of the circuit response ($e_2$) between scenario 1 (with no interactions), and scenario 10 (where population strength and environment interact bidirectionally); See fig. 5.8 for a similar effect.

Figure 5.8 similarly shows another example of this behavior. The figure shows how the intra & extra-cellular circuit responses vary as the population growth rate is changed, when the interactions between the environmental factor and the population are, unidirectional (*left*) and when they are bidirectional (*right*). In this instance we can see how increasing population strength similarly can affect the cellular response observed. This effect is again either weak or stronger depending on whether the interaction between the environmental factors and the population strength is unidirectional or bidirectional. Beyond increasing population strength strengthening extra-cellular circuit response (seen in both scenarios), we can see how unidirectional strengthening of the population level by the environment can enable a stronger extra-cellular response (left plot) vis-a-vis a bidirectional communication where the population degrades the environmental level leading to a weakened response (right plot).

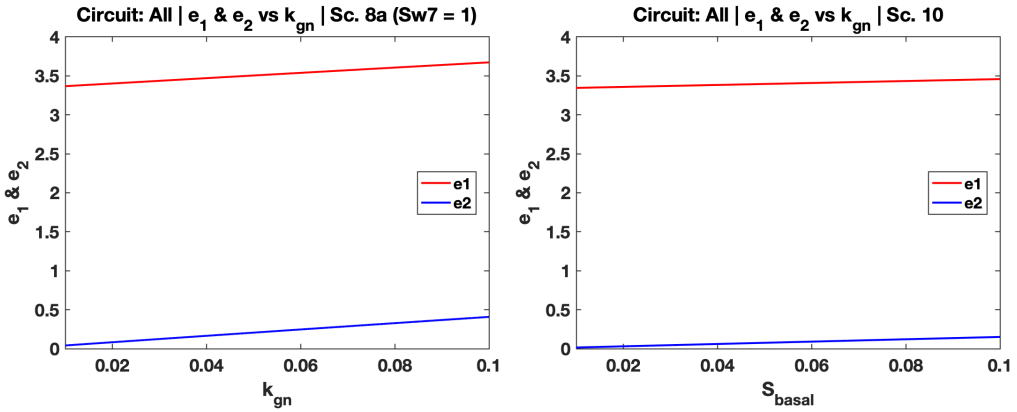



**Figure 5.8: Interactions between population and enviornment level can also modulate extra-cellular circuit response indirectly (fixed circuit strength $S_{basal}$ - changing population growth rate $kg_n$).** The population level and environmental level are can perturb each other as shown in fig. 5.6. In contrast to fig. 5.7 these plots illustrate how as $k_{gn}$ changes (i.e., the population strength changes), the extra-cellular response is affected (even though the circuit response isn't, $c_R$ is constant). This can manifest at both the intra-cellular response $e_1$ and the extra-cellular circuit response $e_2$, implying that even though the population and the environmental level do not affect the circuit in any way, they can modulate it depending on their interaction strength and pattern in unique ways. *Left plot* shows an example of unidirectional effect with the environmental level affecting the population growth, and *right plot* represents a bidirectional interaction with the population level also degrading the environment. We can already observe the strengthening of the extra-cellular response due to increasing population strength. The plots shown are representative of all internal circuit motif choice (since circuit level response is fixed as $k_{gn}$ is changed.

All in all these results show how beyond modulation of the environmental factors by the population strength and vice versa, these simple interactions can have significant consequences for the response seen from the circuit motif both at the intra-cellular level and at the extra-cellular level.

*Effect of intra-cellular circuit interacting with environmental factors*

Similar to the scenarios with interactions between population strength and the cellular circuit, information processing capabilities can transcend the circuit to the environment. Feedback from the environmental factors back to the circuit can further refine this interaction leading to modulated regimes over which qualitative behaviors are observed within the system. The results are similar to when the intra-cellular circuit affects the population strength, however there is no explicit feedback in the unidirectional interaction unlike earlier (between circuit and population strength, arising due to dilution effect discussed earlier), due to the fact that there is no coupling between the environment and the extra-cellular circuit response. Thus the behavior is more direct and intuitive. However such results again have gross implications as they represent a scenario where the environmental factor is being actively modulated by intra-cellular circuitry. The environmental factors can in turn affect other populations, or species or otherwise have vast implications on factors downstream. We also later on explore scenarios where the environment actively engages with the population level and the

circuit level together.

Figure 5.9 shows examples of direct modulation of the environmental factors by the cellular circuit lending it non-linear information processing capabilities. This is shown for both of PFB and NFB circuit motifs.

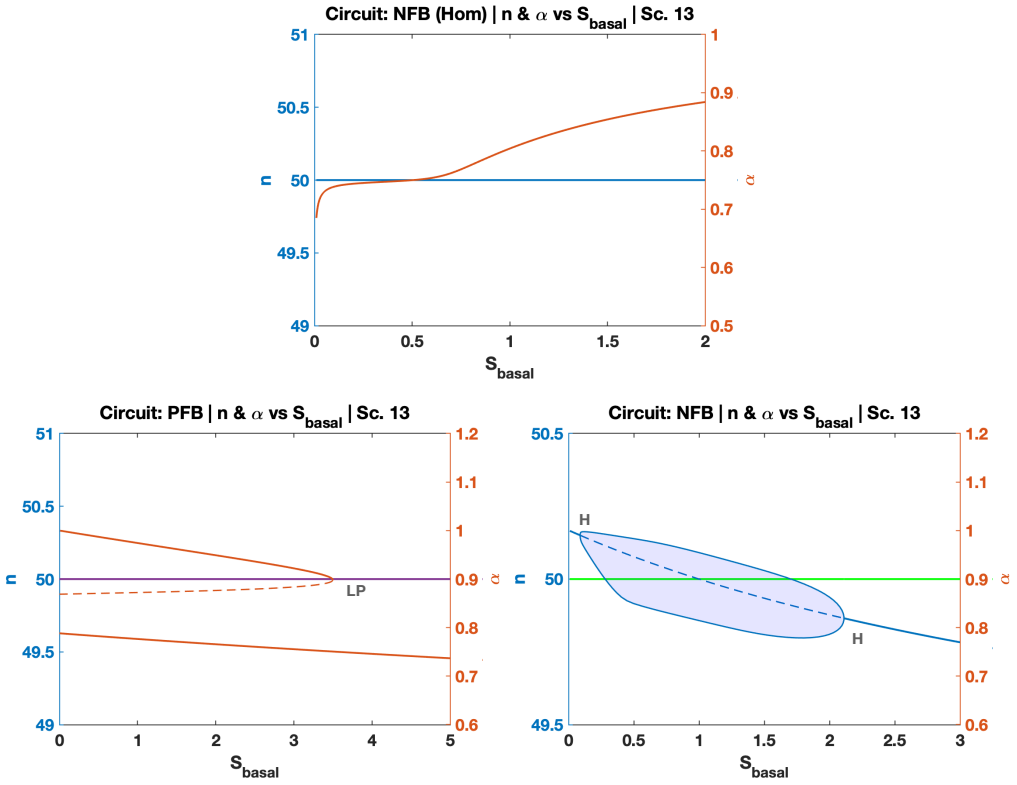



**Figure 5.9: Intra-cellular circuit can modulate and lend behaviors to environmental factors.** Plots show variation in population strength ($n$) and environmental factors ($\alpha$) as basal signal strength to the circuit ($S_{basal}$) is changed for three internal circuits; *Clockwise from top*: NFB (homeostatic), PFB, and NFB (oscillatory). These plots are representative of scenario 13, where the environmental factors and the cellular circuit bidirectionally interact with each other (environmental factors promote circuit response, while circuit degrades environmental factor). The plots above show how the circuit level in such scenarios can lend complex information processing behavior (homeostasis, multistability and oscillations) to the environmental factors. While on one hand intuitive, this information processing can affect factors downstream of the environment and can make a seemingly passive aggregate of environmental factors act as complex regulators of signaling. Exemplars from scenario 13, and results representative of scenarios 11 to 13.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

### *Effect of multi-level interactions*

In our analysis of the multi-level interactions where one level is capable of communicating with multiple others simultaneously (including bidirectional communication), we find that the information processing enabled by such communications is attenuated by the characteristic of the level through which information is channelled. For example,

if the information (qualitative behavior) flows through the population level and back again to the circuit or the environment, dynamic responses such as with oscillations (NFB and relaxation oscillations) from the circuit level are significantly dampened by the slower kinetics of the population level.

Similarly, depending on whether the environment is modulating the flow of information between the circuit and the population and back, environmental effects such as dampening of response are seen in the emergent systems response.

For example, fig. 5.10 shows how depending on whether the signalling characteristics of the circuit response (NFB leading to oscillations, and PFB leading to multi-stability) is channelled through the population, environment and back to the circuit (*top row* in fig. 5.10), or vice-versa (*bottom row* in fig. 5.10), the resulting systems behavior is modulated in a unique way. By comparing the response of the system between scenario 15 and 19, we can see how additional interactions through the environment directly from the circuit registers the circuit response across levels more strongly than when the interactions are through the population and then the environment, in which case we are able to observe the dampened response and behavior (scenario 19).

This dampening and modulating behavior is seen ubiquitously, whenever the circuit response is coupled directly to the population strength, and particularly with dynamic behaviors such as oscillations, where it's more pronounced. A common pitfall here is that, while population level experiments might not pick up on the minute dampened oscillations, the single circuit is still communicating with each other, and the environment, with this active dynamic circuitry.

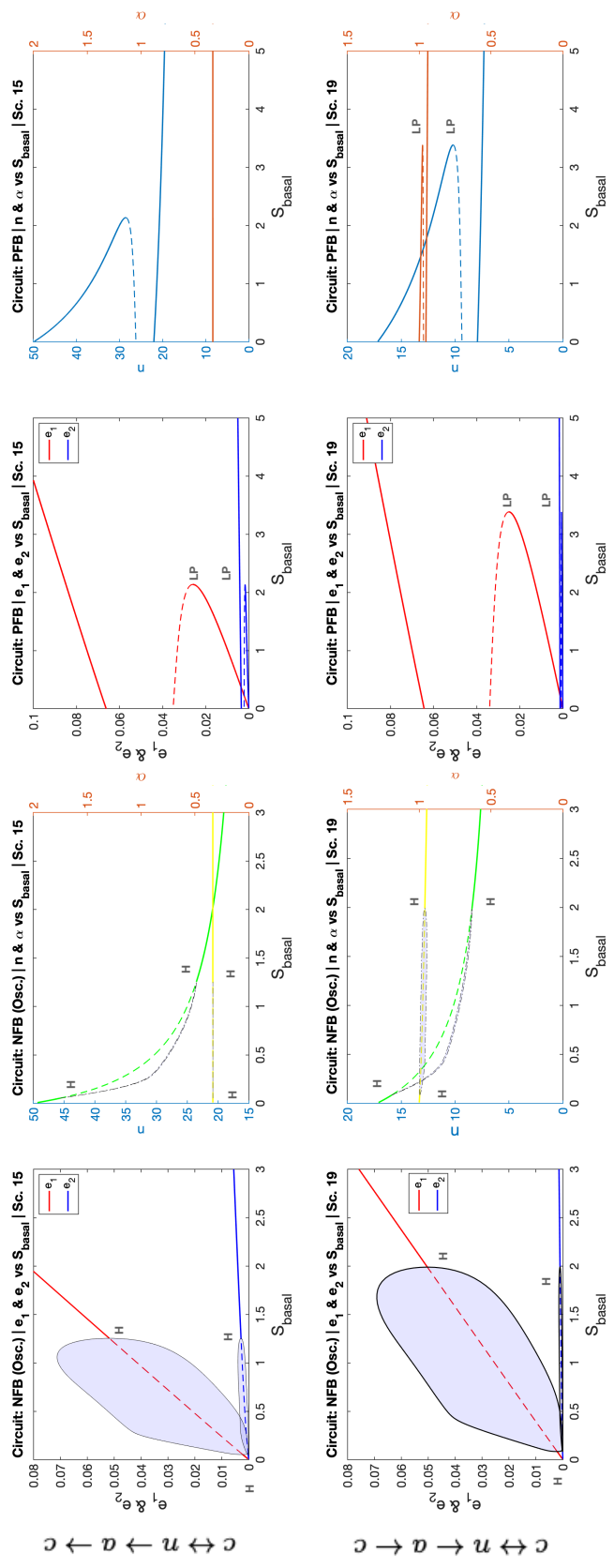



**Figure 5.10: Population level can dampen dynamic circuit response and systems behavior.** The plots above show two possible ways in which the environment can allow for a secondary unidirectional channel of communication between the circuit and the population (where the cellular and the population level are already communicating in a bidirectional manner). *Top row:* population strength affects the environmental factors, which in turn affects the circuit (scenario 15). *Bottom row:* circuit affects the environmental factors which in turn affects the population strength (scenario 19). By comparing the behavior of the system for PFB and NFB motifs when such secondary unidirectional interaction exists between the cellular level and the population, we can see that when the cellular modulates the population strength. The response registered at the system level is more pronounced and strengthens the primary bidirectional communication between the two levels (*bottom row*), as opposed to when the population unidirectional-ly interacts with the cellular level in which case the response is dampened (*top row*). Results representative of systems responses in scenarios 14 to 19.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

However active participation of the circuit level with the other levels leads to more robust and pronounced systems level behaviors at the population and environmental level, while still allowing for the population strength and environmental factors to modulate the circuit response. This is seen for example in the simulations of scenario 31 (see fig. 5.11), where the circuit has bidirectional communication with both environmental factors and the population strength.

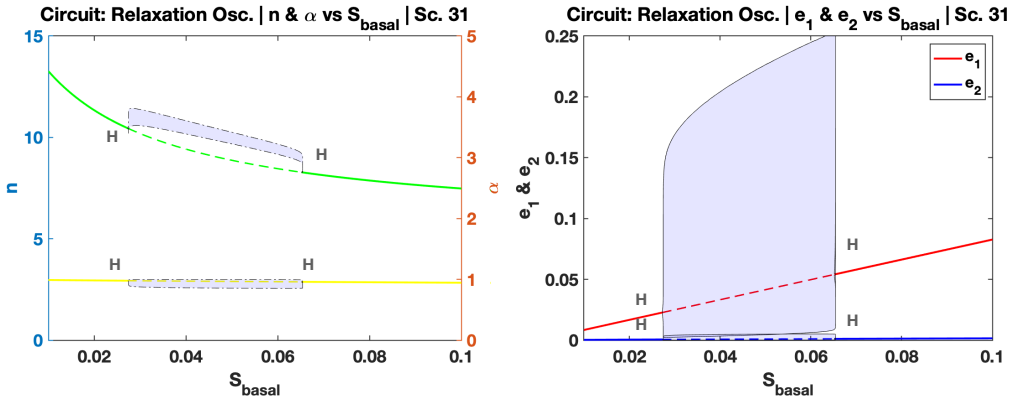



**Figure 5.11: Active and direct interactions with the circuit level allow for pronounced information transfer and communication for the population and environmental factors.** The plots above shows results from scenario 31 where all levels communicate with each other in a bidirectional manner. We can observe that all levels show circuit level response in a robust manner (in comparison to say behavior registered at the population strength $n$ and environmental factors in scenarios shown in fig. 5.10). In the plots above we can see noticeable oscillations in all variables (including environmental factors and population strength). Such bidirectional communication with the environment also allows for further external modulation of the behavior of the system. Representative of behavior in scenarios 15 to 37.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

This behavior is seen even when the multilevel interactions between circuit, population strength and the environment factors are unidirectional (Scenarios 38 to 41). However the strength of the coupling with circuit level is reduced in comparison to a scenario with unidirectional communication, and thus the effect of circuit response is more diluted in the levels further away from the cell, in the network graph (see fig. 5.12). The figure shows two scenarios with unidirectional interactions between the levels (containing a PFB and relaxation oscillator as circuit motifs); top row (scenario 39) - $c \leftarrow n \leftarrow a \leftarrow c$, bottom row (scenario 41) - $c \rightarrow n \rightarrow a \rightarrow c$. We can see how scenario 41 (bottom row) is able to register a stronger multistable and oscillatory behavior in population strength (by virtue of the circuit directly affecting it) than in scenario 39 (top row), where the population strength is further downstream from the circuit. Instead we can see how the environmental factors are more strongly modulated by the circuit by virtue of it being directly affected by the circuit.

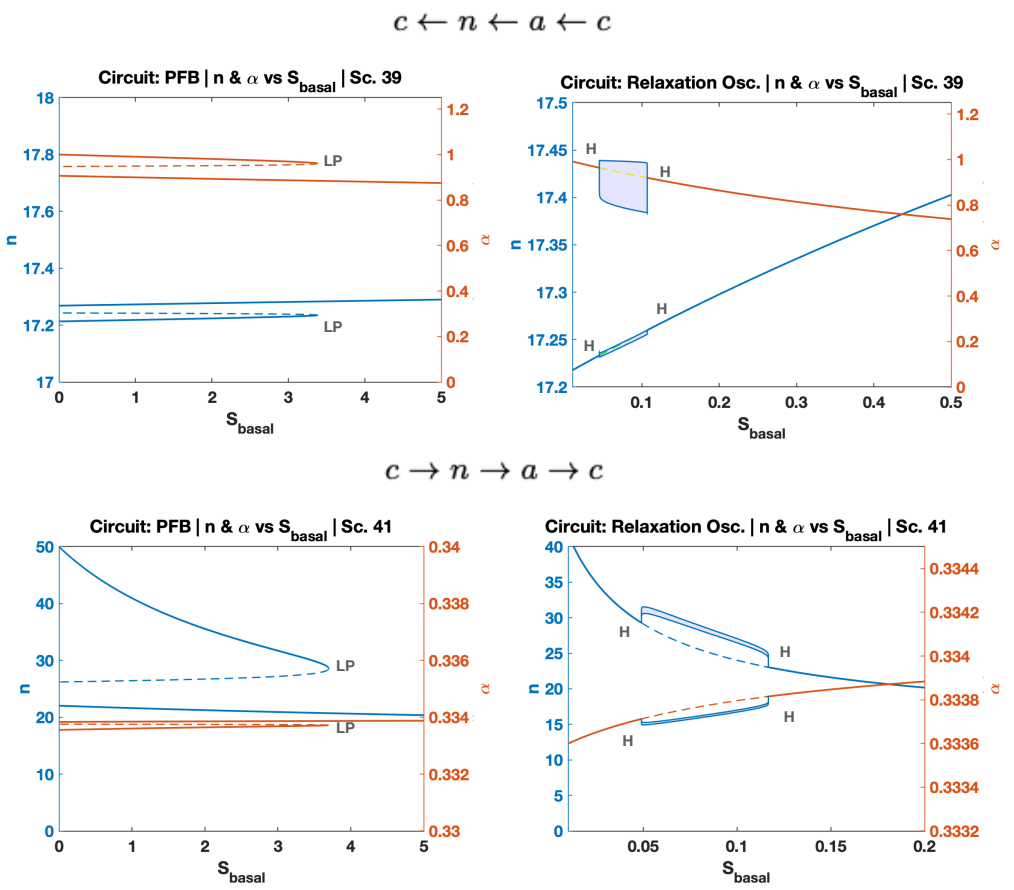



**Figure 5.12: Distance to circuit level dictates strength of information transfer from circuit level to other levels**. Plot shows variation in population strength ($n$) and environmental factors ($\alpha$) for different interaction scenarios and circuit motifs (*column 1*: PFB, *column 2*: Relaxation oscillator) *Top row:* scenario 39 where the circuit directly affects the environmental factor, which then affects the population strength, which then affects the circuit level. *Bottom row:* scenario 41 where the circuit directly affects the population strength, which then affects the environmental factors, which then affects the circuit level. In scenario 39 (*top row*), we can see how there is a pronounced effect on the environmental response and much less so on the population strength which is one step removed from the circuit. This is flipped in scenario 41 (*bottom row*), as is evidenced by the population strength showing a more pronounced response than the environmental factors. These plots reveal how with unidirectional interaction between the levels (even if it forms a closed loop of interactions), the further away a level is from the circuit level, the more diluted the transfer of behavioral characteristics from circuit level to it. This insight is equally relevant to other levels for behavior emergent from them. Results are representative of systems response in scenarios 38 to 41.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

These results indicate how the relative strength of information transfer is diluted by two factors, 1. the distance between the levels interacting (between interactions), 2. the dynamics of the intermediate levels mediating this response. The population level, with its characteristically slower dynamics can suppress and in many cases dilute the response coming from the circuit level. This has vast implications for the the transfer of information processing capability through the population, but also for inferences

on interactions based on experimental observations. Population level circuit response or the population strength itself, in certain cases can be a weak and even misleading proxy for the cellular response, eroding the expected intra & extra-cellular circuit motif response.

*Role of diffusion coefficient and cellular volume*

The role of diffusion coefficient and cellular volume act as further levers that can tune the communication between the cellular level and the environment and the population level. The model currently, as we have seen earlier, does not account for changing cellular volume (despite the population strength changing). However in order to study how this might perturb the systems response and particularly the amplitude/strength of communication between the cell and the population and environment, we performed bifurcation analyses on these parameters.

We observed that increasing diffusion coefficient ($D$) lead to an (steady but saturating) increase in strength of communication from the cell to the other levels. This manifested in the relevant scenarios as a more pronounced response at the population level or environmental level or both. However this effect has a saturating effect, and indefinite increase of diffusion coefficient does not lead to an ever increasing trend in the strength of the communication. This can be observed in fig. 5.13 (top row, scenario 19), where we can see as diffusion coefficient is increased (keeping all else the same), how for the same underlying circuit (NFB - Homeostatic motif), the homeostatic response registered at the population level (and environmental level) is more pronounced at higher diffusion constants. This is directly due to the higher concentrations of extra-cellular species $e_2$ at higher diffusion rates, for the same circuit motif response and signal strength.

The cellular volume ($V_1$) played a similar role in perturbing extra-cellular species concentration. In particular, increasing cellular volume counteracted the effect due to increasing diffusion coefficient. This is intuitive, from the model, as a larger cell size would imply a lower concentration for the same given signalling response from the circuit motif ($cR$), that ultimately affects the extra-cellular species diffusion and eventually the extra-cellular concentration ($e_2$) and the resulting interactions. These behaviors were seen ubiquitously for all cellular circuits and for all scenarios (where interactions between the circuit and the population or the environment exist). fig. 5.13 (bottom row, scenario 19), shows one such example where increasing cellular volume resulted in a reduced population and environmental homeostatic response (borrowed from the NFB intra-cellular circuit motif).

Diffusion coefficient and the cellular volume thus in different ways control the dilution of circuit response species out of the cell ($e_1$ & $e_2$), and thus as a consequence play major roles in modulating the population level response and the emergent collective behavior seen from cells.

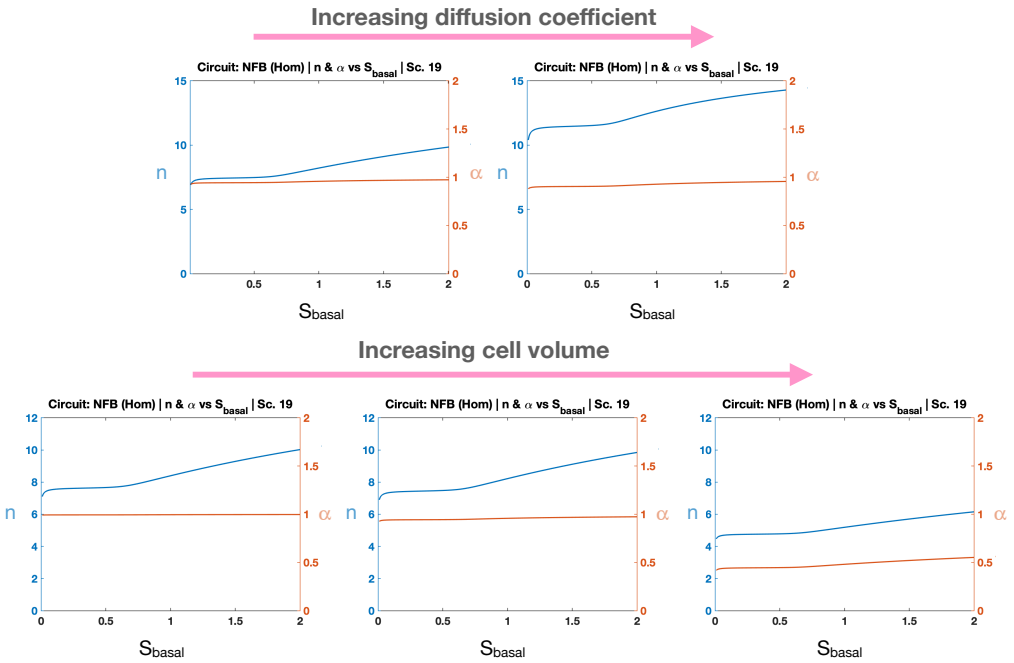



**Figure 5.13: Role of diffusion coefficient and cell volume on strength of interactions between circuit and other levels.** Plot shows variation in population strength ($n$) and environmental factors ($\alpha$) as basal signal strength to the circuit ($S_{basal}$) is changed. The circuit motif is NFB (homeostatic) and the scenario simulated is 19 (where cell bidirectionally interacts with the population, and also affects the environmental factors which in turn affect the population strength). Increasing diffusion coefficient (*top row*) can allow for a strengthened response from the circuit level to the population level, as it immediately addresses the rate at which $e_1$ (intra-cellular circuit response) can diffuse out, and the eventual equilibrium concentration. Thus any cell-environment or cell-population interaction (which is mediated by $e_2$) gets strengthened as shown above. Similarly increasing cell volume in our study has the opposite effect as it essentially dilutes the extra-cellular concentration that diffused out of the cell. This leads to a weakened response from the cell circuit to all other levels (bottom row). Cells however as they grow do not have a fixed volume, and as the population size grows, the average cell volume is assumed to be constant here - see text for discussion. These results are representative of results from all scenarios 1 to 41.

## 5.4 Two populations study

### 5.4.1 Mathematical model

The two population setup can be modelled using a similar higher level description (as in the single population study). The model (and the system) can itself be interpreted as the interplay between five classes of variables, spread between two populations (population 1 & 2, cellular level 1 & 2 and environment). However, our focus here is on cross-population interactions between two populations of species. Further we have also clearly discerned how the environmental level can affect or get affected by a population of cells and any intra-population inter cations. Hence, in this two populations study, in order to better delineate causality and the role of interactions in enabling collective population behavior, we disregard active consideration of the environmental

level. This leaves us with two levels for each of the population of cells, 1. population level ($n$), and 2. the cellular level ($c$ and $e$).

**Cell type 1**

$$
\begin{aligned}
\frac{dn_1}{dt} &= k_{gn1}\left[2+\textcircled{$Sw1^0$}\frac{k_1 n_2}{K1+n_2}+\textcircled{$Sw2^0$}\frac{k2e_{2_{ext}}}{K2+e_{2_{ext}}}+\textcircled{Sw3}\frac{k3e_1}{K3+e_1}\right]-k_{dn1}n_1\\
\frac{dc_1}{dt} &= f_1(S_1)\\
\text{where}\quad S_1 &= S_{basal}\left[1+\textcircled{Sw4}\frac{k_4 n_2}{K_4+n_2}+\textcircled{Sw5}\frac{k_5 n_1}{K_5+n_1}+\textcircled{Sw6}\frac{k_6 e_{2_{ext}}}{K_6+e_{2_{ext}}}\right]\\
\frac{de_1}{dt} &= k_{a_{e_1}}c-k_{d_{e_1}}e_1+D\left[e_{1_{ext}}-e_1\right]\\
\frac{de_{1_{ext}}}{dt} &= k_{e_{1pr}}+D\left(\frac{n_1 V_1}{V_{env}}\right)\left[e_1-e_{1_{ext}}\right]-kd_{e_{1_{ext}}}e_{1_{ext}}
\end{aligned}
\tag{5.2}
$$

**Cell type 2**

$$
\begin{aligned}
\frac{dn_2}{dt} &= k_{gn2}\left[2+\textcircled{$Sw7^0$}\frac{k_7 n_1}{K7+n_1}+\textcircled{$Sw8^0$}\frac{k8e_{1_{ext}}}{K8+e_{1_{ext}}}+\textcircled{Sw9}\frac{k9e_2}{K9+e_2}\right]-k_{dn2}n_2\\
\frac{dc_2}{dt} &= f_2(S_2)\\
\text{where}\quad S_2 &= S_{basal}\left[1+\textcircled{Sw10}\frac{k_{10} n_1}{K_{10}+n_1}+\textcircled{Sw11}\frac{k_{11} n_2}{K_{11}+n_2}+\textcircled{Sw12}\frac{k_{12} e_{1_{ext}}}{K_{12}+e_{1_{ext}}}\right]\\
\frac{de_2}{dt} &= k_{a_{e_2}}c-k_{d_{e_2}}e_2+D\left[e_{2_{ext}}-e_2\right]\\
\frac{de_{2_{ext}}}{dt} &= k_{e_{2pr}}+D\left(\frac{n_2 V_1}{V_{env}}\right)\left[e_2-e_{2_{ext}}\right]-kd_{e_{2_{ext}}}e_{2_{ext}}
\end{aligned}
$$

Similar to the single population study, each *population (1 & 2)* has a basal growth rate ($k_{gn}$) and a death/decay rate ($k_{dn}$). The basal growth rates can get affected/modulated by the other two levels (of the same population or other other population) in many ways,

1. The growth rate is strengthened by the other populations' density (Sw1) = 1
2. The growth rate is weakened by the other populations' density (Sw1) = -1
3. The growth rate is increased by the circuit response of the other population (Sw2) = 1
4. The growth rate is increased by the circuit response of the other population (Sw2) = -1
5. The growth rate is increased by the circuit response of the same population (Sw3) = 1

The switches that have a superscript (0), can take values +1 / -1 / 0 (unlike in the single population study where all switches could take only one of two values: +1 / 0). Thus in this model, this single switch parameter allows for us to account for both positive

and negative interactions between the levels. However, the negative interaction does not change the fundamental sign of the basal growth rate, as the leading constant parameter is scaled to 2, meaning the coefficient of $k_{gn}$ is always positive. In other words, our choice to have a negative switch value of -1, is similar to the introduction of a new switch that possess the following functional structure to indicate negative interactions $\frac{k}{K+x}$, with the inhibition incorporated in the denominator of a function[2].

This allows for a complex set of interactions that can modulate the two populations, with only a minimal model description (and switches). The rest of the model description is akin to the single population study.

*The intra-cellular circuit* is considered to be modular as explained above. The basal signal of the circuit can be strengthened however by the population strengths, and the circuit response (of the other population). An additive structure is assumed here for these effects.

The circuit motif response $c$ (or $cR$) however in turn produces an *intra-cellular circuit response* ($e_1$), which has its own basal linear degradation within the cell ($k_{d_{e1}}$). This intra-cellular circuit response ($e_1$) can diffuse out of the cell as extra-cellular circuit response ($e_2$). We assume that there is a basal production of this extra-cellular circuit response ($e_2$) in our model in the extra-cellular environment (to account for such production from the intra-cellular response of another cell type for example). *The extra-cellular response* $e_2$ is assumed to also degrade (linearly) in the extra-cellular environment.

### Circuits

For this study, we consider the following three intra-cellular circuits as representative examples of intra-cellular circuitry (from the 5 considered in the single population study).

| | |
|---|---|
| 1. Covalent modification system | Monostable |
| 2. Positive feedback system | Multi-stability |
| 3. Negative feedback system | Homeostatic |

In limited scenarios we also study the role of substrate depletion oscillator model to explore the effect of relaxation oscillations on the system. The circuit level response from these motifs is the same as it was for the single population study and is shown in fig. 5.3.

### Interactions

The model described has 12 interactions between the different levels. These interactions are all contained within the master model and can be switched on and off with a

[2]That is since $\left[2 - \frac{x}{K+x}\right] = \left[1 + \frac{K}{K+x}\right]$

suitable choice of parameters for the model variables (Sw1) to (Sw12). Below we list them again and expand on the modelling choice for the interaction term.

Similar to the single population study, each interaction is modelled with a Michaelis–Menten description. This choice is made to account for a saturating effect within the interaction and to ensure there is no run away response from the system.

| | | |
|---|---|---|
| $\boldsymbol{Sw1^0}$ | $Pop_2 \rightarrow Pop_1$ | Population 2 affecting the growth rate of population 1 (effect can be positive or negative) |
| $\boldsymbol{Sw2^0}$ | $Cir_2 \rightarrow Pop_1$ | Extra-cellular circuit response of cell type 2 affecting growth rate of population 1 (effect can be positive or negative) |
| **Sw3** | $Cir_1 \rightarrow Pop_1$ | Intra-cellular circuit response of cell type 1 increases growth rate of population 1 |
| **Sw4** | $Pop_2 \rightarrow Cir_1$ | Population density (cell type 2) increases basal signal activation of circuit within cell type 1 |
| **Sw5** | $Pop_1 \rightarrow Cir_1$ | Population density (cell type 1) increases basal signal activation of circuit within cell type 1 |
| **Sw6** | $Cir_2 \rightarrow Cir_1$ | Extra-cellular circuit response from cell type 2 increases basal signal activation of circuit within cell type 1 |
| $\boldsymbol{Sw7^0}$ | $Pop_1 \rightarrow Pop_2$ | Population 2 affecting the growth rate of population 2 (effect can be positive or negative) |
| $\boldsymbol{Sw8^0}$ | $Cir_1 \rightarrow Pop_2$ | Extra-cellular circuit response of cell type 1 affecting growth of population 2 (effect can be positive or negative) |
| **Sw9** | $Cir_2 \rightarrow Pop_2$ | Intra-cellular circuit response of cell type 2 increases growth rate of population 2 |
| **Sw10** | $Pop_1 \rightarrow Cir_2$ | Population density (cell type 1) increases basal signal activation of circuit within cell type 2 |
| **Sw11** | $Pop_2 \rightarrow Cir_2$ | Population density (cell type 2) increases basal signal activation of circuit within cell type 2 |
| **Sw12** | $Cir_1 \rightarrow Cir_2$ | Extra-cellular circuit response from cell type 1 increases basal signal activation of circuit within cell type 2 |

**Assumptions**

A number of key assumptions are made in the model (akin to the single population study). The main difference however being that, unlike in the single population study (where the focus was on interactions between cell types of a given population and the environment), here we focus on cross-population interactions (cell to cell and cell to population). As such the focus on the environmental effects and intra-population interactions are limited.

1. The **populations (1 & 2)** grow and decay at a basal level which can be affected by circuit responses ($e_1$ or $e_2$) and the environmental factors ($\alpha$).
2. The **circuit** motifs are modular. The circuit responses are $c_1$ & $c_2$ and lead to the production of their respective intra-cellular species ($e_1$ & $e_2$). This is decoupled

to allow for diffusion and sequestration in extra and intra-cellular environment without affecting the modularity of the intra-cellular circuit.

3. The **intra-cellular species** ($e_1$ & $e_2$) can diffuse out of the cell and the dilution is accounted for by taking the volumes (the cell and the extra-cellular environment) in to account. The intra-cellular volume is assumed to be insignificant in comparison to the extra-cellular value.

4. The **extra-cellular environmental volume** is considered fixed, see parameters for more details.

5. The **extra-cellular species** ($e_1$ & $e_2$) have linear rate degradation to account for any environmental degradation.

6. **Both cell types have the same volume** ($V_1$). $V_{env}$ is the volume of the enviornment.

### 5.4.2 Scenarios

As mentioned earlier, in this section the focus is on trying to understand the role of different interactions in enabling effective communication between two populations of cells.

We do this by simulating a number of scenarios with limited interactions and circuitry active. These scenarios are constructed by toggling the switches on and off (and in the case of $Sw1$, $Sw2$, $Sw7$, $Sw8$ taking values -1, 0, 1) in the master model described above in eq. (5.1). The scenarios are built into clusters, starting with the systems with the least interactions and complexity, and building on top of them; at each stage carrying forward the insights from the previous cluster of scenarios to better inform further simulations. These clusters and scenarios are listed in the table below.

Unless explicitly mentioned below the interaction is absent in the model (i.e, every switch (Sw) is equal to 0)

**Table 5.3:** Scenarios used to analyze role of interactions in enabling population level response (two populations study)

| Sc. | Active Levels: Pop.-Pop. $Sw1^0$ | Pop.-Pop. $Sw7^0$ | Cir.-Pop. $Sw2^0$ | Cir.-Pop. $Sw8^0$ | Cir.-Cir. | Interactions | Description | Reference |
|---|---|---|---|---|---|---|---|---|
| 1 | × | | × | | × | None | - | 5.14 |
| | **Only population level interactions** | | | | | | | |
| 2 | 1 | 0 | × | | × | Sw1 | | |
| 3 | -1 | 0 | × | | × | Sw1 | Study only the population level response and focus on | |
| 4 | 1 | -1 | × | | × | Sw1, Sw7 | the emergent population dynamics between the | 5.15 |
| 5 | 1 | 1 | × | | × | Sw1, Sw7 | populations | |
| 6 | -1 | -1 | × | | × | Sw1, Sw7 | | |
| | **Only circuit level interactions** | | | | | | | |
| 7 | × | | × | | ✓ | Sw6 | Study how cellular circuitry interactions between | 5.16 |
| 8 | × | | × | | ✓ | Sw6, Sw12 | different cell types modulates population level circuit response | |
| | **Only population-circuit interactions** | | | | | | | |
| | | | | | | *Circuit to population only* | | |
| 9 | × | | 1 | 0 | × | Sw2 | | |
| 10 | × | | -1 | 0 | × | Sw2 | Study multi-level cross populations, with a focus on | |
| 11 | × | | 1 | -1 | × | Sw2, Sw8 | how circuit to population response modulates | 5.17 and 5.19 |
| 12 | × | | 1 | 1 | × | Sw2, Sw8 | population dynamics | |
| 13 | × | | -1 | -1 | × | Sw2, Sw8 | | |
| | | | | | | *Population to circuit only* | | |

Continued on next page

Table 5.3 – continued from previous page

| Sc. | Active Levels: Pop.-Pop. $Sw1^0$ | Pop.-Pop. $Sw7^0$ | Cir.-Pop. $Sw2^0$ | Cir.-Pop. $Sw8^0$ | Cir.-Cir. | Interactions | Description | Reference |
|---|---|---|---|---|---|---|---|---|
| 14 | × | | × | | ✓ | Sw4 | Study multi-level cross populations, with a focus on how population to circuit response modulates circuit response | 5.20 |
| 15 | × | | × | | ✓ | Sw4, Sw10 | | |
| | | | | | | *Circuit and Population interactions together* | | |
| 16 | × | | 1 | 0 | ✓ | Sw2, Sw4 | Study the role of bidirectional multi-level pop-circ response in modulating population dynamics and cellular response | 5.21 |
| 17 | × | | 1 | 0 | ✓ | Sw2, Sw4, Sw10 | | |
| 18 | × | | -1 | 0 | ✓ | Sw2, Sw4 | | |
| 19 | × | | -1 | 0 | ✓ | Sw2, Sw4, Sw10 | | |
| 20 | × | | 1 | 1 | ✓ | Sw2, Sw8, Sw4 | | |
| 21 | × | | 1 | 1 | ✓ | Sw2, Sw8, Sw4, Sw10 | | |
| 22 | × | | 1 | -1 | ✓ | Sw2, Sw8, Sw4 | | |
| 23 | × | | 1 | -1 | ✓ | Sw2, Sw8, Sw4, Sw10 | | |
| 24 | × | | -1 | -1 | ✓ | Sw2, Sw8, Sw4 | | |
| 25 | × | | -1 | -1 | ✓ | Sw2, Sw8, Sw4, Sw10 | | |
| | **All population-circuit interactions** | | | | | | | |
| 26 | 1 | 1 | 1 | 1 | ✓ | All interactions active | Study the overall systems behavior of two populations that interact with each other at every level and across levels | 5.22 |
| 27 | 1 | -1 | 1 | 1 | ✓ | All interactions active | | |
| 28 | 1 | 1 | 1 | -1 | ✓ | All interactions active | | |

**Exemplar cases** (All level interactions active)

*Assumption: The population to circuit interactions (across population) are ignored in this cluster of scenarios*

Continued on next page

Table 5.3 – continued from previous page

| Sc. | Active Levels: Pop.-Pop. $Sw1^0$ | Pop.-Pop. $Sw7^0$ | Cir.-Pop. $Sw2^0$ | Cir.-Pop. $Sw8^0$ | Cir.-Cir. | Interactions | Description | Reference |
|---|---|---|---|---|---|---|---|---|
| | | | | | | $n_2 \leftrightarrow c_2 \leftrightarrow c_1 \leftrightarrow n_1$ | | |
| 29 | 1 | 1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5 | | |
| 30 | -1 | 1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5 | | |
| 31 | -1 | -1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5 | Study the overall systems behavior of two populations that interact with each other at every level, and within a population across levels, but not across levels and across populations | 5.23 |
| | | | | | | $n_1 \leftrightarrow c_1 \leftrightarrow c_2 \leftrightarrow n_2 \leftrightarrow n_1$ | | |
| 32 | 1 | 1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5, Sw9, Sw11 | | |
| 33 | 1 | -1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5, Sw9, Sw11 | | |
| 34 | -1 | -1 | | × | ✓ | Sw1, Sw7, Sw6, Sw12, Sw3, Sw5, Sw9, Sw11 | | |

**Reasoning**

When two distinct populations interact with each other to elicit a collective populations level systems response, they can do so in multiple ways, including while simultaneously interacting within their own population of cells. Thus in order to study the overall systems landscape, the number of scenarios with combinations of interactions (as we did with the single population study) one might need to consider within our setup is large. To tackle this complexity, we opt to build on fundamental pair wise interactions to explore and anticipate the role of any given class of interaction in the final multi-level cross-population response.

We do so by first focussing on the population level interactions, which represent some of the more well studied and characterized population level interactions between multiple populations [195]. Once a good understanding of this is accomplished, we study the role of intra-cellular circuitry and interactions through diffusing species arising from such cellular response. We then consider both interactions together to explore the role of multi-level, circuit-population interactions. This cluster arguably is combinatorially rich, yet can reveal the multi-level systems behavior of the emergent system in great detail. With this, we then explore other multi-level interactions possible without the need for explicit population-circuit interactions through the study of exemplar cases where the population-circuit interaction within a single population is restored (whose insights have already been explored in the single population study). Again, within these instances, leveraging our insights from the single population study, we focus on the cross population multi-level interactions.

This structured approach allows us to understand emerging behavior when all of these are present together (the final scenarios considered).

### 5.4.3 Questions

Again similar to the single population study, while these questions don't represent the full scope of the insights obtained (see section 5.4.4 for a detailed synthesis), they offer a purview in to the initial focus behind the study.

- How do two populations interact with one and other to generate population level systems response?
- How do circuit interactions modulate population level interactions (when there are more than a single population of cells)?
- How do different cell-population cross interactions (contact, extra-cellular, metabolite) compare in effecting robust population level response?
- How does dilution and diffusion affect such population level response?

### 5.4.4 Results and Discussion

The simulations were run in the order of the scenarios listed in the table above (see table 5.3), using the automation pipeline described in the computational methodology (see section 5.2.1). The results below are organized in the same spirit and style as the scenarios. We first describe how the two populations study was constructed and descriptions of the many scenarios considered before describing the results.

The study considers two distinct populations (1 and 2), that interact in a number of ways with one and other; at the population level, at the circuit (cellular) level, and across levels. The circuit (cellular motif) response of each population is modular. This response then produces intra-cellular species ($e_1$ and $e_2$) that diffuse out of the cell ($e_{1_{ext}}$ and $e_{2_{ext}}$). These species then enable the circuit to interact with the population dynamics (of both its own population and others) and circuit level of the other population. However unlike the single population study, we consider a limited set of cellular motifs, namely, the covalent modification cycle, NFB and PFB systems. Despite the reduced number of motifs the number of ways in which these can be paired up or distributed across two populations significantly expands the combinations of scenarios and models analyzed. The most interesting and illuminating results are presented here, however the computational pipeline generated (see section 5.2.1) contains the entirety of the plots generated.

Thus with this representation in a scenario with no interactions (see fig. 5.14), the population level strengths ($n$) are constant, while the circuit response ($cR$) and its associated intra and extra-cellular species concentration ($e$ & $e_{ext}$) closely follows the circuit motif response. This represents the benchmark against which we compare impact of any interactions between the levels. Note that the circuit response is the same as that in the single population study (since the circuit motif is modular, see fig. 5.4 and fig. 5.3).

As seen in the single population however, there is a dependency between the extracellular response ($e$ and $e_{ext}$) and the population level ($n$), due the diffusion of the species and the dilution effect. Due to this, even with no explicit interaction between the population level and the circuit level, the extra-cellular circuit response $e_{ext}$ (and the intra-cellular response $e$) will trail the population level (as we have seen in the single population study). This behavior is seen and discussed more prominently below.

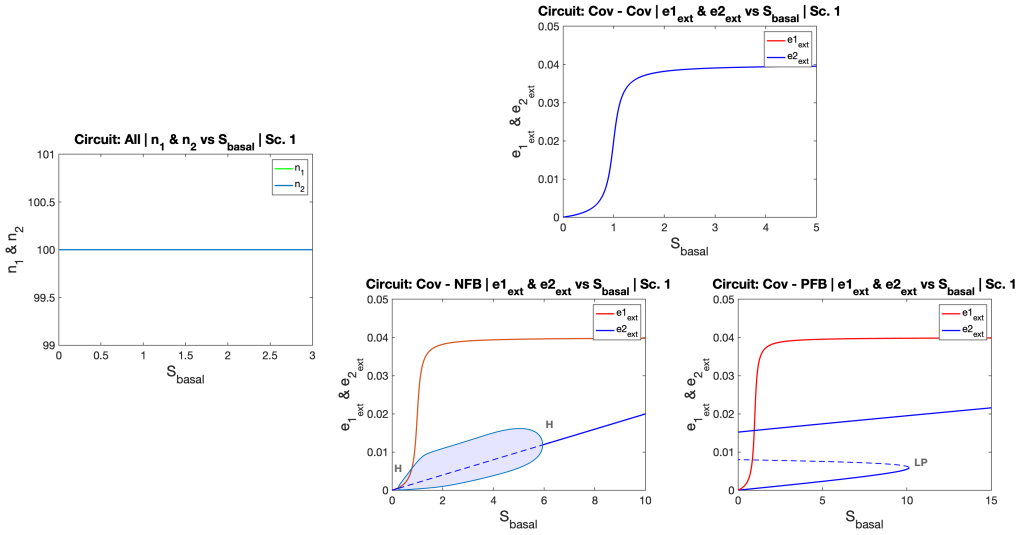



**Figure 5.14: Two populations study: Population strengths and extra-cellular circuit response in the absence of any interactions between the cell, population and the environment.** Results shown for scenario 1 (where there are no interactions between the levels). *Left*: population strengths ($n_1$ & $n_2$) as circuits basal signal activation is changed. *Right*: extra-cellular circuit response ($e1_{ext}$ & $e2_{ext}$) as circuits basal signal activation is changed, for different circuit motifs within the populations. Clockwise from top: Cov & Cov in population type 1 & 2, Cov & NFB in population type 1 & 2, and Cov & PFB in population type 1 & 2. This behavior seen with no interactions serves as a benchmark against influence of interactions across levels and across populations. The population strengths (*left*) are constant due to the absence of any interactions for all circuit level motifs, however depending on the choice of internal circuit motifs, the extra-cellular concentrations ($e1_{ext}$ & $e2_{ext}$) mirrors that of the motif responses ($cR$).
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

### *Role of population level interactions*

The first set of interactions we considered were purely at the population level. Our model allows us to explore the behavior of the system under a number of scenarios where a single population modulates the growth and death kinetics of the other population and vice versa. These scenarios have been analyzed in detail in literature elsewhere, particularly more recently in studies with synthetic microbial communities and microbiomes [196].

However here, unlike the other clusters to be seen later on, the interactions are largely uncoupled from the internal circuits of the cell, and exist purely at the population level. Thus in order to explore the true relevance of the population level only interactions, we choose to perform bifurcations along the growth rate of a single population (arbitrarily chosen as that of population 1). These results are shown in fig. 5.15.

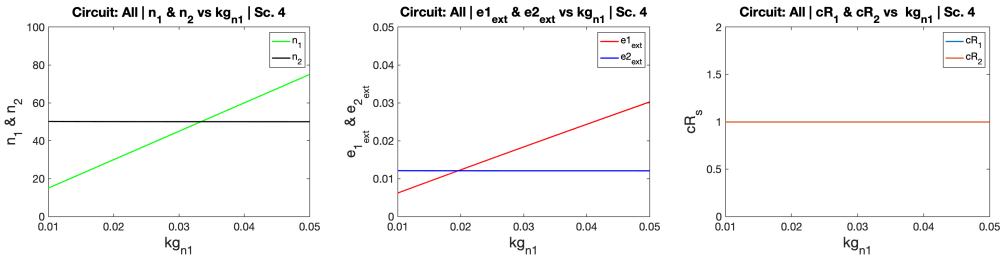



**Figure 5.15: Role of purely population level interactions in modulating cellular response.** *Left:* Increasing population strength of population 1 promotes decay (and weakens) population strength of population 2 (scenario 2). *Middle:* The circuit response is modular and is unaffected by the population level dynamics and is thus constant. *Right:* The extra-cellular circuit response tracks the population response and we see the same effect of increased population strength strengthening the extra-cellular circuit response. These results are indicative of how population level, while uncoupled from the other levels (through the absence of interactions) can still manipulate and perturb the expected response from the circuit both at the intra-cellular level ($e1$ & $e2$) and the extra-cellular level ($e1_{ext}$ & $e2_{ext}$). Results representative of scenarios 2 to 6.

We can observe simple modulation of the population strength of the two populations (*left* plot) as the growth rate of population 1 is changed. This is intuitive as in this set up, the population 1 promotes the decay rate of population 2 as it grows. We see this behavior. In addition we also see how the circuit response (which was designed to be modular) is unaffected by the population level dynamics (*middle* plot). However, we can observe that the extra-cellular circuit responses of the two cell types are modulated by the changing population strengths. This is similar to the dilution effect we observed in the single population study (due to the shuttling of the species inside and outside of the cells), where increasing population strength strengthened extra-cellular circuit response (see fig. 5.8). Together these results reflect how even with two populations the dilution effect between a populations' population level strength and the circuit response can modulate observed intra and extra-cellular circuit response. Our model however does not consider complex highly non-linear interactions between the two population (even with bidirectional communication between them). In our model, the population level interactions and dynamics are governed by linear growth-death dynamics. This allows us to clearly understand and study the different sources of complexity arising from interactions with other levels including feedback, delay, kinetics etc,. However there are several examples of populations interacting with one and other at the population level that can lead to complex highly non-linear dynamics, with the predator prey dynamics being a canonical example (see [195], for more examples of such population level behavior including in synthetic communities). Such complex population level dynamics and its role in enabling collective population level response is discussed in the sub-study (see section 5.5).

*Role of cellular interactions between circuits across populations*

We then considered interactions between the circuits of the cell directly (scenario 7 and 8). These scenarios consisted of both unidirectional and bidirectional communication

between the two populations' circuit levels. Simulations of the system with such interactions revealed how information processing capabilities can be lent from one population of cell types to another to allow for complex patters of information processing to emerge. This is possible in one cell type (as when there is unidirectional communication between one population type to another), and in both cell types (as with bidirectional communication). Figure 5.16 shows how such complex information processing capabilities are possible in the landscape of circuit combinations we have considered, even with simple unidirectional communication. This behavior spills over naturally into the extra-cellular response of the circuit (not shown).

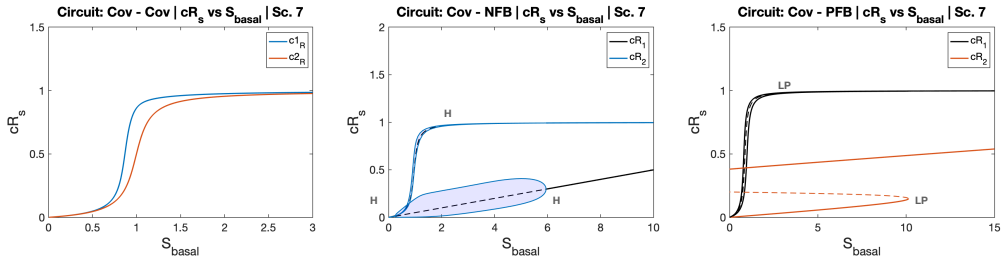



**Figure 5.16: Emergent collective information processing capabilities when cellular circuits communicate across populations (unidirectional).** Results shown from scenario 7, where cellular circuit of cell type 2 (Cov) increases basal activation of the cellular circuit of cell type 1 (*left*: Cov, *middle*: PFB, *right*: NFB). The figure shows how a population (2) with cellular circuit capable of switch-like behavior, bistability, and oscillation can perturb a switch-like response in another population (1) in a non-intuitive way, leading to a range of complex information processing behavior incapable from any given population alone independently. Such characteristic combination of circuit motif response from two levels not only modulates the range over which such qualitative behavior is seen in such systems (perturbation on range and strength of switch-like response in the three cases for example), but also bring additional features to it, such as strengthened switch-like behavior (*left*), oscillations with a switch-like characteristic oscillatory bound (*middle*), and bistability with both stable steady states featuring a switch-like response (*right*). Similar behavior is possible with perturbation of other cellular motifs (PFB and NFB, for example).
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

In fig. 5.16, we can see how a simple negative feedback motif can modulate and introduce oscillations with a switch-like response in the response of another populations' circuit (*middle* plot). Similarly, a PFB cellular circuit can introduce bands of steady states around a switch-like response output (*right* plot). These architectures and signaling patterns can be highly relevant in the context of synthetic biologys efforts to engineer population level response. Simple coupling of the two populations at the circuit level can allow access to a complicated and rich set of information processing tools and capabilities, leveraging qualitative signaling patterns that are otherwise biologically expensive to engineer robustly.

This landscape can be made more rich with the presence of bidirectional communication between the cellular motifs. A striking example of this showing combinations of multistability and oscillations (arising in scenario 8) is shown below (fig. 5.17). We can

observe how PFB from population 1 is able to modulate different regions of oscillations within the NFB motif. This includes multistability in the system behavior along with oscillations for certain ranges of signal activation. This is but an example of how complex behaviors could arise at the collective systems level from circuit level interactions. These behaviors further can modulate other cells/processes and environmental factors downstream transferring this information processing capability further beyond the cellular level.

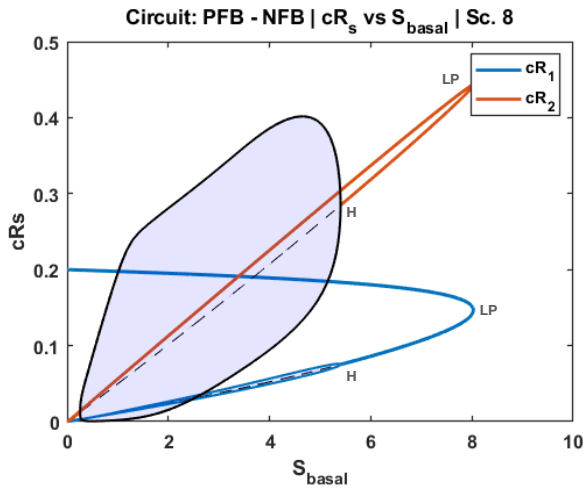



**Figure 5.17: Emergent collective information processing capabilities when cellular circuits communicate across populations (bidirectional).** Result from scenario 8, where the circuit levels of each cell type strengthen the basal activation of their counterparts. Complex information processing capabilities are introduced with bidirectional communication between the cellular circuits of two populations, including a combination of bistability and oscillations. In the above scenario simulated, the cellular circuits (PFB from population 1 and NFB from population 2), interact with one and other directly at the circuit level through the aid of extra-cellular circuit responses. This leads to a complicated information processing capability at both levels, where each level both lends its behavioral characteristic to the other, and gets modulated simultaneously. Here we see multiple hopf points in the system bifurcation and a complicated qualitative systems landscape with multistability and oscillations arising at multi-population level.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

*Circuit level modulating population response*

The effect of cellular level (circuit) modulating population response of the other cell type is intuitive to understand and is studies next. When such cross-level cross-population interactions is one way (scenario 9 and 10, as in sender receiver cells), the population level of cell type 2 mirrors the circuit motif response, albeit with a slight dampening of the features (see fig. 5.18).

Figure 5.18 shows how the population level $n_2$ is able to mirror the circuit response from population 1 (scenario 10). The population response is able to show switch-like response response (*left* plot), oscillatory behavior (*middle* plot) and bistability (*right* plot).

As mentioned earlier however the amplitude and strength of these responses, particularly the dynamic responses are arrested significantly at the population level due to the slower dynamics of the population level. This can be seen more pronounced in the *middle* plot, where NFB circuit motif from population 2 impacts population 1s strength. WE can see the overall effect on reducing population strength as circuit response increases, however, the oscillations (dynamic behavior) is significantly suppressed at the population level, vis-a-vis, that seen at the circuit level (refer fig. 5.3).

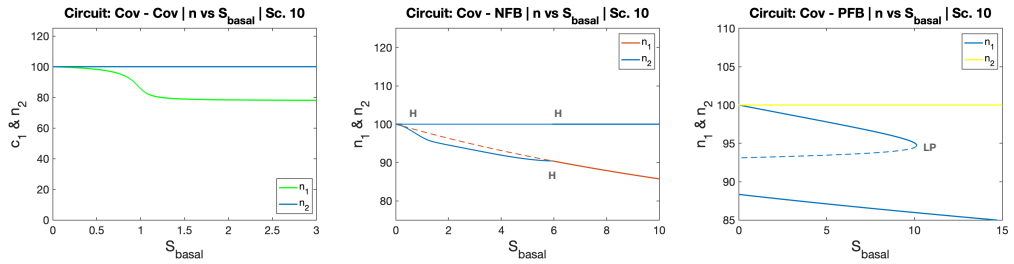



**Figure 5.18: Effect of cellular level interacting at the population level (cross-population & unidirectional from one population).** Results from scenario 10 where the cellular level of population 2 decreases the growth rate of population 1. Figure shows how the cellular level can interact with the population level (in a unidirectional) manner and lend qualitative behaviors to the population. *Left:* shows switch-like response at the population level. *Middle:* shows oscillatory behavior at the population level. *Right:* multi-stability at the population level. The characteristic feature of these responses at the population level is that their strength or amplitude are diminished vis-a-vis the cellular motif response. This is particularly seen in dynamic behavior where the oscillations at the population level have a much smaller amplitude than the cellular motif counterpart (see fig. 5.3). This arises due to the slower dynamics of the population level. Results are representative of scenarios 9 to 10.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

However when there is cross-population communication from both populations, there is a feedback loop created between the populations and the circuit motifs, realized by the effect of the population density modulating the extra-cellular circuit response (seen and discussed earlier). This secondary action is however of a weaker strength than the more direct interaction realizable between the population and the circuit (later analysed with scenarios 16 to 25).

A sample of these results is presented below in fig. 5.19. We can see how the populations vary in the same manner as basal signal activation is increased, when the intracellular circuits are the same (i.e., covalent modification system - *left* plot; in contrast to *left* plot in fig. 5.18 where the circuit of only one population interacted with the population of the other). However this response is further modulated and the switch-like behavior is eroded (albeit slightly) as a consequence of the dilution behavior seen above. This can be observed if we contrast the behavior of the population strengths shown in fig. 5.18 (*left* plot) and fig. 5.19 (*left* plot). The former is an example of unidirectional communication (and thus no dilution effect feeding back) while the latter is an example

of a bidirectional communication with dilution effect modulating the response further.

This effect is also seen with populations of cells containing other complex non-linear signaling motifs. Figure 5.19 shows how the oscillations are further dampened (arising here due to both the slower dynamics of the population level and the shuttling of the species into and out of the cells reducing the concentrations of the extra-cellular motif response). Further, the switch-like behavior also shows small perceptible oscillations due to the feeble feedback from the population level. However the amplitude of these oscillations are too small in comparison to the overall response.

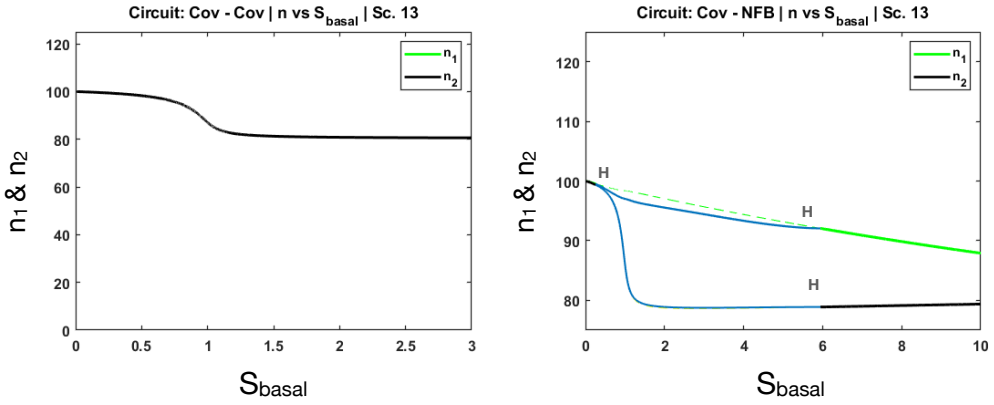



**Figure 5.19: Effect of cellular level interacting at the population level (cross-population & unidirectional from both populations).** Results from scenario 13, where the cellular level of both populations decrease the growth rate of the other population. The figure shows a modulated switch-like behavior (*left*, where both population densities are similar and are superimposed on the plot), and oscillatory response (*right*), arising from the covalent modification system and the NFB oscillator circuit motifs respectively. The bidirectional nature of the interactions and the role of feedback arising from the dilution effect of the extra-cellular circuit response in interactions (due to the species shuttling in and out of the cells) is discernable in the more modulated oscillatory response in the population densities (*right* plot), in comparison to the unidirectional communication between circuit to population across populations seen in fig. 5.18. Results are representative of scenarios 11 to 13.
[H - Hopf Bifurcation; Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

These results once again highlight several pitfalls in making inferences regarding system behavior based on observations of a few key variables. Several layers of communication and feedback are inherent to the complex collective phenomena seen at the level of population of cells, and any of these could be inherently responsible for lending, leveraging and modulating behavior of another level or the behavior seen across the whole system, while being seemingly passive.

*Population strength can modulate circuit response*

For our next analysis, we considered scenarios (14 to 15) where the population level of one cell type modulated the circuit level response from the other cell type. This provides a productive platform to build on the circuit modulating population response

study (seen earlier), to finally study the bidirectional cross-population population-circuit responses.

We observed simple (and intuitive) perturbation of the circuit response and the extra-cellular response from the circuit. This was also the case when there was both population to circuit interactions from cell type 1 to 2 and vice versa. However the one key difference between this and simple modulation is variation in the extra-cellular response, which is directly linked to the total population size (referred to as dilution effect thus far). Thus the extra-cellular response is modulated by two factors, 1. the population size affecting dilution of diffused species, and 2. through changes in the circuit strength increasing concentration of diffusing species (see fig. 5.20 and fig. 5.14 for contrast).

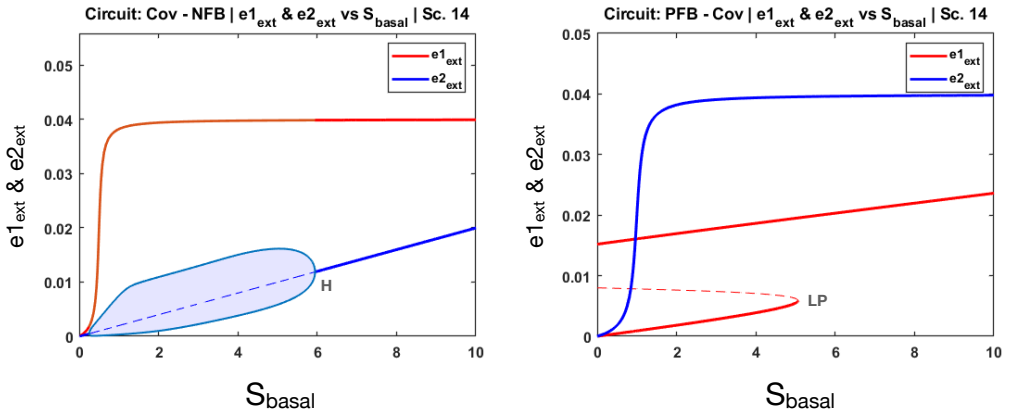



**Figure 5.20: Effect of population circuit communication on extracellular response.** We can observe here the impact of the population level interaction on the extra-cellular circuit response on two axis, 1. population size affecting dilution of species, and 2. through modulation of circuit motif signal strength. The extra-cellular response is strengthened vis-a-vis that obtained without any interaction - see fig. 5.14 for comparison (despite the population level strength remaining constant as circuit basal signal is varied). Results representative of scenario 14 and 15.
[H - Hopf Bifurcation; LP - Limit Point (saddle node bifurcation); Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

*Bidirectional communication between the population level and circuit level across populations*

Explicit bidirectional communication between the population level and the circuit motif, allows for further effective population level response and increasingly richer systems behavior. On one hand the circuit modulates the population and lends its information processing capability across levels and to the other cell type. On the other hand, the population interaction on the circuit buffers and attenuates the amplitude of responses.

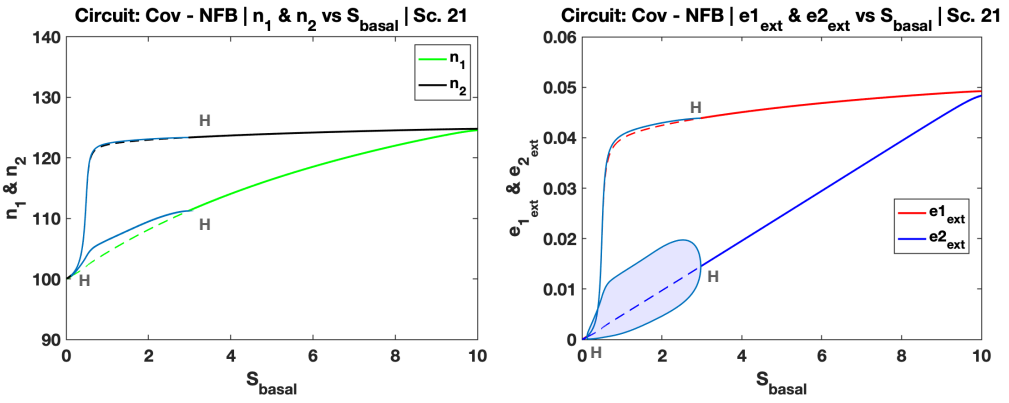



**Figure 5.21: Effect of bidirectional interactions between cellular level and population level.** Figure shows how population strengths ($n_1$ & $n_2$) and extra-cellular circuit responses ($e1_{ext}$&$e2_{ext}$) vary as basal signal activation ($S_{basal}$) is changed. The cellular motifs are covalent modification system and NFB, in populations 1 and 2 respectively. Results from scenario 21 where the cellular circuit and the population level variables of both populations interact bidirectionally. Figures show how on one hand the circuit (Cov and NFB) modulates the population level and lends its information processing capability (oscillations in $n_1$ and $n_2$ - *left*) while on the other hand, the population interacts with the circuit and dampens the amplitude of responses (particularly oscillations) through its slower dynamics. Results representative of 16 to 25.
[H - Hopf Bifurcation; Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

Together, this allows for more richer range of signal activation (and other parameters) over which the qualitative behavior of interest is seen. These results are the sum of the different effects that we have studied individually earlier, including dilution effects, dampened dynamic response etc (and behavior similar to those portrayed in figs. 5.18 to 5.20). An exemplar of the behavior of the system with such bidirectional interaction between the population and the circuit level is shown in fig. 5.21.

*All multi-level multi-population interactions active*

Arguably the richest scenarios in terms of number of interactions, scenarios 26 to 28, represent the most realistic model cases as well, covering the range of cell-population, cell-cell interactions possible. Simulations of these scenarios, reveal the complex nature of the emerging systems behavior, with traces of the pair wise interactions studied above embodied in them.

One of the more unique results seen in this scenario, is how for the same underlying parameters the system with NFB and PFB circuit motifs can have steady states of completely different nature co-exist close to each other (see fig. 5.22). This characteristic combination of circuit response from a PFB and NFB circuit was seen earlier as well (see fig. 5.16). However, the system here more readily presents this close cluster of steady states that may not be clearly discernable in observations.

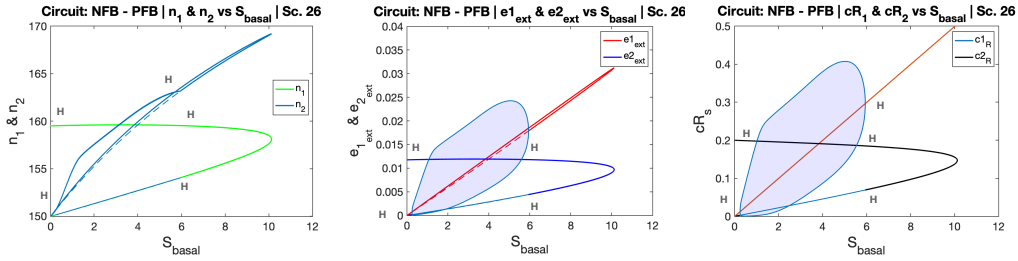



**Figure 5.22: Cross-level interactions allow for simultaneous 'co-existence' of different steady states at similar concentrations at the population level, which otherwise reflect drastically different responses at the cellular level.** Results from scenario 26 where all levels are interacting with each other in a bidirectional manner. *Left*: population strengths ($n_1$ & $n_2$) as basal circuit signal activation is changed. *Middle*: extracellular circuit response ($e1_{ext}$ & $e2_{ext}$) as basal circuit signal activation is changed. *Right*: circuit response ($cR_1$ & $cR_2$) as basal circuit signal activation is changed. The collective emergent behavior from cross-level cross-species interactions can allow for multistability and dynamic behavior such as oscillations. However even though these are systems phenomenon, the behavior may be more reflective in certain species and not in others. We can see one example of this where each level is communicating with the other in a unidirectional way (refer table 5.3), and the system is capable of generating both oscillations and multistability. However we seen above in the plot for $n_1$, the oscillations are more pronounced at the cellular response and not at the population level, where they are negligible and within the margin of error of experiments and observations might be seen as a steady state (despite the system and thus the populations technically oscillating and the steady state being unstable). This illustrates two complexities, 1. how multilevel interactions can allow for the system to share and leverage unique behavior from each level and present a co-existing landscape of behavior, 2. such behaviors however may be more pronounced in few levels, and insignificant or absent in the others.
[H - Hopf point bifurcation, Blue shaded regions represent bounds of oscillations; Dotted lines represent unstable steady states]

Figure 5.22 shows how the overall system has one steady state which is unstable (for ranges 0 to 6 of basal signal, $S_{basal}$), due to the oscillations. But even though the whole connected system is unstable and oscillates, the extra-cellular response concentration $e2_{ext}$ barely registers discernable oscillations (very lower amplitude), while $e1_{ext}$ oscillates with a pronounced amplitude. This directly reveals a pitfall in which measurement of certain variables of a connected system might not reveal the underlying qualitative behavior of the system. This nuance associating with understanding systems behavior through measurement of one variable, is further exacerbated when we consider the stable response further along signal ranges (6 and beyond), where the multiple stable steady states of $e1_{ext}$ are close to each other, while those of $e2_{ext}$ are further away. This is also reflected commensurately in the population level with $n_1$ showing minimal variation between the two steady states, and $n_2$ showing vast differences.

*Multi-level interactions across populations generates novel behavior*

Our simulations of the remaining scenarios (with all multi-level interaction) revealed yet another vista into the rich landscape of systems behavior at this cross section. Multi-level interactions were able to generate tristability in the overall system, when neither circuit level nor any single population as a whole was able to sustain or produce such

qualitative behaviors.

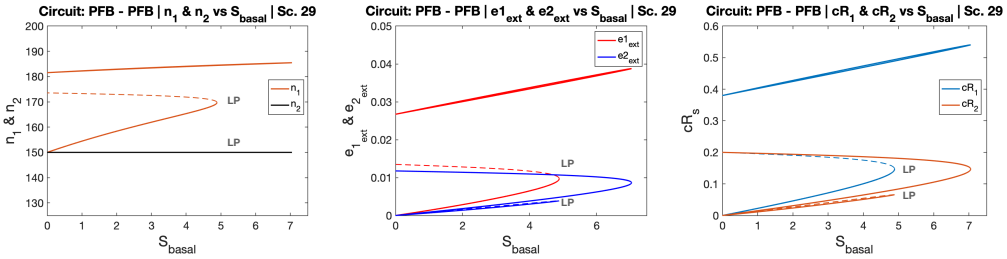



**Figure 5.23: Multi-level interactions generate novel qualitative behavior (tristability) at the systems level, which may not be reflected equally in all species and levels.** Results shown from scenario 29, where each level, apart from the populations, is communicating bidirectionally with the other. Figure shows how such interactions across and between population & circuit levels can enable the introduction of tristability across the system. However, as shown above the tristability is only clearly discernible in key variables despite the system interacting and communicating at every level. For example, $n_2$ does not show significant variation in the steady state behavior between different states. $n_2$ (*left*), in this instance is $\sim 150$ for all steady states while other variables show drastic variations including $n_1$, $cR_1$, $cR_2$, $e1_{ext}$ and $e2_{ext}$ show drastic differences (*left - right*). This is an example of collective behavior being more pronounced in only a few levels of the system, and could lead to incorrect inferences (based on observations of certain levels' or certain populations) of the behavior and its origins.
[LP - Limit point (saddle node bifurcation), Dotted lines represent unstable steady states.]]

In fact, in certain scenarios (see fig. 5.23), despite the system being connected at every level (across populations and across levels), the tristability was only clearly discernable or observable (despite being present) in certain variables. For example, in scenario 29, with all levels except the population levels having cross population bidirectional interaction, we noticed that $n_1$ does not show significant variation in the steady state behavior between different states, while $cR_1$ and $e1_{ext}$ show drastic differences.

## 5.5 Emergent multi-population behavior

In our study so far, we have considered the population level by only capturing its linear growth-death dynamics. The two population study so far considered how populations of cells may interact with other populations of cells in simple ways by affecting the (linear) growth and death rates of each other. This has been because the primary focus of our study has been in trying to understand how cells can enable collective emergent behavior at the population level, including interactions with and through the populations. That is, thus far the perspective has been from the cell, looking outward towards a collective population level response - with complex information processing at the cellular level and we focussed on how this can percolate through other levels by way of cross-level, multi-level interactions. However the population level and its dynamics are also capable of complex information processing and emergent behavior [54, 197]. It is helpful to point out here that, while the collective population level response referred to here is due to the action of the individual cells that make up them, they are funda-

mentally distinct from the intra-cellular signaling circuits. As mentioned earlier, this collective population level response often operates over different timescales and is enabled through distinct cellular processes (apoptosis, toxicity, competition etc) that we don't explicitly consider in our cellular level (being largely restricted to signaling dose responses). Population level interaction could also involve signaling regulation from other factors upstream or beyond our model that enable emergent response. In this sub-study we aim to understand how such complex population level behavior (emergent) can enable and perturb cellular level response.

**Model**

To understand this, we have considered a predator prey model for the population level dynamics, featuring a variant of the canonical Lotka-Volterra model (adapted from [198]). The Lotka–Volterra model is a modelling framework used to capture the dynamics of two species/populations that interact with one and other; where one population is the predator ($n_2$ in our setup), while the other population is the prey ($n_1$) [199]. Using simple assumptions the model can accurately capture the complex dynamics that arise from predator prey population systems. The overall model is described below as a system of ODEs. For this sub-study, we keep the population level dynamics insulated, i.e., We do not consider the intra-population interactions between a cell and its population level, since the focus of this study is to understand how non-linear and complex population level dynamics and interactions affect cellular level response.

The population level of both cell types (predator and prey, $n_2$ & $n_1$) have a predator prey model (adapted from canonical Lotka-Volterra description of such dynamics [198]). The cellular level consists of simple signaling motifs (described below). In line with the previous models, the cellular motifs are still considered modular. The signal activation of these circuits consists of a basal signal than can be amplified by, 1. the population strength of the other cell type, 2. the extra-cellular circuit response of the other cell type. Thus we only consider explicit cross-population activation of the cellular motif. The circuit response $cR$, can (as with the previous models), produce an intra-cellular circuit response $e_1$ & $e_2$, that can diffuse out of the cell as $e1_{ext}$ & $e2_{ext}$. The diffusion is mediated by a diffusion coefficient $D$ (Which is the same for both cellular response species). See assumptions for a detailed enumeration of the various assumptions. The choice of interactions considered within the model is bearing in mind the focus of this sub-study - to explore multi-level multi-population collective behavior; To this end we build on our understanding of intra-population behavior from the two earlier studies, and are primarily interested in this sub-study on understanding on how cross-level cross-population interactions can enable them.

**Cell type 1** $prey$

$$\begin{aligned}
\frac{dn_1}{dt} &= a_0 + \alpha n_1 - \beta n_1 n_2 \\
\frac{dc_1}{dt} &= f_1(S_1) \\
&\quad \text{where} \quad S_1 = S1_{basal}\left[1 + \text{Sw1}\frac{k_1 n_2}{K_1 + n_2} + \text{Sw2}\frac{k_2 e2_{ext}}{K_2 + e2_{ext}}\right] \\
\frac{de_1}{dt} &= k_{a_{e_1}} c - k_{d_{e_1}} e_1 + D\left[e1_{ext} - e_1\right] \\
\frac{de1_{ext}}{dt} &= k_{e1_{pr}} + D\left(\frac{n_1 V_1}{V_{env}}\right)\left[e_1 - e1_{ext}\right] - kd_{e1_{ext}} e1_{ext}
\end{aligned} \tag{5.3}$$

**Cell type 2** $predator$

$$\begin{aligned}
\frac{dn_2}{dt} &= b_0 + \delta n_2 n_1 - \gamma n_2 \\
\frac{dc_2}{dt} &= f_2(S_2) \\
&\quad \text{where} \quad S_2 = S2_{basal}\left[1 + \text{Sw3}\frac{k_3 n_1}{K_3 + n_1} + \text{Sw4}\frac{k_4 e1_{ext}}{K_4 + e1_{ext}}\right] \\
\frac{de_2}{dt} &= k_{a_{e_2}} c - k_{d_{e_2}} e_2 + D\left[e2_{ext} - e_2\right] \\
\frac{de2_{ext}}{dt} &= k_{e2_{pr}} + D\left(\frac{n_2 V_1}{V_{env}}\right)\left[e_2 - e2_{ext}\right] - kd_{e2_{ext}} e2_{ext}
\end{aligned}$$

The fundamental difference between the model described here and those utilized in the two population study (apart from the reduced number of interactions between the levels), is the the emergent behavior capable from the population level dynamics. The Lotka-Volterra predator prey model adapted here is capable of sustained stable oscillations in the concentrations of the prey and predator for suitable choices of the growth and death rates of the respective populations. We study the system in this regime to explore the behavior of the overall system with oscillating populations. However the model is also capable of stable steady states, and the relevance of this in the context our results and insights is discussed later.

### Interactions

*Prey to Predator*

| | | |
|---|---|---|
| **Sw1** | $Pop_2 \rightarrow Cir_1$ | Population strength of the predator strengthening preys' circuit motif response |
| **Sw2** | $Cir_2 \rightarrow Cir_1$ | Extra-cellular circuit response from the predator strengthening preys' circuit motif response |

*Predator to Prey*

| | | |
|---|---|---|
| **Sw3** | $Pop_1 \rightarrow Cir_2$ | Population strength of the prey strengthening predators' circuit motif response |
| **Sw4** | $Cir_1 \rightarrow Cir_2$ | Extra-cellular circuit response from the prey strengthening predators' circuit motif response |

### Assumptions

The various assumptions made in the above model are,

1. The **populations (1 & 2)** grow and decay following predator prey dynamics (Lotka-Volterra model, $n_1$ - prey, $n_2$ - predator).
2. **Lotka-Volterra model assumptions** apply for the population dynamics, including; the prey finds ample food at all times, predator population depends entirely on the prey population, the rate of change of the populations (prey and predator) is proportional to its size, and finally, the predators have limitless appetite.
3. The **population dynamics** does not explicitly depend on environment.
4. The **circuit** motifs are modular. The circuit responses are $c_1$ & $c_2$ and lead to the production of their respective intra-cellular species ($e_1$ & $e_2$). This is decoupled to allow for diffusion and sequestration in extra and intra-cellular environment without affecting the modularity of the intra-cellular circuit.
5. The **intra-cellular species** ($e_1$ & $e_2$) can diffuse out of the cell and the dilution is accounted for by taking the volumes (the cell and the extra-cellular environment) in to account. The intra-cellular volume is assumed to be insignificant in comparison to the extra-cellular value and is thus fixed, see parameters for more details.
6. The **extra-cellular species** ($e_1$ & $e_2$) have linear rate degradation to account for any environmental degradation.
7. **Both cell types have the same volume** ($V_1$). $V_{env}$ is the volume of the enviornment.

**Comment on parameters:** In our other studies, conscious of the fact that the dynamics of the different levels (population and cellular level) are over differing timescales, we chose representative parameter values within our models, where the population level parameters were 2 orders of magnitude lower than the cellular counterparts. However, for our current sub-study, since the primary focus of interest is to understand how complex emergent behavior at the population level (between populations) can percolate down to the cellular level, and affect the collective behavior of cells and cellular responses, we relax our earlier assumption to couple these levels more strongly within our study. We also wish to highlight here that, the population level dynamics can also often be of similar timescales as the cellular level.

### Circuit motif considerations

Similar to the one population study and the two population study, we consider different circuit motifs to understand the role of population level behavior and dynamics in affecting the collective emergent behavior. We do so by considering the following circuit motif combinations. The model descriptions for these circuit motifs are presented in appendix B.2 (and the responses from these modular cellular circuits are shown in fig. 5.3). As an additional exemplar case study to explain specific insights, we also consider a mutual inhibition circuit motif that is capable of reversible bistability. The model for this is also presented in appendix B.2.

| Population 1 | Population 2 |
|---|---|
| $n_1$ — *prey* | $n_2$ — *predator* |
| Cov | Cov |
| PFB | NFB |
| NFB | PFB |

**Table 5.4: Circuit choice for two population study with complex emergent population level dynamics**

Unlike the earlier studies, the two populations are fundamentally distinct in their dynamics and behavior; i.e., the effect and behavior of a system with a PFB motif within a predator is fundamentally different from a PFB motif within a prey (even if the circuit motif does not affect its own population strength). Thus it is vital to understand the difference between insights from when one motif is part of a predator and when it is part of a prey. To explore this further we consider a PFB and a NFB circuit motif for the predator and prey respectively (and vice versa). A simpler study focussing on the behavior of this system, where both populations contain a covalent modification cycle as a intra-cellular circuit is considered to provide a benchmark scenario, where all the non-linear complexity is arising from purely the population level dynamics.

### Scenarios

In this section the focus is on trying to understand the role of emergent population level dynamics on collective behavior of cells and populations containing different cellular circuitry.

| Row No. | Interactions | Description | Reference |
|---|---|---|---|
| 1 | - | Studying only population level dynamics | 5.24 |
| | | Population $\rightarrow$ circuit | |
| 2 | Sw1, Sw3 | Studying downstream effects of population level dynamics on cellular response | 5.25 and 5.26 |
| | | Circuit $\rightarrow$ circuit | |
| 3 | Sw2 | Studying role of circuit interactions and upstream | 5.27 |
| 4 | Sw4 | population level dynamics | |
| | | All interactions together | |
| 5 | Sw1, Sw2, Sw3, Sw4 | - | 5.27 and 5.28 |

**Table 5.5: Circuit choice for two population study with complex population level dynamics**

We do this by simulating a number of scenarios with primarily cross-population interactions. These scenarios are constructed by toggling the switches describing the in-

teractions on and off (taking values 0 or 1 respectively) in the master model described above in eq. (5.3), and are listed in the table table 5.5. Unless explicitly mentioned in the scenario the interaction is absent in the model (i.e, (Sw) = 0)

**Results and discussion**

The defining feature of the Lotka Volterra predator prey model (adopted here) is the oscillatory behavior of the predator and prey population strengths. The variant of the Lotka-Volterra model considered above is also capable of non-emergent stable steady state responses (for suitable non-zero choices of parameters for $a_0$ and $b_0$, see [198] for more details). However the focus of our study here, predominantly, is in understanding how emergent features of the population level can percolate through the system and affect other levels and the observed collective behavior. Hence we consider the population level dynamics to be oscillatory and thus choose suitable parameter values, i.e., $a_0 = 0$ & $b_0 = 0$. We also do not consider any interactions or influence from the other levels to the population level (beyond the predator prey population level dynamics). This is since the focus of the sub-study is to understand cross-population cross-level interaction particularly in the presence of emergent behavior at the population level. However, population level dynamics as explained in the model description can and does affect the levels further downstream. We simulate the system (see appendix B.2 for parameters) according to the scenarios described in table 5.5, and the results are presented below.

**Population dynamics affects observed intra & and extra-cellular circuit response.**

We initially simulate the scenario 1, where there are no interactions between the levels. Similar to the simple dilution effect observed in the single population study and the two population study earlier (due to the intra & extra-cellular response shuttling in and out of the cell), here again, the changing population level is able to modulate the intra & extra-cellular circuit response. However unlike the earlier examples, where we reported this behavior at steady state concentrations of the populations, they can also be observed in the dynamic response (see fig. 5.24). Figure 5.24 (row 1) shows how the intra and extra-cellular circuit response can oscillate despite the circuit motif response being steady due to the population level dynamics. Here the circuit motifs contained within the populations (prey and predator) are simple covalent modification systems. With no interactions affecting its activation, the motif provides a steady constant response (row 1, column 2). However, due to the changing population strengths the extra-cellular and intra-cellular concentration observed oscillates with the population strengths. The fact that such behavior is seen both in the concentration within the cell and that of the species outside of the cell in the dynamic response implies that dynamic changes picked up by experiments, including single cell experiments, might be reflective of behavior borrowed from another level rather than inherent to the internal circuitry of the cell. This dilution effect can further complicate the landscape

of behaviors seen when oscillations from within the cellular circuit get modulated by the population level oscillations (with different phases and time periods). This complicated behavior, a superposition of the two oscillations arising from the motif and the population level is shown in row 2 and 3 of fig. 5.24.

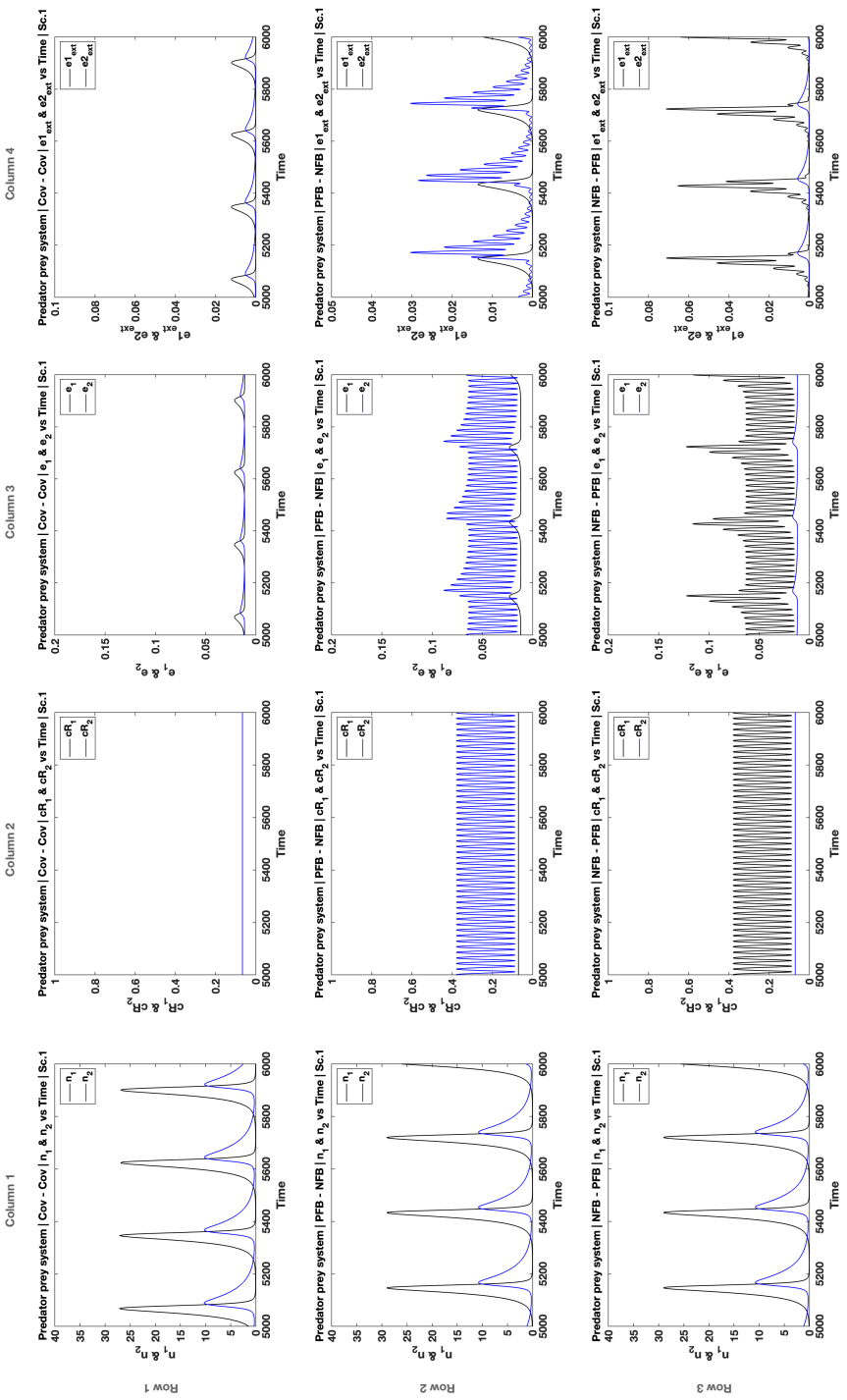



**Figure 5.24: Population dynamics modulating intra & extra-cellular circuit response through dilution.** Figure shows simulations from scenario 1, where there is no explicit interactions between any of the levels. *Column 1*: shows population level dynamics, with oscillating population strengths of predator population ($n_2$) and prey population ($n_1$). *Column 2*: shows the modular circuit motif response for each circuit motif. *Column 3*: shows the concentration of the intra-cellular circuit response. *Column 4*: shows the concentration of the extra-cellular circuit response. We can see how despite the levels not communicating with one and other, the intra & and extra-cellular responses trail the respective population level strengths (for the different populations and motifs) *(column 2 & 3)*. This is due to the dilution effect that has been discussed in detail within the two population study. However, here in the context of dynamic changes to the population, this can allow for a simple steady circuit response (such as in the covalent modification system motif t the circuit level) to also oscillate (with different characteristics to the population level oscillations) - see *row 1, left to right*. These oscillations can further modulate the oscillations arising from the circuit motif response leading to superimposition of oscillations arising from different levels - see *row 2 and 3*.

Another aspect of complexity is that the dynamics of the oscillations is different for the predator and the prey systems. The prey has a gradual rise in its population level strength before starkly decaying (due to a abundance of predators). This behavior is juxtaposed with the dynamics of the predator which rises more sharply and falls gradually. This difference in dynamics is seen in the way the oscillatory behavior of the NFB motif response is superimposed on the population dynamics of the prey and the predator (row 2 and 3).

Put together, a simple population level dynamic (such as oscillations) can fundamentally shape the way the responses from within the cell are picked up or observed both within the cell and outside of it, while retaining characteristic features of the population type to which they belong.

**Interactions between the population level and cellular level: Upstream regulation.**

The next scenario we considered was one involving the population levels regulating the circuit level response through signal activation (scenario 2). This offers a basic exploration of how the predator prey dynamics can directly affect the cellular circuit. The results from this analysis are presented in fig. 5.25.

In this analysis, we can see how the dynamics of the population level is more apparent in the intra and extra-cellular response (in comparison to the scenario where there were no interactions). This is due to the fact that unlike in the case with no interactions where the resulting oscillations and perturbations were due to the relatively weaker channel of dilution effects, here it is due to active modulation of the signal to the circuit. This increased register of the oscillatory dynamics is apparent in all three circuit motif combinations (row 1 to 3 of fig. 5.25). We again see superimposed oscillations arising from the interactions with the NFB. However, interestingly we can note how the modulation from the population has switched the PFB response to a higher steady state (row 2 & 3 of fig. 5.25). This is apparent from the fact that the system (which originally had basal activation that sustained a lower steady state - see fig. 5.3), has moved to response ranges reflective of the higher steady state.

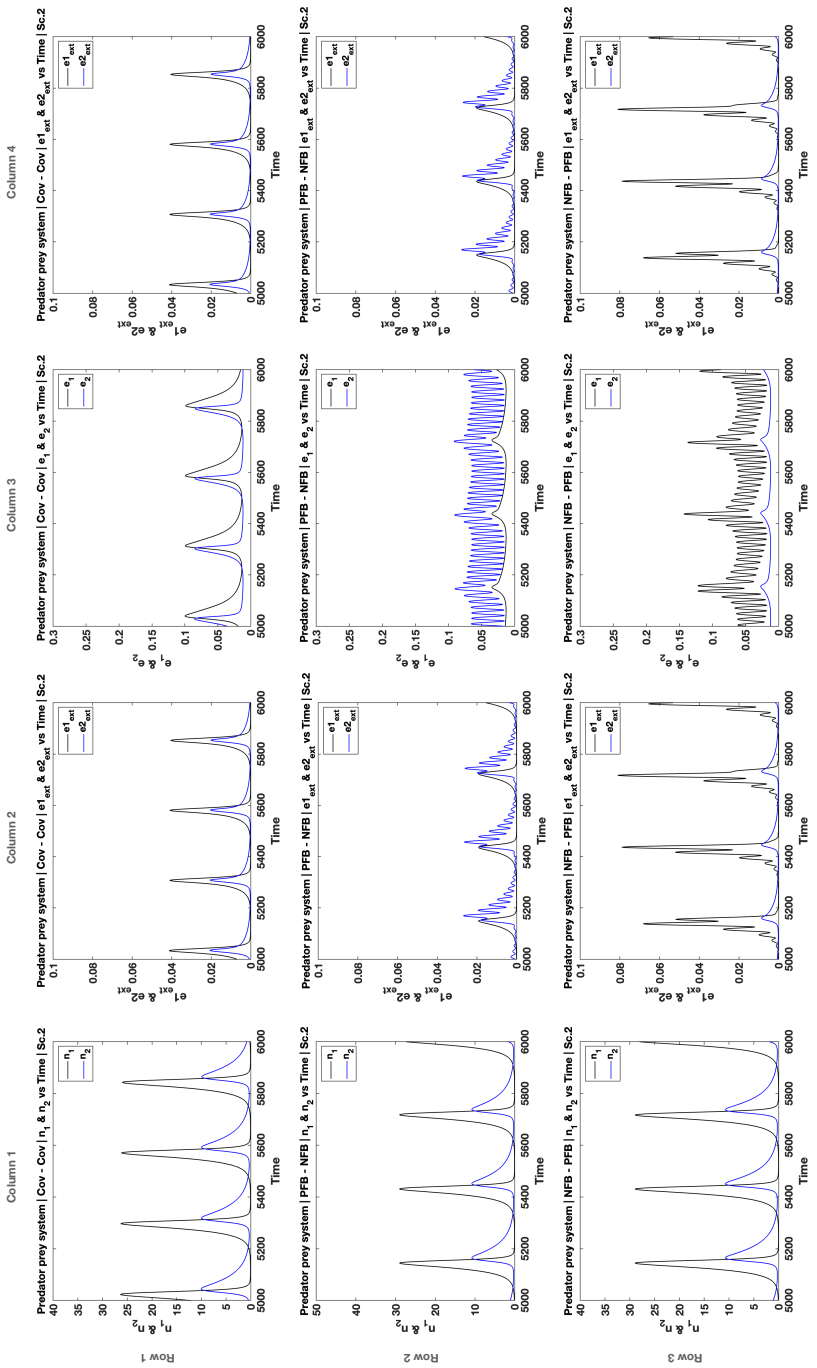



**Figure 5.25: Population dynamics directly modulates circuit activation (cross-population).** Figure shows simulations from scenario 2, where there is direct interaction in the form of signal activation between the population level of predator and prey on the circuit level of the prey and predator populations respectively. *Column 1*: shows population level dynamics, with oscillating population strengths of predator population ($n_2$) and prey population ($n_1$). *Column 2*: shows the modular circuit motif response for each circuit motif. *Column 3*: shows the concentration of the intra-cellular circuit response. *Column 4*: shows the concentration of the extra-cellular circuit response. We can see how direct signal activation is capable of registering stronger oscillatory behavior at the circuit level (in comparison to no interactions as shown in fig. 5.24) - see *column 2*. We can also observe stronger superimposed oscillations in the intra & extra-cellular circuit responses. With PFB motifs (*row 2& 3*), we can see how the interaction is able to switch the circuit to the higher steady state response (in comparison to the no interaction scenario shown in fig. 5.24, where the system remained at the lower steady state consistently).

Note that the motif we used in the analysis was a irreversible bistable switch (PFB - mutual activation motif). To illustrate how this switching can be achieved dynamically with each oscillation, in a reversible manner, we simulated a similar population level setup with a reversible bistable switch (mutual inhibition motif). The model for the motif is adapted from Tyson et al., [52] and is provided in appendix B.2. The results from this simulation are shown in fig. 5.26. This illustrates how the dynamics of the population level (over larger time frames can affect stability of the internal circuit temporally and at steady state).

Figure 5.26, column 1 shows the expected input output response from the MI motif (reversible bistability in the output response). In our analysis, this motif was considered as the cellular motif of both the predator and prey populations. We then considered a scenario 2 of interactions where the population levels where actively regulating the signal activation of the circuit motif in the other population. The figure shows clearly how, the population of the prey ($n_1$) and predator ($n_2$) oscillate with time aligned with how the respective MI circuit motif responses change. It is worth recapping that the predator population affects the preys circuit and vice versa. This is quite apparent in the phases of the oscillations between the population level and the circuit response. But row 2 of the plot very clearly is able to show how the motif responses are both in the higher stable branch momentarily as the population strengths regulating them move beyond the threshold bringing the signal to higher ranges without bistability. As the population then wanes (along with the oscillations), we can see the circuit motifs holding the higher steady state before falling abruptly to the lower steady state response. This hysteresis fall is very clearly seen in the response of the preys' circuit motif $cR_1$ (row 2, column 2). A similar response is indeed present in the predators response but however is difficult to discern since the preys drops drastically, modulating the signal sharply within it.

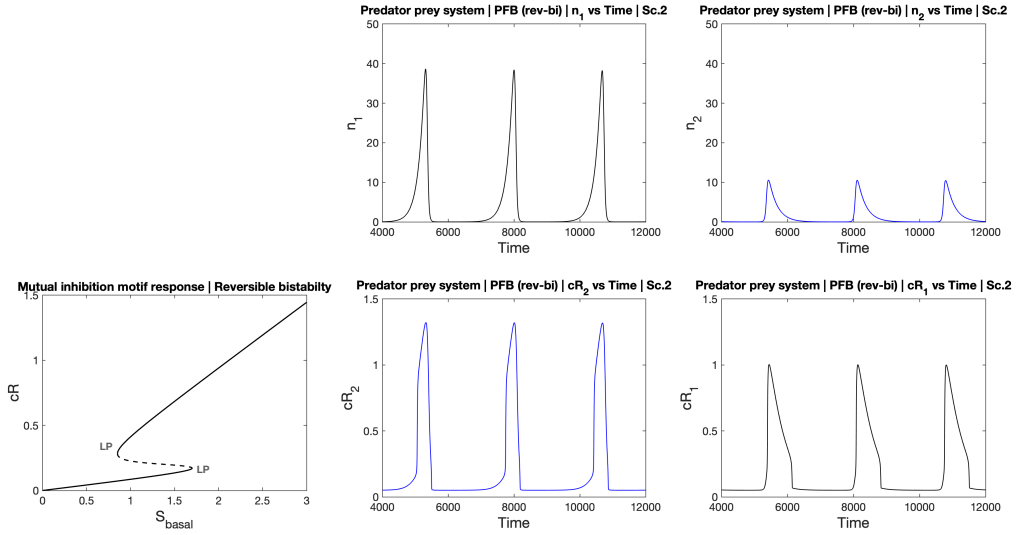



**Figure 5.26: Population dynamics can enable reversible switching in circuit motif response (MI motif).** Plot shows population dynamics (*column 1*) and the transient behavior of the circuit level responses of both the predator (*column 3*) and prey populations (*column 2*). Both populations have MI motif at the intra-cellular level. Figure shows how population level regulation of a MI signaling motif (capable of reversible switching) can reversibly switching the circuit motif response between the different steady state branches as the population strength oscillates. The scenario simulated is scenario 2, where the populations are actively regulating the signal activation of the circuit motifs in their counterparts' cells. i.e., the predators population strength regulating the signal activation of the prey and vice versa. *Column 1* shows the expected signal output response from the MI motif. *Row 1* presents the population level dynamics of the predator prey system. *Row 2* of the plot very clearly is able to show how the motif responses are both in the higher stable branch momentarily as the population strengths regulating them move beyond the threshold bringing the signal to higher ranges without bistability. As the population then wanes (along with the oscillations), we can see the circuit motifs holding the higher steady state before falling abruptly to the lower steady state response. This exemplar study consists of a reversible MI signaling motif (mutual inhibition - see appendix B.2

This insight however highlights how higher level dynamics and emergent behavior can sometimes register strongly in the circuit level of the populations' cells. From fundamentally modulating the cells simple response, to regulating oscillations and enabling reversible switching, the population level is capable of much more than simple dilution of the secreted species.

**Cross-population interactions between and among the population level and cellular level.**

The scenario with the cellular circuits interacting with one and other (scenario 3-5) offers more complexity of the flavour seen above. A composite plot showing the manifestation of these characteristic interaction effects on the different levels is shown for scenario 5 (all cross-population interactions active - see table 5.4) in fig. 5.27.

The possibility of cross-circuit interactions (across-populations) enables an additional layer of complexity in the form of superimposition of oscillations. In addition to the

direct effect of the population level oscillations regulating the circuit level (for either populations' circuit), the indirect effect through the corresponding population offers a albeit feebler delayed oscillatory effect. This works to partially weaken the characteristics of the oscillations emerging from the circuit level (altered amplitude and phase etc), due to it belonging to a prey or a predator and being regulated by either, as the circuit is now regulated by the system, i.e., both and itself (as the circuit can also affect the other circuit). This presents a truly systems phenomenon, where the emergent response is a combination of different factors and effects including,

- Direct regulation from the population level through signal activation
- Dilution effect from the population level affecting intra & and extra-cellular responses
- Indirect regulation of the circuit level by the population through the regulation of extra-cellular responses
- Regulation of the circuit level by each other directly due to signal activation
- Additional complexity arising due to emergent behavior at the circuit levels

Figure 5.27 shows how each of these factors come together to enable a truly systems landscape of behaviors with even populations containing simple the simple covalent modification cycles showing superimposition of oscillations in their intra and & extra cellular responses (row 1, column 3 & 4).

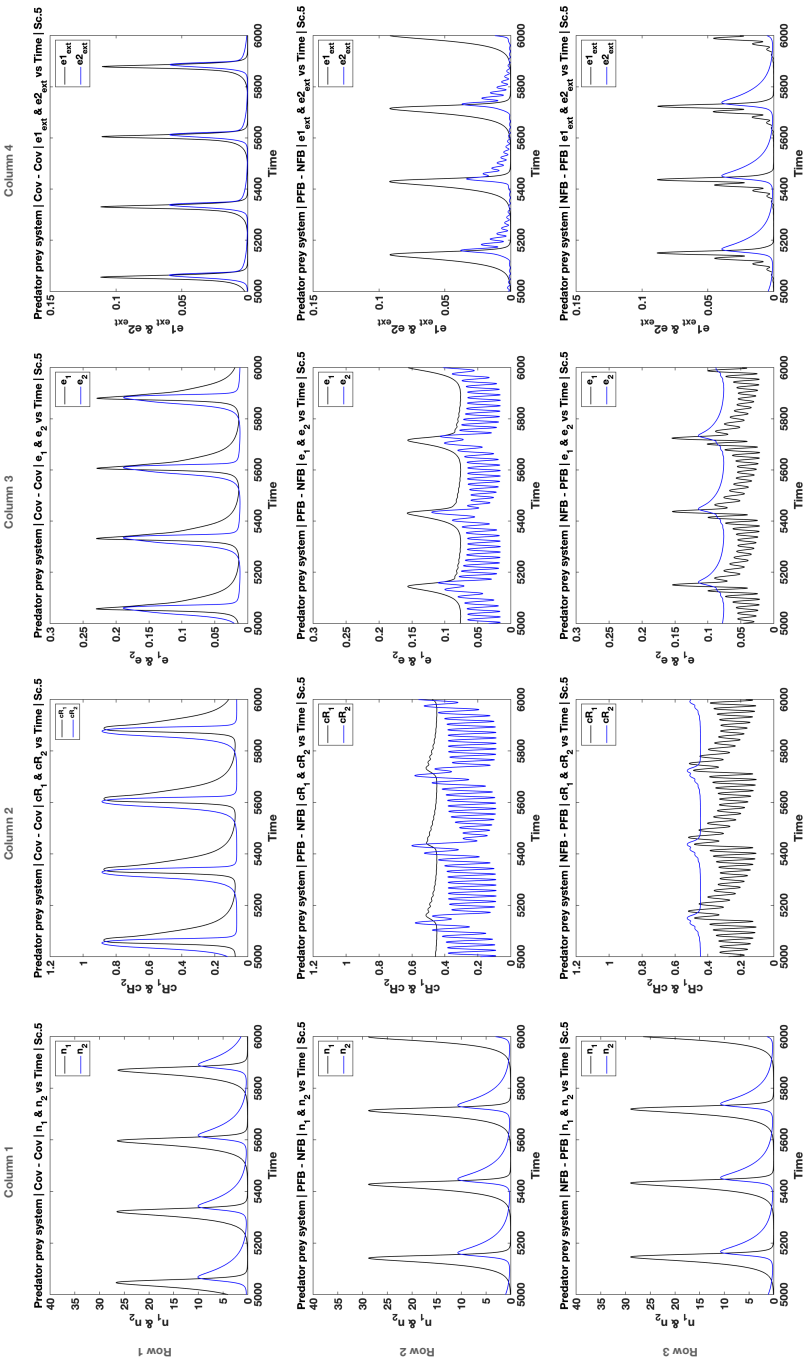



**Figure 5.27: Population dynamics and cross species circuit interactions in predator prey model.** Figure shows simulations from scenario 5, where there are multiple interactions 1. signal activation by the population level of predator and prey of the circuit level of the prey and predator populations respectively, 2. circuit levels activating each other (cross-species) *Column 1*: shows population level dynamics, with oscillating population strengths of predator population (1) and prey population (2). *Column 2*: shows the modular circuit motif response for each circuit motif. *Column 3*: shows the concentration of the intra-cellular circuit response. *Column 4*: shows the concentration of the extra-cellular circuit response. The figure shows how the existence of interactions at the circuit level (across different populations) can enable further channels of interactions through which the population dynamics can affect the collective behavior seen. We see how even in the simplest case with only covalent modification systems and a steady expected response at the circuit level and beyond, we can get superimposed oscillations at the circuit level, and at the level of the intra and extra-cellular response (*row 1*). The systems with more complicated circuit motifs present this more strongly (*row 2 & 3*), across all species.

These higher level modulations of the cellular circuit and its discernable responses both at the intra-cellular and at the extra-cellular level are highly relevant to the inference of underlying networks and systems based on observations. Not only can population dynamics (which can be much slower and indiscernible over the time scale of an experiment) affect the steady state of the system, it can also modulate the system dynamically. A pressing example of this is shown in fig. 5.28. This is an example of a predator prey system where the NFB circuit is being actively modulated by the population strength of the other species (scenario 5). The circuit in this system is operating with a signal activation that is in the range $\sim 10$ to $\sim 20$. Under this range, the cellular circuit is stable (see fig. 5.3), however due to the dynamic and transient nature of signal regulation (arising due to population dynamics), the system still registers damped oscillations throughout leading to a behavior alike to superimposed oscillations, even though the internal circuit operates in a regime of steady state. It is striking to note that the internal circuit is also being regulated through signal activation in the range where the system is stable. Thus purely transient oscillatory signal regulation from the population strengths (or in other instances other circuits) can enable a cellular motif to behave as if it is emergent and present with damped oscillations that together with other emergent behavior might be inferred as different.

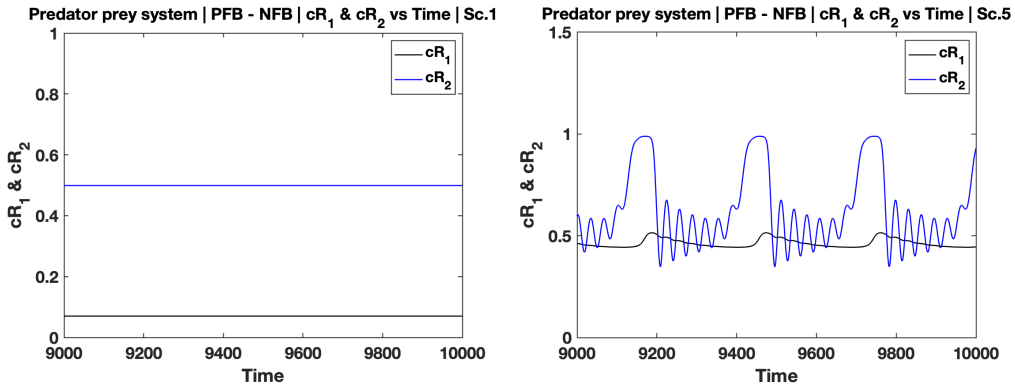



**Figure 5.28: Population dynamics can seemingly introduce inaccessible (for given signal activation) behaviors to the circuit motif.** *Left:* shows simulation (scenario 1) of a NFB motif response within the predator population where there are no interactions between levels. *Right:* shows simulation (scenario 5) of a NFB motif response within the predator population where the circuits signal is being modulated by the population strength of the prey. In both cases, the basal signal activation of the system is outside the range over which oscillations are observed for the NFB motif, however we can see that the dynamic population level regulation is capable of inducing instability in the system (leading to damped oscillations and eventually a fixed steady state, see *left*) that present together with the overall population level oscillatory dynamics as superimposed oscillations (see *right*). See text for more discussion.

The picture can get further complicated inferences when we consider time averaged sampling (either at the intra-cellular level or at the extra-cellular level). These insights together reveal the number of ways in which complexity at the population level, even with differing timescales of activity can percolate to the overall system and affect fun-

damental behavior of circuit motifs, and their intra & extra-cellular responses. In our study so far, we have only considered scenarios where the population level affects the cellular response and the cellular response interacts with each other. There is no explicit consideration of cellular level feedback back to the population level. This has been primarily because the focus in this sub-study has been to understand the role of emergent non-linear behavior at the population level and how it affects the collective behavior of cells. However when there is such feedback as we have considered in the two population study (see table 5.3), the cellular level and the collective behavior of them can perturb the population level response further - modulating the transition between emergent and non-emergent regimes of steady states at the population level. This can interact with the existing landscape of behavior emerging from the population level to circuit level interactions (modulation of oscillations and switching behavior for example), and can lead to a rich landscape of behaviors at the systems level. Such an analysis would be highly fruitful and is considered in the future work.

## 5.6 Conclusions and future work

In this chapter we studied cell-cell communication and other interactions amoung and between, the population level, environment factors and cellular circuits, and how they come together to enable a collective population level response.

We began by studying a single population; where we considered a population of cells (of a given type) containing a cellular circuit motif (covalent modification system, PFB, NFB, relaxation oscillator, and homeostatic motif). We considered the variety of interactions that can exist between these different levels. Using a master model incorporating these interactions and communication channels, and an automated computational pipeline, we generated the behavior of these systems under different combinations of interactions and cellular motifs. We then studied these systems to obtain insights on how the population level response is dependent on the different levels and the interactions between them. Building on these insights, we proceeded with a similar study, this time focussing on two populations and the cross population interactions. We explored the role of cross-population cross-level interactions (circuit-population) in enabling the collective emergent behavior of cells (both at the intra-cellular level and the extra-cellular level). Finally we explored complex population level dynamics by studying such interactions in a focussed manner to understand how complex population level dynamics and other emergent phenomenon can percolate to other levels and affect the collective behavior of the system.

That together we presented a structured framework grounded in these studies to understand and unravel the complex population (and multi-population) level behavior of cells, looking both from the perspective of cells coming together to enable behavior, population level dynamics affecting cellular responses and circuits, and the environment, which is both an active modulator of this collective response and can get

perturbed by it. This study revealed a number of key insights which are summarised below.

**Borrowed behavior across levels.**

The fundamental feature of cross-level cross-population interactions is the fact that they allow for information processing capabilities to be transferred/borrowed across levels. We repeatedly saw in our simulations of the single population and two population study, The circuit motif from these cells for example, is able to lend qualitative information processing capabilities such as oscillations, multistability, homeostasis, and switch like behavior to the population level (see figs. 5.5 and 5.7,fig. 5.17). Depending on the nature of these interactions, these behaviors can register strongly carrying characteristics of the original level to which the behavior belongs; for example, in the single population study where the circuit directly modulates the response at the population level (fig. 5.5); or the behavior can register with characteristics of an other level (through which the interaction was mediated); for example, in the single population study where the circuit modulates the environment through the population. In this instance the strength of the borrowed behaviors are weaker, and the population in turn modulates the information processing capability by mildly suppressing the amplitude of the oscillations (fig. 5.11). An even more stark example of this is population strength modulating the observed intra & extra-cellular response from the circuit motif through dilution. From simple modulation at steady state to driving oscillations into steady state circuit responses (see fig. 5.24), the borrowed behavior can be both dynamic and at steady state.

This implies that a qualitative behavior observed at one level could in fact trace its origin to another level, with many factors, including the level mediating this interaction and the nature of the interaction, that can affect the nature of the collective behavior seen from these systems.

**Introduction of systems behavior and in some and all variables within a system.**

Cross-level interactions can also lead to novel behavior at the systems level, unseen (and sometimes incapable of) in each level/population by itself. For example, we showed the introduction of tristability from two populations with PFB intra-cellular circuitry interacting with each other (see fig. 5.23). Such behavior is exemplified when each level is capable of highly non-linear information processing or emergent behavior. When the population level is capable of emergent behavior (oscillations) we observed how the resulting system through various interactions can present with a superimposition of multiple oscillations in various species (see fig. 5.28). The relevance of such borrowed and novel behavior is even more pronounced and interesting given that not all species in the system will manifest the behavior. As we saw in fig. 5.23, the population size of species 2 can present with a completely different behavior (monostable), while the population and intra & extra cellular species concentration show a completely different response (bi and tri-stability). These interactions can also lead to non-trivial

combinations of novel behavior including multiple regions of oscillations and bistability emerging in the overall system, from the interactions between the circuit level of cells belonging to two distinct populations (see fig. 5.23).

**Modularity and downstream regulation.**

The implications of these interactions and the emergent landscape of behavior from cells and populations is further underscored by the fact that these cells often regulate further processes downstream. Even the simplest scenario considered in our one population study with the cell circuitry unidirectional-ly regulating environmental factors can lead to further non-trivial regulation of other processes (through the environmental level which in our model is passive). Thus beyond the immediate model concerned it is worth bearing in mind that these systems are further connected with the larger cellular circuitry and each (if not all species) is involved in regulating other key processes downstream. Hence behavior borrowed across levels or novel behavior introduced within the system from this collective behavior can transcend the immediate system and register and transfer such behavior to wider cellular processes and pathways.

**Relevance to inference and data.**

From an inference perspective, these results reveal how both observations of specific variables cannot allow for extrapolation to the behavior seen in the other variables or even the system. For both observations made dynamically (instantaneously) and time averaged, our insights and results indicate how inferences from them can be quite misleading with regard to the underlying behavior. For example, akin to the behavior referenced in fig. 5.23, in fig. 5.22 we showed how when there are cross population-cell interactions, the population level response can be manifestly different and from the cellular level response (even with dynamic behavior such as oscillations). Thus experiments that treat extra-cellular circuit response or other variables as a proxy to determine response or overall system behavior might make misleading insights regarding the nature of the underlying system. Our study reveals the multitudes of ways in which the different levels can not only perturb responses seen at other level, but can also fundamentally mask qualitative behavior from certain levels.

Future work will focus on sensitives in these scenarios to other kinetic parameters and signal strengths. In particular the relative strengths of interactions which has so far not been considered or explored will be given attention. Such a study will allow us to understand how the different channels of interactions differ from one and other (beyond the behavioral characteristics lent), from a quantitative perspective. We aim to build on the insights obtained in this chapter with these focussed explorations. The computational pipeline generated to automate the bifurcation analysis will be integral to exploring the complexity of the scenarios, interactions and parametric sensitivities. Finally we wish to extend these insights to concrete biological contexts, with a key focus on systems and populations that feature in synthetic biologys' efforts to leverage population level interactions for engineering systems behavior.

# Chapter 6

# Conclusions

Cellular information processing is a multi-level systems phenomenon, that actively involves the cell communicating with other cells (of the same cell type or of different types), and reacting to the environment. Internally, the cell fulfils this role by leveraging a vast chemical reaction network encompassing numerous proteins and gene regulatory networks [2]. Some of these biochemical agents are self-contained within the cell, and some diffuse in and out, allowing it to further modulate, adapt and react to (as well as affect) environmental factors. At the population level, beyond diffusible agents, a cell can communicate with other cells in a number of ways including through contact, close neighbours, ligands, long range diffusible agents, and through the environment [200, 186, 201]. For example, populations of cells can interact with each other more directly, through competition for resources, inhibition of growth rates, and can even have toxic effects on each other [197]. Put together, the cell both internally and externally leverages agents across multiple levels (each operating with its own complexity) to generate a complex information processing landscape at the population level. While there has always been focus on a given level, the multi-level nature of these interactions (both internally and across cells) has not been given much attention. However, with recent advents in single-cell experimentation (and analytical tools), there is a renewed interest in understanding the role of multi-levelness in cellular signaling. This is aided in part by efforts to engineer behavior across levels in synthetic biology [202, 173]. In this thesis, we tackled this complexity in a targeted manner by focussing on four distinct strands of research.

### Tackling the systems complexity within the cell

In chapters 2 and 3, we focussed on the mechanism of protein modification and its intrinsic behavioral capabilities. Given that modifications are an integral cog in the cells' internal machinery to react and adapt to changes, we explored the behavior and capabilities of these networks from two axes,

**Leveraging symmetry and symmetry breaking to explore capability and features of multisite modification networks**

In chapter 2 (Symmetry and symmetry breaking in multisite modification), we studied the role of symmetry and symmetry breaking in a canonical cellular mechanism and pathway, multisite protein modification. We focussed on a suite of double site modification networks, and using the vista of symmetry in the chemical reaction kinetics, we showed how protein modification networks with specific features can break symmetry (or certain symmetries) under specific conditions. DSP networks were capable of three classes of symmetries. These symmetries have multiple biological consequences. For example, case 2 symmetry (symmetry in the modification steps of both modification sites) is seen when there is no bias in the ordering and directionality of modifications, and case 3 symmetry (symmetry between the rate of first & second modification and first & second demodification at the same site) is hypothesized to play a key role in maintaining robust circadian rhythms [102, 103].

For these symmetries in canonical double site modification models, using a combination of numerical simulations, bifurcation analysis and analytical work, we provide answers to

- Whether symmetric kinetics implies a symmetric output, i.e. whether certain symmetries break?
- If such symmetries do break, what features of the system and what parameter regimes enable them?
- Conversely, if it doesn't break, what features of the system parameter regimes preclude symmetry breaking?
- What novel features does symmetry breaking engender to the symmetry broken state?

Our study considered a suite of double site modification systems, and included all combinations of ordered (random and ordered), mechanism (distributive and mixed) and enzyme action (same or different enzymes). By studying this landscape of models, and analyzing which systems break symmetry and which don't, we discerned features (structural) of the modification systems which enabled symmetry breaking. For example, we showed that commonality of enzymes and random distributive modifications promote symmetry breaking. In all cases through our analytical work we unambiguously ascertained which symmetry can/does break, and determined conditions (necessary and sufficient), based on the kinetic parameters, which guarantee or preclude behavior.

Symmetry breaking can lend features and behaviors to the symmetry broken state, and to these underlying systems. Beyond biasing or ordering of modifications, this can include complex information processing behavior such as multi-stability in the system. Our study shows how from starting with a system with no bias to the sequence of

modifications, symmetry breaking is potentially one possible way through which such networks can evolve to gain ordering and directionality (case 2 symmetry breaking). Symmetry breaking was also shown to present with Absolute Concentration Robustness - a sophisticated form of mathematically exact homeostasis, where concentration of variables are mathematically exact and unchanged even as parameters (as in this case, total amount of substrate) are changed. ACR has been seen earlier with bifunctional enzymes [113]), but to our knowledge this is the first time it's been reported in multisite modification systems. ACR through symmetry breaking was seen in the concentrations of specific substrate forms, and in sums of concentrations of pairs of substrate forms (see figs. 2.2 and 2.3). Our analytical work characterised these behaviors and we obtained analytical expressions which determine the homeostatic concentrations based on kinetic parameters and total amount of enzymes and substrates. In addition to ACR, symmetry breaking also lead to robust oscillations in the system. Case 3 symmetry breaking leading to oscillations has been observed in the context of robustness oscillations in circadian pathways [99]. Our analysis complements findings there and reveals how kinetics emerging from symmetry considerations can enable robust oscillations in canonical pathways.

The importance of multisite modification (and modification systems) stems from the fact that they confer functionality to proteins and in turn regulate key processes within the cell. Thus the consequences of symmetry breaking in multisite modification can have vast implications for networks and pathways regulated by and containing them. The sequestration of a substrate species in other pathways, cross-talk, and coupling (through enzyme and substrate sharing) can lead to a myriad of interesting behaviors and ways in which symmetry breaking can transcend the modular modification system and reflect across pathways and the cell. Through computation, we showed how all these insights are also relevant to systems that are near symmetric. Our analysis in simplified systems (TSP) showed how even in networks that are not symmetric, we can obtain approximate ACR, and other features of symmetry breaking. Thus, symmetry breaking can truly be a vista to explore the complexity of higher order systems, and behavioral features such as ACR. We explored these behavioral characteristics and further analyzed the biological relevance for the study of networks containing multisite modification. A summary of these different results is provided in section 2.5.

All in all, this study revealed the complicated nature of information processing within multi-site protein modification. By focusing on the vista of symmetry and symmetry breaking, we revealed a path to untangling the systems complexity and non-linearity inherent to the system and reveal novel behaviors, their unambiguous binary existence/possibility, and sourced the origins of behaviors. This approach of using symmetry and symmetry breaking could be of relevance to other complex networks that are otherwise shrouded in parametric, network and systems complexity.

**Exploring origins of biphasic response in protein modification networks**

Biphasic dose responses (non-monotonous bell shaped response to a monotonous input) are ubiquitous in signaling and are seen across multiple scales and levels, both in dynamic response and in steady state dose response. Such biphasic responses are typically seen as the cell maintaining an "optimal" range of activation/response in output, and are often associated with (or encountered in the context of) incoherent feedforward network motifs. In chapter 3, we explored the propensity of networks made of up of building blocks of protein modifications to present with such steady state biphasic dose responses.

Starting with a simple modification-demodification cycle (covalent modification system) and widely used modelling descriptions for enzymatic modifications, we constructed models of canonical modification networks. These included different topologies such as ordered double site modification, cascaded modification systems, and coupled modification systems (coupled by common enzyme or enzymes). In each of these networks, we considered different classes of enzyme action including common and separate enzymes effecting modifications.

A unified approach was adopted to explore biphasic response in both substrate and enzyme dose response. Our approach incorporated a hierarchy in the parameters; drawing a distinction between intrinsic kinetics (parameters that are largely inherent to the system), and total amounts of enzymes and substrates (parameters that can be easily tuned/modified). This hierarchy and synthesis of the parametric complexity allowed us to carefully answer questions regarding biphasic dose response, particularly since doses (total amounts) in the context of steady state dose response act as natural levers and dose for the analysis. Using a combination of analytical work (symbolic computer algebra) and bifurcation analysis, we showed how certain networks can present with substrate biphasic or enzyme biphasic, or biphasic response to both doses, depending on the underlying kinetics and total enzyme and substrate amounts. We revealed how certain networks can more readily present substrate or enzyme biphasic responses across all of kinetic parameter space for some combination of total amounts of enzymes and substrates, revealing how the behavior is available and accessible in all such modification networks inbuilt by default. For example, the two simple covalent modification cycles coupled through enzyme sharing, can readily present with substrate biphasic dose response for all underlying kinetics and total amounts of enzymes. Such results show even seemingly isolated systems with little complexity or cross talk - in this case only through enzyme sharing (a ubiquitous feature of cellular proteins), can give rise to complex and unexpected behavior. A simplistic description or understanding of such networks, particularly one that discounts the coupling or cross talk present, might disregard the possibility of such complex information processing from a simple covalent modification cycle.

This comphrehensive analysis of a suite of modification systems and analytical work

grounded in our hierarchical approach to parameters and a unified purview of substrate and enzyme biphasic dose responses, reveals the landscape of features at a network level (such as enzyme sharing and commonality of enzymes effecting modification) and at the biochemical modification level (parametric constraints including kinetic and total amounts), that enable or preclude the behavior. In particular, in key systems we show how biphasic behavior is counter poised with other complex information processing behavior such as multistability, revealing the richness of information processing capabilities and decision making that a simple enzymatic modification mechanism can confer. For example, the kinetic constraints that guarantee the absence of enzyme biphasic response in the ordered distributive DSP (with common enzyme and common phosphatase), is sufficient to guarantee multistability in the system for some total amounts of substrate and enzyme (see [74]). Thus our analysis can serve not only to engineer or preclude behavior such as biphasic responses but can also serve to eliminate or engineer interactions with other potentially interesting qualitative signaling.

We looked beyond these simple building blocks and explored our results in the context of the multi-level interactions taking place within cells. We studied the consequence of modification networks' to present biphasic readily in the context of wider network motifs in cellular circuits and showed how biphasic responses within such motifs, could subvert and introduce unexpected behavior to the network motif. For example, we showed how biphasic within interactions between nodes of a network motif such as negative feedback (with characteristic homeostatic input-output response) could enable increased multi-stability from the system. Similarly our results from analysis of such biphasic from within a incoherent feedforward interaction (with characteristic biphasic response itself) showed how it could subvert the expected biphasic response by introducing multiphasic response and even erode it.

Such results underscore the importance of including the complexity of biochemistry in the study cellular networks (when present within nodes or in arrows). We extended this analysis and approach to concrete systems by showing using modelling the robust presence of biphasic responses (to both enzyme and substrate doses) in the canonical ErK pathway. A summary of these different results is provided in section 3.5. Given the significant efforts recently to engineer biochemical circuits to obtain desired behavior, including synthetic gene regulatory incoherent feedforward networks to obtain biphasic responses, our analysis and insights present an alternative to leverage basic building blocks that are pervasive in biochemical networks to engineer such behavior. Analytical work, that is able to predict the presence, absence of biphasic responses and the conditions under which such behavior is guaranteed will be of particular interest to the synthetic biology community. In particular, if the over expression of specific substrates is undesired, modification systems can also be leveraged to engineer (by operation in suitable regimes of total enzyme amounts and substrates for example) robust safeguarding mechanisms against in.

Together these insights reveal how complex behavior at the level of the biochemistry (be it symmetry/symmetry breaking or biphasic dose response) can percolate up the levels affecting multi-level interactions, to present with drastic deviations from expected behavior. Protein modification is a pervasive theme in cellular signaling and given the multi-functionality of proteins and enzymes, there is a high degree of cross talk and coupling such that no network or protein modification is ever isolated and modular. Thus our insights and study underscores how biphasic responses could be leveraged by the cell from these simple ingredients to present robust biphasic and or other complex phenomena in cellular networks. Not accounting for this biochemical complexity, or pursuing a network analysis that ignores the complex input output characteristics of these systems could lead to a incomplete understanding. The relevance of these studies in particular to systems biologist, experimental work in synthetic biology and for network inference from data is underscored in the work.

## Intra-cellular signaling: A study of networks & network-motifs

Building on our understanding of the role that protein modification mechanism plays in cellular signaling (from chapters 2 and 3), and on work done by ex-research group member, Dr. *Thapanar Suwanmajo* in his thesis, in chapter 4 we explored their role as part of netowrk motifs more explicitly.

We studied this by considering a suite of double site modification systems (with different enzyme combinations) participating in canonical network motifs such as positive and negative feedback, and incoherent and coherent feedforward networks. Understanding this interplay is a substantial systems challenge which involves addressing the complexity of both the substrate modification level (sequestration, enzyme sharing, multisite), the network modification level (feedback and feedforward and the ways in which they can be effected through multisite modification), and the systems complexity of their interplay. To do this we adopted a dedicated three pronged methodology,

1. **Analytical work:** designed to provide unambiguous answers to the possibility or impossibility of systems response
2. **Targeted computational work:** revealing the multitudes of behaviors present
3. **Semi analytical work:** bridging the gap between insights from analytical and computational work

By using generic descriptions for widely used systems as representative exemplars of each level, we built a suite of models involving substrate modification and the network motif. This consideration by itself revealed the different nature of interactions possible between the two levels. For example, our discussion and results reveal how feedback through specific substrates participating in multisite modification can allow for mixed feedback network structures even when the substrate form is promoting the production of the kinase enzyme (which would otherwise be a positive feedback). Thus consider-

ation of (multisite) substrate modification and network motif level can fundamentally expand the capabilities of and the structure of the network level.

Using our three pronged approach we showed how behaviors from the two levels, can combine in non-intuitive ways to both generate new behavior at the systems level (that sometimes cannot be attributed to any given level), and subvert and erode existing behavior from either level. Our analysis revealed multitudes of such systems level phenomena arising in this interplay. For example, we showed how negative feedback, which is seen to characteristically present or induce oscillations in a system, can erode the strength and oscillations arising from the modification level (see fig. 4.3). In a similar light, incoherent feedback network which is seen as a biphasic modulator of the system can in some cases remove the biphasic behavior in the system (see fig. 4.5). Multi-level interaction here can also enable new behavior at the systems level. Examples of such introduction of behavior include the generation of increased multistability and tri-stability from negative feedback motifs (see fig. 4.3) and complex behaviors such as *isolas* from incoherent feedforward motifs (see fig. 4.5).

In addition to this, our work also revealed when one level (network or modification system) can enable transitions to different qualitative behaviors present in the behavioral landscape of the other level. Our semi-analytical work in this instance allows one to predict when such transitions are possible, or will be encountered. This methodology and framework can be a vital toolkit to engineer (or preclude) behavior at the systems level through suitable modulation/tuning/engineering of one of the levels.

Extensive analytical work, through considerations of simpler/reduced models (with key substrate-enzyme complexes removed) was pursed to source the origins of the systems behavior. This reduced analysis through analytical work (symbolic computational algebra) allowed us to clearly ascertain the role of key features of the substrate modification level (such as commonality of enzymes and complex formation for example) and the network motif level (direct/mediated feedback for example). Our framework and methodology can be adapted to studies of a multi-level nature such as this, to carefully discern the source of systems level behavior and its characteristics. A summary of these different results is provided in section 4.5.

Taken together the study establishes unambiguously how the modification level affects behavior from the network regulation level, and vice versa. From an inference perspective our study reveals the several pitfalls associated with extrapolating between levels, and presents a methodological framework to tackle multi-level information processing, with a clear focus on origins of behaviors, guaranteeing or tracing transitions between behaviors from one level, and predicting and engineering emergent systems behavior from the interplay. These aspects are not only critical for the understanding of cellular networks, but provide a fresh perspective of their capabilities and open up many possibilities for engineering interesting behavior from biochemical networks in synthetic biology.

## Tackling multi-level interactions in cells and populations

In chapter 5, we looked beyond the cell and studied multi-level complexity across communicating cells and populations.

Our focus was grounded in two sub-studies focussing on, 1. interactions between cells within a population of a single cell type (single population study) and 2. interactions between cells types belonging to different populations. The first study (single population study) focussed on understanding the multi-level interactions between the three distinct levels, 1. cellular circuit, 2. population level, and 3. environmental level. In the two population study, we focussed in a similar spirit on the cross population interactions, particularly between circuits and populations of different cell types. As an exemplar we also considered a sub-study with emergent behavior at the population level through interactions between them, and studied how these percolate through and affect the collective behavior seen.

A systematic framework was built where in we considered in increasing complexity the role of interactions between these different levels - eventually building to a scenario where all interactions between these different levels and populations were present. To do this, we built generic master models of single and two populations of cells communicating across levels, that can capture the effect of multi-level cross population interactions. We then used a computational automation pipeline to generate bifurcations of the systems behavior as key parameters are changed to better understand the emergence of systems level behavior and their dependence on cellular interactions.

The single population study explored the role of the various interactions in effecting population level response. In particular, the study revealed several instances in which population response can deviate from the cellular response drastically and other instances in which the population response might mask the underlying amplitude and strength of cellular response. These careful explorations revealed the consequences of being restricted in thinking to any one level, even when the system as a whole is interconnected. Building on these the two population study further revealed how these effects get exasperated with cross population interactions. This alike to the systems level behavior seen at the intra-cellular level, can give rise to novel behavior that cannot be solely attributed to any one population or circuit (for example, combination of oscillation and multistability, and increased multi-stability). The pitfalls with inference can also further get exacerbated; where different levels show different degrees and amplitudes of response. This allows for the system to mask the underlying response. For example, as seen in fig. 5.23, different steady states could register same concentration at the population level of one population type while allowing for completely different cellular level responses and population level responses at the other cell type.

Our framework and analysis reveals several key insights of broad relevance to the understanding of cellular circuits and information processing through them. Our analysis shows several examples of borrowed qualitative behavior transcending levels (and

populations); sometimes registering directly, and sometimes registering with characteristic features of the levels that enabled such behavior. We show how key interactions between and across levels is a recipe to generate novel and interesting systems behavior that cannot be otherwise obtained from the individual levels. From a synthetic biology perspective engineering these interactions can often be the simplest (and most robust) way to obtain these behaviors at the systems level. We also expand on the relevance of our results to inferences particularly in light of the many results that suggest how through dilution and characteristics of a level, various species concentrations could be masked or could mask underlying behavior. They underscore the pitfalls with treating variables from different levels (or even the same level) as proxies to draw inferences on the underlying behavior of the system. That these networks don't exist in isolation and often regulate or participate in other key cellular pathways exemplifies the aspect of further potential downstream regulation lending behavior and features seen here from multi-level interactions to other cellular processes.

All in all, our approach and framework, using minimal assumptions and generic models in chapter 5 presents a unified first step towards tackling the complexity of these interactions present in cellular signaling.

**Future work**

Our analysis along these research strands revealed several pitfalls associated with thinking rooted in a single level. However our results and synthesis show value in extending our framework to explore the following directions of study.

*Leveraging the vista of symmetry in modification systems beyond two sites*

Our approach to employ symmetry as a new vista to explore the intrinsic capabilities of multisite modification (random double site modifications sites) revealed that such systems through symmetry breaking could provide interesting features such as ACR, multi-stability and oscillations. ACR, a sophisticated version of exact homeostasis (as key parameters or as in this case total amounts of substrate was changed) was in fact not only seen in concentrations of substrates but in sums of concentrations of substrates.

The majority of this study however was restricted to protein networks with two modification sites. However as mentioned earlier, proteins undergo a multitude of such modifications (by either the same enzyme or by different enzymes). A typical protein like the *p53* contains 100s of modification sites [203]. Thus in addition to capturing a more realistic scenario of protein modification in vivo, such an exploration of how this approach scales is likely to

1. *Offer a multiplicity of symmetric vistas to study complexity of multisite networks:* number of symmetries scale rapidly with increasing modification sites
2. *Reveal complex features:* potentially reveal complex forms of ACR and other sys-

tems behaviors

A result to this end, suggestive of such pockets of interesting insights was shown in fig. 2.12 (chapter 2), where a triple site ordered system also showed shared ACR between the concentrations of the partially modified substrate forms. This potential is substantiated by the richness of the non-linearity (through sequestration) afforded by more modification sites, along with the increase in the combinations of enzyme action possible on those sites (common and/or separate). In addition to this, the mechanisms of modifications can also interact and combine to reveal several classes of distributive, processive and mixed mechanisms, each of which increase the combinatorial possibilities of vistas of symmetry.

Taken together such a study building the current work, will represent more realistic scenarios of modification networks in-vivo and can be fruitful in revealing the complex information processing capabilities of multisite modification networks.

*Biphasic responses from enzymatic modifications - beyond the building blocks, the distributive mechanism, and exemplars:*

In chapter 3, we looked at how simple networks built from enzymatic modification mechanism as building blocks, when operating in a distributive manner, was capable of presenting with biphasic responses in a robust manner. However, distributive is but one potential mechanism with which enzymatic modification is effected (within multisite modification). Similarly, the basic topologies explored there are bare examples of the complicated ways in which modifications present within cellular networks. A pertinent example of such complexity at the network level is our study of the *ERK* pathway with (even its minimal network description) containing a random modification topology and combinations of reversible and irreversible modifications. Such combinations of building blocks of modification systems recur heavily in cellular signaling pathways and thus a future study focussing such higher order systems, reflecting possible other mechanisms (processive or mixed), topologies, cross-talk and coupling, can shine more light into the origins of the behavior in key cellular signaling pathways, and the propensity of modification systems to enable them.

One possible focus of the study could be the canonical *MAPK* pathway. Akin to the *ERK* pathway, *MAPK*, with its cascaded network structure, represents a cornerstone of cellular signaling, and participates in and regulates key cellular processes including cell proliferation, motility and survival [204]. Biphasic responses have been observed in the pathway [205], and thus makes it a suitable candidate for our study. In addition, there is abundant evidence of network regulation within the pathway, including feedback and feedforward [146, 51]. Focusing on such regulation within this exemplar, will allow for a convergence in our understanding of modification systems ability to present biphasic responses, as well as their behavior in relation to network regulation.

Another fruitful avenue is the extension of our analysis on consequences of biphasic

responses to network motif response and to networks more broadly. Given the ubiquity of modification systems in cellular signaling, a majority of the network description of any cellular pathway or process contains multiplicities of them in different nodes and arrows. A study focussing on how distributions of multiple biphasic responses within a network or network motif affects the expected response will not only provide relevant insights into the functioning of (and inferences for) multiple cellular pathways, but also reveal the full extent of the implication of such biphasic responses from inert sources of complexity such as modification systems.

*Exploring the role of sequestration in modification systems and & network motif:*

In chapter 4, we showed how simple modification systems when part of network motifs can fundamentally subvert expected network motif response eroding expected response and/or introduce new behavior. Our analysis of this however didn't consider the role of sequestration of the modified substrate (or the species involved in modification) at the network regulation level. While our study served as a first principle study to explore the consequences of such multi-level interactions, multi-level interactions often involve sequestration of the species at the other level. This is only underscored by the fact that such sequestration can also stem from other sources such as cross talk and coupling, given that a substrate or a component of the network level might be fulfilling another role in a distinct process (reflecting the multiplicity in the functionality of an enzyme). Network regulation can also involve and or interact with such cross talk. For example, studies on Hedgehog signaling have revealed how cross talk can contribute to feedback regulation of processes (and interact with feedback regulation) [206]. Thus, an exploration of our systems using our framework and three pronged methodology with such sequestration accounted for will present a foray into tackling this complexity and reveal the complete scope of possibilities that emerge from the resulting systems behavior.

*Role of interaction strengths in cellular communication:*

In chapter 5, we studied how cells interact with one and other and revealed a landscape of interesting features arising from such interactions both at the single cell level, and at the level of populations of cells. However the majority of the analysis relied on the interactions considered (as part of the scenarios there in) to be of equal strength. However, in reality, these interactions are often of different strengths. Immediate next steps should involve analysis focussing on the role of parameters of the population level and the environmental level in perturbing observed population and cellular level response, particularly exploring how the relative strengths of the different interactions can affect observed systems response would better reveal the facets of multi-level interactions.

*Testable predictions in inter-population communication using engineered systems*

With recent advances over the last two years in synthetic biology, there is finally significant progress in efforts to engineer population level behavior [54, 58, 207, 208, 209].

Such a landscape is ripe for leveraging to study our predictions and understanding of cell-cell communication. A study focussing on an exemplar of engineered population level response (such as [173]), superimposed with the framework of modelling and analysis developed in chapter 5, will allow us to better refine our insights, and also predict and aid in synthetic engineering of such responses at each level.

## Summary

Together these chapters (and this thesis) reveal the pitfalls of studies focussing on and drawing inferences from any one level (be it simple biochemistry such as in protein modification, network motifs, behavior from and of a single cell, population response or by observing environmental factors). We showed how even within a single level (say, biochemistry of modification), there is much to be understood about modular response, and by bringing new vistas (symmetry) and methodology (a three pronged computational, analytical and semi-analytical toolkit), our work has revealed not just interesting results in the context of biochemistry of modification, but how such tools and methodologies can tackle the systems challenges present at this cross-section and help provide better understanding of cellular information processing.

The unifying theme in these studies, is our implementation of a structured systems framework that tackles this complexity in a bite sized manner and designed to build on past insights as more complexity is added on to the system. These approaches and studies will be of great relevance to researchers in the systems and synthetic biology communities, and is guaranteed to help and equip researchers to better understand multi-level networks and inference from them.

# Appendix A

# Parameters

Unless otherwise stated all parameter values (including total amounts) are assumed to be 1.

## A.1 Symmetry and symmetry breaking in multisite modification

**Figure 2.2:** ***Case 1 and case 2 symmetry breaking in various double site phosphorylation networks***

A *Case 1 - Ordered Distributive DSP*
$k_1$ = 0.1; $k_2$ = 0.5; $k_3$ = 0.1; $k_4$ = 0.5; $K_{Total}$ = 0.1; $P_{Total}$ = 0.1;

B *Case 1 - Random System 1 (Common kinase and phosphatase)*
$k_1$ = 0.1; $k_3$ = 0.1; $a_1$ = 0.25; $a_2$ = 0.4; $a_3$ = 0.25; $a_4$ = 0.4; $K_{Total}$ = 1; $P_{Total}$ = 1;

C *Case 1 - Random System 3 (Separate kinase and phosphatase)*
$k_1$ = 0.1; $k_3$ = 0.1; $a_1$ = 0.5; $a_2$ = 1.5; $a_3$ = 0.5; $a_4$ = 1.5; $K1_{Total}$ = 1; $P1_{Total}$ = 1; $K2_{Total}$ = 1; $P2_{Total}$ = 1;

D *Case 2 - Random System 3 (Separate kinase and phosphatase)* $k_1$ = 2.35; $k_2$ = 0.46; $k_3$ = 1.86; $k_4$ = 1.1; $a_1$ = 2.35; $a_2$ = 0.46; $a_3$ = 1.86; $a_4$ = 1.1; $K1_{Total}$ = 1; $P1_{Total}$ = 1; $K2_{Total}$ = 1; $P2_{Total}$ = 1;

E *Case 2 - Mixed Random 2 (Separate kinase and processive dephosphorylation by common phosphatase*
$k_1$ = 2; $k_2$ = 0.1; $k_3$ = 0.75; $a_1$ = 2; $a_2$ = 0.1; $a_3$ = 0.75; $K1_{Total}$ = 0.1; $K1_{Total}$ = 0.1; $P_{Total}$ = 0.2;

**Figure 2.3:** ***Case 3 symmetry breaking in various double site phosphorylation networks***

A *Panel 1 - Hopf and Pitchfork Bifurcation*
$k_1$ = 100; $k_2$ = 2; $k_3$ = 0.01; $k_4$ = 20; $a_1$ = 0.01; $a_2$ = 20; $a_3$ = 100; $a_4$ = 2; $k_{b1}$ = 100; $k_{b3}$ = 100; $k_{b4}$ = 0.1; $a_{b1}$ = 100; $a_{b3}$ = 100; $a_{b2}$ = 0.1; $K_{Total}$ = 1.25; $P_{Total}$ = 1.25;

A *Panel 2 - Pitchfork Bifurcation*
$k_1$ = 100; $k_2$ = 2; $k_3$ = 0.01; $k_4$ = 20; $a_1$ = 0.01; $a_2$ = 20; $a_3$ = 100; $a_4$ = 2; $k_{b1}$ = 100; $k_{b3}$ = 100; $k_{b4}$ = 0.1; $a_{b1}$ = 100; $a_{b2}$ = 0.1; $a_{b3}$ = 100; $K_{Total}$ = 10; $P_{Total}$ = 10;

B *Panel 1 - Hopf Bifurcation*
$k_1$ = 150; $k_2$ = 50; $k_3$ = 1; $k_4$ = 10; $a_1$ = 1; $a_2$ = 10; $a_3$ = 150; $a_4$ = 50; $k_{b1}$ = 100; $k_{b3}$ = 0.01; $k_{b4}$ = 500; $a_{b1}$ = 0.01; $a_{b2}$ = 500; $a_{b3}$ = 100; $K1_{Total}$ = 1; $P1_{Total}$ = 1; $K2_{Total}$ = 1; $P2_{Total}$ = 1;

B *Panel 2 - Pitchfork Bifurcation*
$k_1$ = 10; $k_2$ = 1; $k_3$ = 2; $k_4$ = 5; $a_1$ = 2; $a_2$ = 5; $a_3$ = 10; $a_4$ = 1; $K1_{Total}$ = 1; $P1_{Total}$ = 1; $K2_{Total}$ = 1; $P2_{Total}$ = 1;

**Figure 2.5:** ***Schematic representation of the mixed random ordered DSP network***

$k_1$ = 0.9; $k_2$ = 0.8; $k_3$ = 2; $a_1$ = 0.9; $a_2$ = 0.8; $a_3$ = 2; $K1_{Total}$ = 0.1; $K2_{Total}$ = 0.1;

**Figure 2.6:** ***Case 3 symmetry breaking in random ordered DSP with common kinase and common phosphatase acting on each modification site***

A *Panel 1 - Hopf Bifurcation*
$k_1$ = 100; $k_2$ = 20; $k_4$ = 2; $a_2$ = 2; $a_3$ = 100; $a_4$ = 20; $k_{b1}$ = 100; $k_{b3}$ = 100; $a_{b1}$ = 100; $a_{b3}$ = 100; $K_{Total}$ = 1; $P_{Total}$ = 1;

B Panel 2 - Dynamic simulation:
Parameter set in Figure 2.3A where $A_{Total}$ = 12.95;

C *Panel 3 - Pitchfork Bifurcation*
$k_1$ = 0.1; $k_3$ = 2; $k_4$ = 5; $a_1$ = 2; $a_2$ = 5; $a_3$ = 0.1; $K_{Total}$ = 1; $P_{Total}$ = 1;

**Figure 2.7:** ***Case 1 symmetry breaking in the random ordered DSP with common kinase and common phosphatase acting on each modification site, even when one of the legs is incapable of breaking symmetry by itself***

$k_1$ = 0.1; $k_3$ = 0.1; $K_{Total}$ = 1; $P_{Total}$ = 1;

**Figure 2.8:** ***Case 3 symmetry breaking through a supercritical pitchfork bifurcation in the random ordered DSP with separate kinase and separate phosphatase acting on each modification site***

$k_2$ = 2; $k_4$ = 20; $a_2$ = 20; $a_4$ = 2; $k_{b1}$ = 10; $k_{b3}$ = 10; $a_{b1}$ = 10; $a_{b3}$ = 10; $K1_{Total}$ = 5; $P1_{Total}$ = 5; $K2_{Total}$ = 5; $P2_{Total}$ = 5;

**Figure 2.9:** ***Features of symmetry and symmetry breaking in systems near symmetric***

$k_1$ = 1.25; $k_2$ = 1.1; $k_3$ = 2.5; $k_4$ = 0.4; $a_1$ = 1.25; $a_3$ = 2.5; $a_4$ = 0.4; $K1_{Total}$ = 1; $P1_{Total}$ = 1; $K2_{Total}$ = 1; $P2_{Total}$ = 1;

**Figure 2.10:** ***Approximate concentration robustness shown by systems that deviate from exact symmetry***

$k_2$ = 0.5; $k_3$ = 0.1; $k_4$ = 0.5; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $K_{Total}$ = 0.1; $P_{Total}$ = 0.1;

**Figure 2.11:** ***Total enzyme concentrations are an additional lever (apart from basic network kinetics) that can independently tune symmetry breaking behavior in MSP networks***

- *Left Panel*
  $k_1$ = 30; $k_2$ = 2; $k_3$ = 0.3; $k_4$ = 20; $a_1$ = 0.3; $a_2$ = 20; $a_3$ = 30; $a_4$ = 2; $k_{b1}$ = 100; $k_{b3}$ = 100; $k_{b4}$ = 0.1; $a_{b1}$ = 100; $a_{b2}$ = 0.1; $a_{b3}$ = 100; $K_{Total}$ = 1; $P_{Total}$ = 1;
- *Right Panel*
  $K_{Total}$ = 20; $P_{Total}$ = 20;

**Figure 2.12:** ***Case 1 symmetry breaking in distributive ordered TSP network***

$k_1$ = 0.1; $k_2$ = 1.5; $k_3$ = 2; $k_4$ = 0.1; $k_5$ = 1.5; $k_6$ = 2; $K_{Total}$ = 0.1; $P_{Total}$ = 0.1;

**Figure 2.13:** ***Absolute concentration robustness in ordered DSP***

$k_1$ = 1; $k_2$ = 2; $k_3$ = 0.5; $k_4$ = 1.2; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 11; $K_{Total}$ = 1; $P_{Total}$ = 1;

## A.2 Biphasic responses in cell signalling: A unified approach

**Figure 3.2:** ***Presence and absence of substrate and enzyme biphasic dose responses in the commonly observed enzymatic cellular signaling systems***

A *Enzyme biphasic dose response in PPI model*
$k_1$ = 1; $k_4$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 20; $k_{b4}$ = 1; $k_{ub1}$ = ; $k_{ub2}$ = 1; $k_{ub4}$ = 1;
$A_{Total}$ = 30; $P_{Total}$ = 1;

B *Enzyme and Substrate biphasic in DSP with common enzyme action*
$k_1$ = 0.1; $k_2$ = 1; $k_3$ = 1; $k_4$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $K_{Total}$ = 2.5; $P_{Total}$ = 1;

C *Substrate biphasic in DSP with common kinase and separate phosphatase action*
$k_1$ = 3.5; $k_2$ = 50; $k_3$ = 100; $k_4$ = 100; $k_{b1}$ = 20; $k_{b2}$ = 75; $k_{b3}$ = 50; $k_{b4}$ = 30; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $K_{Total}$ = 6; $P1_{Total}$ = 1; $P2_{Total}$ = 1;

D *Enzyme biphasic in DSP with separate kinase and common phosphatase action*
$k_1$ = 200; $k_2$ = 40; $k_3$ = 200; $k_4$ = 5; $k_{b1}$ = 250; $k_{b2}$ = 100; $k_{b3}$ = 100; $k_{b4}$ = 75; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $A_{Total}$ = 50; $K1_{Total}$ = 18; $P_{Total}$ = 30;

E *Substrate biphasic in the cascaded enzymatic model with common phosphatase*
$k_1$ = 80; $k_2$ = 35; $k_{b1}$ = 200; $k_{b2}$ = 200; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $p_1$ = 5; $p_3$ = 60; $p_{b1}$ = 40; $p_{b2}$ = 25; $p_{ub1}$ = 1; $p_{ub2}$ = 1; $A_{Total}$ = 45; $K_{Total}$ = 10; $P_{Total}$ = 16;

F *Substrate biphasic in coupled covalent modification cycles with common enzymes*
$k_1$ = 4; $k_2$ = 0.1; $k_{b1}$ = 0.9; $k_{b2}$ = 8; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $p_1$ = 0.1; $p_3$ = 3; $p_{b1}$ = 1; $p_{b2}$ = 1; $p_{ub1}$ = 1; $p_{ub2}$ = 1; $A_{Total}$ = 8; $K_{Total}$ = 4; $P_{Total}$ = 3;

**Figure 3.3:** ***Role of biphasic interactions in network and motif response***

A *Upstream biphasic regulation of covalent modification cycle*
$u_a$ = 1; $u_b$ = 0.01; $u_d$ = 1; $b_1$ = 2.42*4; $b_2$ = sqrt(2)/20; $k_1$ = 4; $k_2$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $A_{Total}$ = 20; $K_{Total}$ = 5; $P_{Total}$ = 3;

B *Upstream biphasic regulation of double site modification system (with common enzymes)*

(a) $u_a$ = 1; $u_b$ = 0.01; $u_d$ = 1; $b_1$ = 2.42*1; $b_2$ = sqrt(2)/3; $k_1$ = 1; $k_2$ = 2; $k_3$ = 1; $k_4$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1;
$A_{Total}$ = 20; $K_{Total}$ = 4; $P_{Total}$ = 3;

(b) $u_a$ = 0.1; $u_b$ = 0.01; $u_d$ = 0.1; $b_1$ = 2.42*0.1; $b_2$ = sqrt(2)/8; $k_1$ = 1; $k_2$ = 5; $k_3$ = 1; $k_4$ = 8; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1;
$A_{Total}$ = 20; $K_{Total}$ = 5; $P_{Total}$ = 2;

(c) $u_a$ = 0.02; $u_b$ = 0.001; $u_d$ = 40; $b_1$ = 2.42*0.5; $b_2$ = sqrt(2)/2; $k_1$ = 0.5; $k_2$ = 0.1; $k_3$ = 10; $k_4$ = 1; $k_{b1}$ = 10; $k_{b2}$ = 10; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1;
$A_{Total}$ = 60; $K_{Total}$ = 75; $P_{Total}$ = 3;

C *Incoherent feedforward regulation of double site modification system (with common enzymes)*

(a) $u_a$ = 0.02; $u_b$ = 0.001; $u_d$ = 40; $d_a$ = 0.1; $d_b$ = 0.001; $d_d$ = 0.4; $k_1$ = 3; $k_2$ = 0.1; $k_3$ = 10; $k_4$ = 1; $k_{b1}$ = 10; $k_{b2}$ = 10; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1;
$A_{Total}$ = 60; $K_{Total}$ = 75; $P_{Total}$ = 3;

(b) $u_a$ = 0.2; $u_b$ = 0.001; $u_d$ = 30; $d_a$ = 4; $d_b$ = 0.001; $d_d$ = 20; $k_1$ = 0.5; $k_2$ = 0.1; $k_3$ = 10; $k_4$ = 1; $k_{b1}$ = 10; $k_{b2}$ = 10; $k_{b3}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $A_{Total}$ = 100; $K_{Total}$ = 75; $P_{Total}$ = 2;

D *Incoherent feedforward network motif*
$b_1$ = 2.42*2; $b_2$ = sqrt(2)/2; $k_1$ = 1; $k_2$ = 1; $K_1$ = 0.1; $K_2$ = 1; $p_1$ = 1; $p_2$ = 1; $P_1$ = 0.1; $P_2$ = 4; $d_1$ = 2; $d_2$ = 2; $D_1$ = 0.1; $D_2$ = 1; $A_{Total}$ = 1; $B_{Total}$ = 2; $R_{Total}$ = 4;

E *Positive feedback network motif*
$b_1$ = 2.42*0.1; $b_2$ = sqrt(2)/3; $k_0$ = 0.1; $k_1$ = 0.4; $K_1$ = 10; $k_2$ = 2; $K_2$ = 2; $k_3$ = 1; $k_4$ = 1; $K_4$ = 0.1; $A_{Total}$ = 3; $R_{Total}$ = 10;

F *Negative feedback network motif*
$b_1$ = 13.0266*0.1; $b_2$ = 1; $k_1$ = 0.1; $K_1$ = 2; $k_2$ = 0.1; $k_3$ = 0.1; $k_4$ = 0.1; $K_4$ = 0.5; $R_{Total}$ = 50; $A_{Total}$ = 5;

**Figure 3.4:** ***Presence and consequence of biphasic dose responses in cellular signaling arising from enzyme mediated substrate modification***

A *Substrate biphasic in the Erk regulation model*
$k_1$ =0.0747; $k_2$ =0.0957; $k_4$ =0.0104; $k_{b1}$ =0.18; $k_{b2}$ =0.18; $k_{b4}$ =0.18; $k_{ub1}$ =0.27; $k_{ub2}$ =0.27; $k_{ub4}$ =0.27;
$p_2$ =0.0338; $p_3$ =0.00851; $p_4$ =0.0106; $p_{b2}$ =0.18; $p_{b3}$ =0.18; $p_{b4}$ =0.18; $p_{ub2}$ =0.27; $p_{ub3}$ =0.27; $p_{ub4}$ =0.27;
$MeK_{Total}$= 0.12; $P_{Total}$ = 1;

B *Enzyme biphasic in the Erk regulation model*
$k_1$ = 2; $k_2$ = 1; $k_4$ = 1; $k_{b1}$ = 0.18; $k_{b2}$ = 0.18; $k_{b4}$ = 0.18; $k_{ub1}$ = 0.27; $k_{ub2}$ = 0.27; $k_{ub4}$ = 0.27;
$p_2$ = 0.1; $p_3$ = 5; $p_4$ =0.1; $p_{b2}$ = 0.18; $p_{b3}$ = 0.18; $p_{b4}$ = 0.18; $p_{ub2}$ = 0.27; $p_{ub3}$ = 0.27; $p_{ub4}$ = 0.27;
$Erk_{Total}$ = 100; $P_{Total}$ = 5;

C *Simultaneous presence of substrate and enzyme biphasic response in the Erk regulation model*
$k_1$ = 0.1; $k_2$ = 1; $k_4$ = 10; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub4}$ = 1;
$p_2$ = 0.01; $p_3$ = 40; $p_4$ = 1; $p_{b2}$ = 1; $p_{b3}$ = 1; $p_{b4}$ = 1; $p_{ub2}$ = 1; $p_{ub3}$ = 1; $p_{ub4}$ = 1;
$P_{Total}$ = 5;

**Figure 3.5:** ***Coexistence of multi-stability and biphasic dose response in various enzymatic models for the same underlying kinetic regime***

A *Enzyme biphasic - DSP with common kinase and common phoshpatase*
Multi-stability: $k_1$ = 100; $k_2$ = 6; $k_3$ = 5; $k_4$ = 20; $k_{b1}$ = 0.5; $k_{b2}$ = 30; $k_{b3}$ = 100; $k_{b4}$ = 750; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $A_{Total}$ = 65; $P_{Total}$ = 30;
Enzyme biphasic: $k_1$ = 100; $k_2$ = 6; $k_3$ = 5; $k_4$ = 20; $k_{b1}$ = 0.5; $k_{b2}$ = 30; $k_{b3}$ = 100; $k_{b4}$ = 750; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $A_{Total}$ = 80; $P_{Total}$ = 30;

B *Substrate biphasic - DSP with common kinase and common phosphatase*
$k_1$ = 0.1; $k_2$ = 1; $k_3$ = 1; $k_4$ = 1; $k_{b1}$ = 10; $k_{b2}$ = 1; $k_{b3}$ = 0.1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $K_{Total}$ = 2.5; $P_{Total}$ = 1;

C *Substrate biphasic - DSP with common kinase and separate phosphatase*
$k_1$ = 0.1; $k_2$ = 1; $k_3$ = 1; $k_4$ = 1; $k_{b1}$ = 10; $k_{b2}$ = 1; $k_{b3}$ = 0.1; $k_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $K_{Total}$ = 5; $P1_{Total}$ = 1; $P2_{Total}$ = 2;

D *Enzyme biphasic - DSP with separate kinase and common phosphatase*
$k_1$ = 30; $k_2$ = 5; $k_3$ = 50; $k_4$ = 1; $k_{b1}$ = 25; $k_{b2}$ = 10; $k_{b3}$ = 10; $k_{b4}$ = 7; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $A_{Total}$ = 200; $K1_{Total}$ = 0.1; $P_{Total}$ = 6;

E *Substrate biphasic response - Cascaded enzymatic network*
$k_1$ = 80; $k_2$ = 35; $k_{b1}$ = 600; $k_{b2}$ = 400; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $p_1$ = 5; $p_3$ = 60; $p_{b1}$ = 40; $p_{b2}$ = 25; $p_{ub1}$ = 1; $p_{ub2}$ = 1; $A_{Total}$ = 45; $K_{Total}$ = 10; $P_{Total}$ = 16;

**Figure 3.6:** ***Biphasic responses in random double site modification network with separate enzymes effecting each modification***

A *Substrate biphasic in random ordered DSP with separate enzymes*
$k_1$ = 1; $k_2$ = 1; $k_3$ = 1; $k_4$ = 1; $a_1$ = 1; $a_2$ = 4; $a_3$ = 1; $a_4$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $a_{b1}$ = 1; $a_{b2}$ = 1; $a_{b3}$ = 1; $a_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $a_{ub1}$ = 1; $a_{ub2}$ = 1; $a_{ub3}$ = 1; $a_{ub4}$ = 1;
$K1_{Total}$ = 2; $K2_{Total}$ = 2; $P1_{Total}$ = 3; $P2_{Total}$ = 3;

B *Enzyme biphasic in random ordered DSP with separate enzymes*
$k_1$ = 3; $k_2$ = 1; $k_3$ = 0.1; $k_4$ = 1; $a_1$ = 0.1; $a_2$ = 10; $a_3$ = 1; $a_4$ = 1; $k_{b1}$ = 1; $k_{b2}$ = 1; $k_{b3}$ = 1; $k_{b4}$ = 1; $a_{b1}$ = 1; $a_{b2}$ = 1; $a_{b3}$ = 1; $a_{b4}$ = 1; $k_{ub1}$ = 1; $k_{ub2}$ = 1; $k_{ub3}$ = 1; $k_{ub4}$ = 1; $a_{ub1}$ = 1; $a_{ub2}$ = 1; $a_{ub3}$ = 1; $a_{ub4}$ = 1; $A_{Total}$ = 50; $K1_{Total}$ = 2; $P1_{Total}$ = 12; $P2_{Total}$ = 12;

**Figure 3.7:** ***Effect of biphasic response in interactions within feedback network motifs (open systems)***

A *Positive feedback motif*
$b_1$ = 0.1; $b_2$ = 5; $k_0$ = 0.01; $k_1$ = 0.5; $k_2$ = 1; $k_3$ = 1; $k_4$ = 1; $k_4$ = 0.5; $K_4$ = 0.1; $A_{Total}$ = 3;

B *Negative feedback motif*
$b_1$ = 0.5, $b_2$ = 0.8; $k_0$ = 0.1; $k_1$ = 0.1; $k_2$ = 0.001; $k_3$ = 0.1; $k_4$ = 2; $K_3$ = 0.01; $K_4$ = 0.01; $A_{Total}$ = 3;

**Figure 3.8:** ***Effect of biphasic response in interactions within integral feedback control motif: Perturbation of homeostatic response***

A *Expected system response*
$ka_d$ = 1; $kp_a$ = 1; $kp_d$ = 1; $km_a$ = 1; $km_d$ = 1;

B *Upstream biphasic regulation by signal*
$b_2$ = 0.1; $b_2$ = 0.8; $b_2$ = 2.5;

C *Biphasic in interaction*
$b_2$ = 0.1; $b_2$ = 0.8; $b_2$ = 2.5;

## A.3 Network regulation meets substrate modification chemistry

**Figure 4.2:** ***Impact of positive feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system***

A *Perturbation of monostable response*
$u_a$ = 0.1, $u_d$ = 100, $u_{beta}$ = 0.01, $d_a$ = 0.01, $d_d$ = 1, $k_1$ = 1, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 5, $X_{Total}$ = 1, $k_{PFB}$ = 0,1,2,3,5.

B *Bistability enabled by positive feedback*
$u_a$ = 0.1, $u_d$ = 100, $u_{beta}$ = 0.01, $d_a$ = 0.01, $d_d$ = 1, $k_1$ = 1, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 5, $X_{Total}$ = 1, $k_{PFB}$ = 5,10,15,20.

C *Perturbation of biphasic response*
$u_a$ = 1, $u_d$ = 100, $u_{beta}$ = 0.01, $d_a$ = 1, $d_d$ = 1, $k_1$ = 1, $k_2$ = 0.01, $k_3$ = 10, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 70, $X_{Total}$ = 5, $k_{PFB}$ = 0,0.1.

D *Perturbation of bistable response*
$u_a$ = 0.01, $u_d$ = 125, $u_{beta}$ = 0.001, $d_a$ = 20, $d_d$ = 1, $k_1$ = 0.01, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 100, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 70, $X_{Total}$ = 2, $k_{PFB}$ = 0,0.1.

E *Tri-stability enabled by positive feedback*
$u_a$ = 0.01, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 18, $d_d$ = 12, $k_1$ = 3, $k_2$ = 60, $k_3$ = 40, $k_4$ = 12, $k_{b1}$ = 40, $k_{b2}$ = 2250, $k_{b3}$ = 250, $k_{b4}$ = 300, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 70, $P_{Total}$ = 1, $K_{Total}$ = 2, $X_{Total}$ = 5, $k_{PFB}$ = 6.1.

F *Perturbation of oscillatory response*
$u_a$ = 0.01, $u_d$ = 1, $u_{beta}$ = 0.001, $d_a$ = 200, $d_d$ = 1, $k_1$ = 1, $k_2$ = 10, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 40, $k_{b2}$ = 5, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 2, $X_{Total}$ = 2, $k_{PFB}$ = 0,0.001,0.002

**Figure 4.3:** ***Impact of negative feedback mediated by the fully modified substrate $A_{pp}$ on the behavior of the double site modification system***

A *Perturbation of monostable response*
$u_a$ = 1, $u_d$ = 20, $u_{beta}$ = 0.01, $d_a$ = 200, $d_d$ = 1, $k_1$ = 0.5, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 30, $k_{b2}$ = 10, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 3, $X_{Total}$ = 5, $k_{NFB}$ = 0,10,25,50,100.

B *Perturbation of biphasic response*
$u_a$ = 1, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 1, $d_d$ = 1, $k_1$ = 0.5, $k_2$ = 0.1, $k_3$ = 10, $k_4$ = 1, $k_{b1}$ = 100, $k_{b2}$ = 100, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 40, $X_{Total}$ = 2, $k_{NFB}$ = 0,10,25,50,100.

C *Perturbation of bistable response*
$u_a$ = 1, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 1, $d_d$ = 1, $k_1$ = 0.5, $k_2$ = 1, $k_3$ = 0.9, $k_4$ = 50, $k_{b1}$ = 1, $k_{b2}$ =

1, $k_{b3}$ = 200, $k_{b4}$ = 3, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 10, $X_{Total}$ = 4, $k_{NFB}$ = 0,5,10,15.

D *Increased multistability (tri-stability) and oscillations enabled by negative feedback*
$u_a$ = 1, $u_d$ = 4, $u_{beta}$ = 0.01, $d_a$ = 2, $d_d$ = 1, $k_1$ = 1, $k_2$ = 1, $k_3$ = 0.9, $k_4$ = 100, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 200, $k_{b4}$ = 3, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 20, $X_{Total}$ = 10, $k_{NFB}$ = 0,2,4.

E *Perturbation of oscillatory response*
$u_a$ = 0.01, $u_d$ = 2, $u_{beta}$ = 0.001, $d_a$ = 200, $d_d$ = 1, $k_1$ = 1, $k_2$ = 10, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 40, $k_{b2}$ = 5, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 0, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 2, $X_{Total}$ = 2, $k_{NFB}$ = 0,0.1,0.25.

**Figure 4.4:** ***Impact of feedback mediated by the partially modified substrate $A_p$ (through activation or inactivation of kinase) on the behavior of the double site modification system***

A, E *Perturbation of monostable response*
$u_a$ = 0.1, $u_d$ = 100, $u_{beta}$ = 0.01, $d_a$ = 0.01, $d_d$ = 1, $k_1$ = 1, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 5, $X_{Total}$ = 1, $k_{Ap-Act}$ = 0,1,2,3,5. $k_{Ap-InAct}$ = 0,50,100,250,500.

B, F *Perturbation of biphasic response*
$u_a$ = 1, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 1, $d_d$ = 1, $k_1$ = 0.5, $k_2$ = 0.1, $k_3$ = 10, $k_4$ = 1, $k_{b1}$ = 100, $k_{b2}$ = 100, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 40, $X_{Total}$ = 2, $k_{Ap-Act}$ = 0,0.01,0.02,0.05. $k_{Ap-InAct}$ = 0,1,5,25.

C, G *Perturbation of bistable response*
$u_a$ = 1, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 1, $d_d$ = 1, $k_1$ = 0.5, $k_2$ = 1, $k_3$ = 0.9, $k_4$ = 50, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 200, $k_{b4}$ = 3, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 10, $X_{Total}$ = 4, $k_{Ap-Act}$ = 0,0.01,0.05,0.1. $k_{Ap-InAct}$ = 0,1,2,5.

D, H *Perturbation of oscillations* $u_a$ = 0.01, $u_d$ = 2, $u_{beta}$ = 0.001, $d_a$ = 200, $d_d$ = 1, $k_1$ = 1, $k_2$ = 10, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 40, $k_{b2}$ = 5, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 0, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 2, $X_{Total}$ = 2, $k_{Ap-Act}$ = 0,0.001,0.005. $k_{Ap-InAct}$ = 0,0,1,0.25.

**Figure 4.5:** ***Impact of incoherent feedforward network regulation on the double site modification system***

A *Perturbation of monostable response*
$u_a$ = 1, $u_d$ = 1, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 0.01, $k_1$ = 1, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 1, $k_{iFFW}$ = 0,0.1,0.25,0.5,1,2,5,10.

B *Inducing bistability*
$u_a$ = 1, $u_d$ = 4, $u_{beta}$ = 0.01, $d_d$ = 2, $d_{beta}$ = 0.01, $k_1$ = 1, $k_2$ = 1, $k_3$ = 0.4, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1.5, $K_{Total}$ = 1, $k_{iFFW}$ = 0,0.1,0.25,0.5,1,2.

C *Inducing combination of bistability and biphasic behavior*
$u_a$ = 1, $u_d$ = 8, $u_{beta}$ = 0.01, $d_d$ = 4, $d_{beta}$ = 0.01, $k_1$ = 0.1, $k_2$ = 1, $k_3$ = 1, $k_4$ = 1, $k_{b1}$ = 40, $k_{b2}$ = 10, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1.1, $K_{Total}$ = 1, $k_{iFFW}$ = 0,0,05,0.1,0.25,1.

D *Incoherent feedforward expanding and suppressing range of oscillations*
$u_a$ = 1, $u_d$ = 5, $u_{beta}$ = 0.01, $d_d$ = 2, $d_{beta}$ = 0.01, $k_1$ = 0.8, $k_2$ = 1, $k_3$ = 1, $k_4$ = 40, $k_{b1}$ = 10, $k_{b2}$ = 1, $k_{b3}$ = 100, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 20, $K_{Total}$ = 15, $k_{iFFW}$ = 0,0.01,0.05

E *Incoherent feedforward creating distinct zones of oscillations*
$u_a$ = 1, $u_d$ = 1, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 0.01, $k_1$ = 0.8, $k_2$ = 2, $k_3$ = 1, $k_4$ = 10, $k_{b1}$ = 3, $k_{b2}$ = 1, $k_{b3}$ = 4, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 10, $K_{Total}$ = 8, $k_{iFFW}$ = 0,0.1,0.15.

F *Perturbing bistable response leads to combination of biphasic and bistable behavior*
$u_a$ = 1, $u_d$ = 3, $u_{beta}$ = 0.01, $d_d$ = 75, $d_{beta}$ = 1, $k_1$ = 0.01, $k_2$ = 1, $k_3$ = 0.01, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 10, $k_{b3}$ = 25, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 5, $K_{Total}$ = 2, $k_{iFFW}$ = 0,5,25,40.

G *Disconnected bistable steady states (Isola)*
$u_a$ = 3, $u_d$ = 4, $u_{beta}$ = 0.01, $d_d$ = 100, $d_{beta}$ = 1, $k_1$ = 0.01, $k_2$ = 1, $k_3$ = 0.05, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 5, $K_{Total}$ = 2, $k_{iFFW}$ = 0,0.25.

**Figure 4.6:** ***Impact of coherent feedforward network regulation on the double site modification system***

A *Perturbation of monostable response*
$u_a$ = 1, $u_d$ = 1, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 0.01, $k_1$ = 1, $k_2$ = 0.1, $k_3$ = 1.5, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 1, $k_{cFFW}$ = 0,0.1,0.25,1,2.

B *Perturbation of biphasic response*
$u_a$ = 1, $u_d$ = 10, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 1, $k_1$ = 1, $k_2$ = 0.1, $k_3$ = 100, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 1, $K_{Total}$ = 70, $k_{cFFW}$ = 0,0.1,0.25,0.5,1.

C *Perturbation of bistable response*
$u_a$ = 1, $u_d$ = 3, $u_{beta}$ = 0.01, $d_d$ = 0.75, $d_{beta}$ = 0.01, $k_1$ = 0.01, $k_2$ = 1, $k_3$ = 0.01, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 10, $k_{b3}$ = 25, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 5, $K_{Total}$ = 2, $k_{cFFW}$ = 0,5,25.

D *Perturbation of oscillations*
$u_a$ = 1, $u_d$ = 5, $u_{beta}$ = 0.01, $d_d$ = 2, $d_{beta}$ = 0.01, $k_1$ = 0.8, $k_2$ = 1, $k_3$ = 1, $k_4$ = 40, $k_{b1}$ = 10, $k_{b2}$ = 1, $k_{b3}$ = 100, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 20, $K_{Total}$ = 15, $k_{cFFW}$ = 0,50,250.

E *Transition: Monostable to oscillatory response*
$u_a$ = 1, $u_d$ = 1, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 0.01, $k_1$ = 0.1, $k_2$ = 0.5, $k_3$ = 1, $k_4$ = 20, $k_{b1}$ = 5, $k_{b2}$ = 1, $k_{b3}$ = 25, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 15, $K_{Total}$ = 15, $k_{cFFW}$ = 0,1.

F *Transition: Monostable to bistable response*
$u_a$ = 1, $u_d$ = 5, $u_{beta}$ = 0.01, $d_d$ = 10, $d_{beta}$ = 1, $k_1$ = 0.01, $k_2$ = 1, $k_3$ = 0.05, $k_4$ = 1, $k_{b1}$ = 2, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 5, $K_{Total}$ = 2, $k_{cFFW}$ = 0,2.

G *Transition: Bistable to oscillatory response*
$u_a$ = 1, $u_d$ = 5, $u_{beta}$ = 0.01, $d_d$ = 0.1, $d_{beta}$ = 0.01, $k_1$ = 0.1, $k_2$ = 1, $k_3$ = 0.9, $k_4$ = 100, $k_{b1}$ = 2, $k_{b2}$ = 1, $k_{b3}$ = 100, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 3, $K_{Total}$ = 7, $k_{cFFW}$ = 0,5.

**Figure 4.7:** ***Negative feedback from $A_{pp}$ can enable a combination of multistability and oscillations*** $u_a$ = 0.1, $u_d$ = 10, $u_{beta}$ = 0.001, $d_a$ = 1, $d_d$ = 0.75, $k_1$ = 4, $k_2$ = 0.5, $k_3$ = 0.25, $k_4$ = 75, $k_{b1}$ = 0.5, $k_{b2}$ = 2, $k_{b3}$ = 150, $k_{b4}$ = 5, $k_{ub1}$ = 0.5, $k_{ub2}$ = 0.5, $k_{ub3}$ = 0.5, $k_{ub4}$ = 0, $A_{Total}$ = 200, $P_{Total}$ = 5, $K_{Total}$ = 10, $X_{Total}$ = 10, $k_{NFB}$ = 40.

**Figure 4.8:** ***Incoherent feedforward regulation can generate multiple homeostasis capabilities in substrate through both simple network regulation and by exploiting and tuning intrinsic substrate modification capabilities such as symmetry breaking*** $u_a$ = 1, $u_d$ = 1, $u_{beta}$ = 0.01, $d_d$ = 1, $d_{beta}$ = 0.01, $k_1$ = 0.6, $k_2$ = 1, $k_3$ = 0.6, $k_4$ = 1, $k_{b1}$ = 1, $k_{b2}$ = 1, $k_{b3}$ = 1, $k_{b4}$ = 1, $k_{ub1}$ = 1, $k_{ub2}$ = 1, $k_{ub3}$ = 1, $k_{ub4}$ = 1, $A_{Total}$ = 40, $P_{Total}$ = 5, $K_{Total}$ = 5,
A. $k_{iFFW}$ = 0.8,1,1.2 B. S = 2

## A.4 Multi-level complexity of cells and populations

All figures have the same underlying parameters unless specified otherwise. For circuit motif definition, unless otherwise specified all parameters and total amounts are 1.

### Single population study

- **Interactions**

  (1) : $k_1 = 1$; $K_1 = 0.05$; (2) : $k_2 = 1$; $K_2 = 0.05$; (3) : $k_3 = 1$; $K_3 = 0.05$;
  (4) : $k_4 = 1$; $K_4 = 0.05$; (5) : $k_5 = 1$; $K_5 = 0.05$; (6) : $k_6 = 1$; $K_6 = 0.05$;
  (7) : $k_7 = 1$; $K_7 = 0.05$; (8) : $k_8 = 1$; $K_8 = 0.05$;

- **Population level**

  $kg_{n1} = 0.1$; $kd_{n1} = 0.002$; $ka_{e1} = 1$; $kd_{e1} = 1$; $kpr_{e1} = 0$; $kd_{e1ext} = 1$;

- **Cellular level**

  See circuit motif parameters in chapter appendix

- **Environmental level**

  $D = 5$; $V_1 = 1$; $V_{env} = 2000$;

### Two population study

- **Interactions**

  (1) : $k_1 = 1$; $K_1 = 0.05$; (2) : $k_2 = 1$; $K_2 = 0.05$; (3) : $k_3 = 1$; $K_3 = 0.05$;
  (4) : $k_4 = 1$; $K_4 = 0.05$; (5) : $k_5 = 1$; $K_5 = 0.05$; (6) : $k_6 = 1$; $K_6 = 0.05$;
  (7) : $k_7 = 1$; $K_7 = 0.05$; (8) : $k_8 = 1$; $K_8 = 0.05$; (9) : $k_9 = 1$; $K_9 = 0.05$;
  (10) : $k_{10} = 1$; $K_{10} = 0.05$; (11) : $k_{11} = 1$; $K_{11} = 0.05$; (12) : $k_{12} = 1$; $K_{12} = 0.05$;

- **Population level**

  *Population 1*: $kg_{n1} = 0.1$; $kd_{n1} = 0.002$; $ka_{e1} = 1$; $kd_{e1} = 1$; $kpr_{e1} = 0$; $kd_{e1ext} = 1$;
  *Population 2*: $kg_{n2} = 0.1$; $kd_{n2} = 0.002$; $ka_{e2} = 1$; $kd_{e2} = 1$; $kpr_{e2} = 0$; $kd_{e2ext} = 1$;

- **Cellular level**

  *Cell type 1*: See circuit motif parameters in chapter appendix
  *Cell type 2*: See circuit motif parameters in chapter appendix

- **Environmental level**

  $D = 5$; $V_1 = 1$; $V_{env} = 2000$;

## Complex population dynamics (predator-prey study)

- **Interactions**

  (1) : $k_1 = 1$; $K_1 = 5$; (2) : $k_2 = 1$; $K_2 = 0.05$;

  (3) : $k_3 = 1$; $K_3 = 15$; (4) : $k_4 = 1$; $K_4 = 0.05$;

- **Population level**

  *Population 1* (predator): $\alpha = 0.055$; $\beta = 0.02$;

  *Population 2* (prey): $\delta = 0.005$; $\gamma = 0.02$;

- **Cellular level**

  *PFB*: $S_{basal} = 6$; *NFB*: $S_{basal} = 4$; *Cov*: $S_{basal} = 0.5$;

  *Cell type 1*: See circuit motif parameters in chapter appendix

  *Cell type 2*: See circuit motif parameters in chapter appendix

- **Environmental level**

  $D = 5$; $V_1 = 1$; $V_{env} = 100$;

## Circuit motif parameters

*Cov* Switch-like behavior
$k_1 = 1$; $k_2 = 1$; $R_p + c = 1$, $K_{m1} = 0.05$; $K_{m2} = 0.05$;

*PFB* Irreversible bistability
$k_0 = 0.4$; $k_1 = 0.01$; $k_2 = 1$; $k_3 = 1$; $k_4 = 0.2$; $K_3 = 0.05$; $K_4 = 0.05$;
$E + E_p = 1$;

*NFB* Homeostasis
$k_0 = 1$; $k_2 = 1$; $k_3 = 0.5$; $k_4 = 1$; $K_3 = 0.01$; $K_4 = 0.01$; $E + E_p = 1$;

*NFB* Oscillations
$k_0$=0, $k_1$=1, $k_2$=0.01, $k_{21}$=10, $k_3$=0.1, $k_4$=0.2, $k_5$=0.1, $k_6$=0.05,
$Y + Y_p = 1$; $R + R_p$=1; $K_3 = 0.01$; $K_4 = 0.01$; $K_5 = 0.01$; $K_6 = 0.01$;

*Rlx* Relaxation oscillations
$k_{01} = 0.01$; $k_0 = 0.4$; $k_1 = 1$; $k_2 = 1$; $k_3 = 1$; $k_4 = 0.3$; $K_3 = 0.05$; $K_4 = 0.05$; $E + E_p = 1$;

*PFB* Reversible bistability
$k_0 = 0$; $k_1 = 0.05$; $k_2 = 0.1$; $k_{21} = 0.5$; $k_3 = 0.2$; $k_4 = 1$; $K_3 = 0.05$; $K_4 = 0.05$; $E + E_p = 1$;

**Figure 5.27 - Population dynamics can enable reversible switching in circuit motif response (MI motif)**
$\alpha = 0.0055$; $\beta = 0.004$; $\delta = 0.001$; $\gamma = 0.004$; $k_1 = 5$; $K_1 = 10$; $k_3 = 5$; $K_3 = 15$; (Rest unchanged)

# Appendix B

# Mathematical models and analytical results

## B.1 Symmetry and symmetry breaking in multisite modification

We now present the kinetic descriptions of our modification networks. In all cases we employ ODE based kinetic descriptions as these focus on the central aspects of interest (the modifications and their sequence) for the purposes of our study. Spatial and stochastic aspects can be studied as extensions of this.

Consider enzyme action on a protein substrate with two modification sites and a common enzyme acting on both sites. The mechanism of modification could be either distributive or processive manner as described below:

**Enzyme action : Double site modification (Processive)**

| **Substrate** | + | **Enzyme** | $\rightleftharpoons$ | **Complex** | $\rightleftharpoons$ | **Mod. Com.1** | $\rightarrow$ | **Mod. Sub.** | + | **Enzyme** |
|---|---|---|---|---|---|---|---|---|---|---|
| $[A_{00}]$ | | [K] | | $[A_{00}K]$ | | $[A_{01}K]$ | | $[A_{11}]$ | | [K] |
| | | | | | OR | | | | | |
| $[A_{00}]$ | | [K] | | $[A_{10}K]$ | | $[A_{10}K]$ | | $[A_{11}]$ | | [K] |

**Enzyme action : Double site modification (Distributive)**

| **Substrate** | + | **Enzyme** | $\rightleftharpoons$ | **Complex1** | $\rightarrow$ | **Mod. Sub.** | + | **Enzyme** |
|---|---|---|---|---|---|---|---|---|
| $[A_{00}]$ | | [K] | | $[(A_{00}K)_2]$ | | $[A_{01}]$ | | [K] |
| $[A_{01}]$ | | [K] | | $[A_{01}K]$ | | $[A_{11}]$ | | [K] |
| | | | | | OR | | | |
| $[A_{00}]$ | | [K] | | $[(A_{00}K)_2]$ | | $[A_{10}]$ | | [K] |
| $[A_{10}]$ | | [K] | | $[A_{10}K]$ | | $[A_{11}]$ | | [K] |

In each case (processive and distributive) the enzyme (in the context of a phosphorylation & dephosphorylation - kinase & phosphatase respectively) binds reversibly to the

protein substrate to form a enzyme substrate complex. If the enzyme action is **processive**, the enzyme is bound to the substrate (in multiple complexes) until all modifications are effected, after which it irreversibly detaches to give the completely modified protein and the free enzyme. In the context of a random double site modification as shown, there are two possible ways in which the modifications can be (processively) effected, corresponding to the ordering in modifications i.e. first site being modified followed by the second site and vice versa. If the enzyme action is **distributive**, the enzyme detaches from the substrate after effecting each modification before reversibly binding again with the partially modified substrate to form a new complex corresponding to the additional modification. Similar to the processive enzyme action, there are two pathways (corresponding to different ordering) in which modifications can be effected in the double site module as shown, each leading to a unique partially modified substrate. Combinations of processive and distributive enzyme action are also possible. Furthermore, additional variations in multisite modifications are possible depending on enzyme specificity. In the example above, we considered a case where both modification sites were effected by a common enzyme (kinase/phosphatase). However,(de)modification of each modification site can be effected by a different enzyme, and thus the possibility of having common/separate enzymes performing different modifications represent a suite of basic modification scenarios.

These complexities (for the double site modification of substrate) are captured in the models depicted in Figure 2.1. The networks are modelled as a system of ODEs generated by describing the individual elementary reactions (as shown above in Figure 2.1 and Figure 2.14) using generic mass action kinetic descriptions. This makes no assumptions on the kinetic regime of modification.

The nomenclature used for modelling is as follows. The kinetic constants of the binding and unbinding reactions are denoted by the letters $k_{bi}$ and $k_{ubi}$ while that of the irreversible catalytic reaction of the complex is denoted by $k_i$; where $i$ is an index which stands for the reaction number in a given model. For the sake of clarity and brevity, models represented in the figures in the main text (refer Figure 2.1) use $\kappa_i$ and $\alpha_i$ to concisely represent binding, unbinding and catalytic reaction rate constants involved in a particular modification. An individual modification/demodification step involves the enzyme reversibly binding to the substrate to form a complex, which then involves an irreversible catalytic reaction to produce the modified form (and release the enzyme): this represents a widely employed model of substrate modification. The detailed model description for the networks considered in chapter 2 are constructed by employing such mass action kinetic descriptions of the elementary reactions for all modification/demodification steps (refer to fig. 2.14 for detailed schematic denoting independent reactions and their kinetic nomenclature). All models are presented in dimensionless form.

### Ordered Distributive Multi Site Phosphorylation

**DSP - Common Kinase Common Phosphatase**: The ordered distributive DSP with common kinase common phosphatase acting on both modification sites (refer Figure 2.1C) is realized when the order of phosphorylation and order of dephosphorylation are opposite. This results in a single partially modified (phosphorylated) substrate. In the model description $[A]$ represents the substrate while its subscript designation represents the number of modifications effected, i.e. $[A_p]$ represents the partially modified substrate and $[A_{pp}]$ represents the completely modified substrate. As shown in the schematic, the substrate ($[A]$) first reversibly binds to the kinase ($[K]$) to form an enzyme-substrate complex ($[AK]$). The modification is then effected with the complex irreversibly producing the modified substrate ($[A_p]$) and the free enzyme ($[K]$). Phosphorylation and dephosphorylation proceeds in a distributive manner at both modification sites. Using mass action kinetics and the kinetic nomenclature described earlier, we can model the system as a set of ODEs as follows.

$$\begin{aligned}
\frac{d[A]}{dt} &= -k_{b1}[A][K] + k_{ub1}[AK] + k_4[A_pP] \\
\frac{d[A_p]}{dt} &= -k_{b2}[A_p][K] - k_{b4}[A_p][P] + k_1[AK] + k_{ub2}[A_pK] + k_{ub4}[A_pP] + k_3[A_{pp}P] \\
\frac{d[A_{pp}]}{dt} &= -k_{b3}[A_{pp}][P] + k_2[A_pK] + k_{ub3}[A_{pp}P] \\
\frac{d[AK]}{dt} &= k_{b1}[A][K] - (k_{ub1} + k_1)[AK] \\
\frac{d[A_pK]}{dt} &= k_{b2}[A_p][K] - (k_{ub2} + k_2)[A_pK] \\
\frac{d[A_{pp}P]}{dt} &= k_{b3}[A_{pp}][P] - (k_{ub3} + k_3)[A_{pp}P] \\
\frac{d[A_pP]}{dt} &= k_{b4}[A_p][P] - (k_{ub4} + k_4)[A_pP] \\
\frac{d[K]}{dt} &= -k_{b1}[A][K] + (k_{ub1} + k_1)[AK] - k_{b2}[A_p][K] + (k_{ub2} + k_2)[A_pK] \\
\frac{d[P]}{dt} &= -k_{b3}[A_{pp}][P] + (k_{ub3} + k_3)[A_{pp}P] - k_{b4}[A_p][P] + (k_{ub4} + k_4)[A_pP]
\end{aligned} \tag{B.1}$$

The total substrate and enzyme concentrations are also conserved in this system. This is represented mathematically as follows:

$$\begin{aligned}
A_{Total} &= [A] + [A_p] + [A_{pp}] + [AK] + [A_pK] + [A_{pp}P] + [A_pP] \\
K_{Total} &= [K] + [A_pK] + [AK] \\
P_{Total} &= [P] + [A_{pp}P] + [A_pP]
\end{aligned} \tag{B.2}$$

**Triple Site Phosphorylation (TSP) - Common Kinase Common Phosphatase**: While double site phosphorylation networks are the primary focal point of the study, we use the ordered distributive TSP to make some specific points. The ordered distributive TSP with common kinase and common phosphatase acting on all modification sites is realized when the order of phosphorylation and dephosphorylation are reversed (see fig. 2.12). The network is modelled as a system of ODEs using mass action kinetic description of the elementary reactions involved. Similar to the ordered distributive DSP, the substrate is represented by $[A]$, while the subscript represents the number of phosphorylated modification sites. The equations are presented in appendix C.

**Random Double site modification**

The random DSP network is realized when there is no preferential ordering for either the modification / demodification (phosphorylation / dephosphorylation). In chapter 2 we have considered multiple variations of random DSP, depending on whether the enzymes (kinases/phosphatases) effecting the modifications are the same or different (namely system 1-3, refer Figure 2.1).

The general nomenclature used in these models is as follows. The substrate is denoted by the letter $A$, while the subscript with binary digits is used to denote the nature of the specific modification sites. '1' represents a phosphorylated modification site, while '0' represents that the specific site is unphosphorylated. Thus $[A_{00}]$ and $[A_{11}]$, represents the completely unmodified and completely modified substrates, while $[A_{01}]$ and $[A_{10}]$ represent partially modified substrates with only the second and first site modified respectively. Where required, the enzyme-substrate complexes are further designated with subscript numbers (outside curved brackets) to differentiate between complexes where the enzymes are associated with different modification (eg.. $[(A_{00}K)_1]$ and $[(A_{00}K)_2]$).

The ODE models are constructed in a similar manner to those of the ordered mechanism, by combining mass action kinetic descriptions of the elementary steps. For a detailed schematic including kinetic representation of the elementary binding and unbinding reactions, please refer to Figure 2.14.

**System 1 - Common Kinase Common Phosphatase**: System 1 has a single kinase and phosphatase that effects phosphorylation and dephosphorylation respectively on both the modification sites. Using the above nomenclature this system can be represented

as a system of ODEs as follows.

$$
\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K] - a_{b1}[A_{00}][K] + k_{ub1}[(A_{00}K)_1] + a_{ub1}[A_{00}K_2] + a_4[A_{10}P] + k_4[A_{01}P] \\
\frac{d[A_{10}]}{dt} &= -a_{b2}[A_{10}][K] - a_{b4}[A_{10}][P] + a_1[(A_{00}K)_2] + a_{ub2}[A_{10}K] + a_{ub4}[A_{10}P] + a_3[(A_{11}P)_2] \\
\frac{d[A_{01}]}{dt} &= -k_{b2}[A_{01}][K] - k_{b4}[A_{01}][P] + k_1[(A_{00}K)_1] + k_{ub2}[A_{01}K] + k_{ub4}[A_{01}P] + k_3[(A_{11}P)_1] \\
\frac{d[A_{11}]}{dt} &= -k_{b3}[A_{11}][P] - a_{b3}[A_{11}][P] + a_2[A_{10}K] + k_2[A_{01}K] + k_{ub3}[(A_{11}P)_1] + a_{ub3}[(A_{11}P)_2] \\
\frac{d[(A_{00}K)_1]}{dt} &= k_{b1}[A_{00}][K] - (k_{ub1} + k_1)[(A_{00}K)_1] \\
\frac{d[A_{01}K]}{dt} &= k_{b2}[A_{01}][K] - (k_{ub2} + k_2)[A_{01}K] \\
\frac{d[(A_{00}K)_2]}{dt} &= a_{b1}[A_{00}][K] - (a_{ub1} + a_1)[(A_{00}K)_2] \\
\frac{d[A_{10}K]}{dt} &= a_{b2}[A_{10}][K] - (a_{ub2} + a_2)[A_{10}K] \\
\frac{d[(A_{11}P)_1]}{dt} &= k_{b3}[A_{11}][P] - (k_{ub3} + k_3)[(A_{11}P)_1] \\
\frac{d[A_{01}P]}{dt} &= k_{b4}[A_{01}][P] - (k_{ub4} + k_4)[A_{01}P] \\
\frac{d[(A_{11}P)_2]}{dt} &= a_{b3}[A_{11}][P] - (a_{ub3} + a_3)[(A_{11}P)_2] \\
\frac{d[A_{10}P]}{dt} &= a_{b4}[A_{10}][P] - (a_{ub4} + a_4)[A_{10}P] \\
\frac{d[K]}{dt} &= -k_{b1}[A_{00}][K] + (k_1 + k_{ub1})[(A_{00}K)_1] - a_{b1}[A_{00}][K] \\
&\qquad + (a_{ub1} + a_1)[(A_{00}K)_2] - a_{b2}[A_{10}][K] + (a_{ub2} + a_2)[A_{10}K] \\
&\qquad - k_{b2}[A_{01}][K] + (k_2 + k_{ub2})[A_{01}K] \\
\frac{d[P]}{dt} &= -k_{b3}[A_{11}][P] + (k_{ub3} + k_3)[(A_{11}P)_1] - a_{b3}[A_{11}][P] \\
&\qquad + (a_{ub3} + a_3)[(A_{11}P)_2] - a_{b4}[A_{10}][P] + (a_{ub4} + a_4)[A_{10}P] \\
&\qquad - k_{b4}[A_{01}][P] + (k_{ub4} + k_4)[A_{01}P]
\end{aligned}
\tag{B.3}
$$

The total substrate and enzyme concentrations are conserved in the system. This is represented mathematically as follows.

$$
\begin{aligned}
A_{Total} &= [A_{00}] + [A_{10}] + [A_{01}] + [A_{11}] + [(A_{00}K)_1] + [A_{01}K] + [(A_{00}K)_2] \\
&\qquad + [A_{10}K] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P] \\
K_{Total} &= [K] + [(A_{00}K)_1] + [A_{10}K] + [(A_{00}K)_2] + [A_{01}K] \\
P_{Total} &= [P] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P]
\end{aligned}
\tag{B.4}
$$

**System 2 - Separate Kinase Common Phosphatase**: System 2 has a distinct kinase acting on each modification site, while a common phosphatase effects dephosphorylation on both sites. Using the nomenclature discussed earlier, the system is modelled as a set

of ODEs as shown below.

$$
\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K_1] - a_{b1}[A_{00}][K_2] + k_{ub1}[A_{00}K_1] + a_{ub1}[A_{00}K_2] + k_4[A_{01}P] + a_4[A_{10}P] \\
\frac{d[A_{01}]}{dt} &= -k_{b2}[A_{01}][K_2] - k_{b4}[A_{01}][P] + k_1[A_{00}K_1] + k_{ub2}[A_{01}K_2] + k_{ub4}[A_{01}P] + k_3[(A_{11}P)_1] \\
\frac{d[A_{10}]}{dt} &= -a_{b2}[A_{10}][K_1] - a_{b4}[A_{10}][P] + a_1[A_{00}K_2] + a_{ub2}[A_{10}K_1] + a_{ub4}[A_{10}P] + a_3[(A_{11}P)_2] \\
\frac{d[A_{11}]}{dt} &= -k_{b3}[A_{11}][P] - a_{b3}[A_{11}][P] + k_2[A_{01}K_2] + a_2[A_{10}K_1] + k_{ub3}[(A_{11}P)_1] + a_{ub3}[(A_{11}P)_2] \\
\frac{d[A_{00}K_1]}{dt} &= k_{b1}[A_{00}][K_1] - (k_1 + k_{ub1})[A_{00}K_1] \\
\frac{d[A_{10}K_1]}{dt} &= a_{b2}[A_{10}][K_1] - (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[A_{00}K_2]}{dt} &= a_{b1}[A_{00}][K_2] - (a_{ub1} + a_1)[A_{00}K_2] \\
\frac{d[A_{01}K_2]}{dt} &= k_{b2}[A_{01}][K_2] - (k_{ub2} + k_2)[A_{01}K_2] \\
\frac{d[(A_{11}P)_1]}{dt} &= k_{b3}[A_{11}][P] - (k_{ub3} + k_3)[(A_{11}P)_1] \\
\frac{d[A_{10}P]}{dt} &= a_{b4}[A_{10}][P] - (a_{ub4} + a_4)[A_{10}P] \\
\frac{d[(A_{11}P)_2]}{dt} &= a_{b3}[A_{11}][P] - (a_{ub3} + a_3)[(A_{11}P)_2] \\
\frac{d[A_{01}P]}{dt} &= k_{b4}[A_{01}][P] - (k_{ub4} + k_4)[A_{01}P] \\
\frac{d[K_1]}{dt} &= -k_{b1}[A_{00}][K_1] + (k_1 + k_{ub1})[A_{00}K_1] - a_{b2}[A_{10}][K_1] + (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[K_2]}{dt} &= -a_{b1}[A_{00}][K_2] + (a_{ub1} + a_1)[A_{00}K_2] - k_{b2}[A_{01}][K_2] + (k_{ub2} + k_2)[A_{01}K_2] \\
\frac{d[P]}{dt} &= -k_{b3}[A_{11}][P] + (k_{ub3} + k_3)[(A_{11}P)_1] - a_{b4}[A_{10}][P] + (a_{ub4} + a_4)[A_{10}P] \\
&\quad - a_{b3}[A_{11}][P] + (a_{ub3} + a_3)[(A_{11}P)_2] - k_{b4}[A_{01}][P] + (k_{ub4} + k_4)[A_{01}P]
\end{aligned}
\tag{B.5}
$$

The total substrate and enzyme concentrations are conserved in this system. Note that each distinct kinase is associated with a conservation condition.

$$
\begin{aligned}
A_{Total} &= [A_{00}] + [A_{10}] + [A_{01}] + [A_{11}] + [A_{00}K_1] + [A_{01}K_2] + [A_{00}K_2] \\
&\quad + [A_{10}K_1] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P] \\
K1_{Total} &= [K_1] + [A_{00}K_1] + [A_{10}K_1] \\
K2_{Total} &= [K_2] + [A_{00}K_2] + [A_{01}K_2] \\
P_{Total} &= [P] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P]
\end{aligned}
\tag{B.6}
$$

**System 3 - Separate Kinase Separate Phosphatase**: System 3 has distinct kinases and phosphatases effecting phosphorylation & dephosphorylation on each modification

site. The system is modelled as a set of ODEs as shown below.

$$\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K_1] - a_{b1}[A_{00}][K_2] + k_{ub1}[A_{00}K_1] + a_{ub1}[A_{00}K_2] + k_4[A_{01}P_1] + a_4[A_{10}P_2] \\
\frac{d[A_{01}]}{dt} &= -k_{b2}[A_{01}][K_2] - k_{b4}[A_{01}][P_1] + k_1[A_{00}K_1] + k_{ub2}[A_{01}K_2] + k_{ub4}[A_{01}P_1] + k_3[A_{11}P_2] \\
\frac{d[A_{10}]}{dt} &= -a_{b2}[A_{10}][K_1] - a_{b4}[A_{10}][P_2] + a_1[A_{00}K_2] + a_{ub2}[A_{10}K_1] + a_{ub4}[A_{10}P_2] + a_3[(A_{11}P_1] \\
\frac{d[A_{11}]}{dt} &= -a_{b3}[A_{11}][P_1] - k_{b3}[A_{11}][P_2] + k_2[A_{01}K_2] + a_2[A_{10}K_1] + a_{ub3}[A_{11}P_1] + k_{ub3}[A_{11}P_2] \\
\frac{d[A_{00}K_1]}{dt} &= k_{b1}[A_{00}][K_1] - (k_1 + k_{ub1})[A_{00}K_1] \\
\frac{d[A_{01}K_2]}{dt} &= k_{b2}[A_{01}][K_2] - (k_2 + k_{ub2})[A_{01}K_2] \\
\frac{d[A_{11}P_2]}{dt} &= k_{b3}[A_{11}][P_2] - (k_3 + k_{ub3})[A_{11}P_2] \\
\frac{d[A_{01}P_1]}{dt} &= k_{b4}[A_{01}][P_1] - (k_4 + k_{ub4})[A_{01}P_1] \\
\frac{d[A_{00}K_2]}{dt} &= a_{b1}[A_{00}][K_2] - (a_1 + a_{ub1})[A_{00}K_2] \\
\frac{d[A_{10}K_1]}{dt} &= a_{b2}[A_{10}][K_1] - (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[A_{11}P_1]}{dt} &= a_{b3}[A_{11}][P_1] - (a_3 + a_{ub3})[A_{11}P_1] \\
\frac{d[A_{10}P_2]}{dt} &= a_{b4}[A_{10}][P_2] - (a_4 + a_{ub4})[A_{10}P_2] \\
\frac{d[K_1]}{dt} &= k_{b1}[A_{00}][K_1] - (k_1 + k_{ub1})[A_{00}K_1] + a_{b2}[A_{10}][K_1] - (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[K_2]}{dt} &= a_{b1}[A_{00}][K_2] - (a_1 + a_{ub1})[A_{00}K_2] + k_{b2}[A_{01}][K_2] - (k_2 + k_{ub2})[A_{01}K_2] \\
\frac{d[P_1]}{dt} &= a_{b3}[A_{11}][P_1] - (a_3 + a_{ub3})[A_{11}P_1] + k_{b4}[A_{01}][P_1] - (k_4 + k_{ub4})[A_{01}P_1] \\
\frac{d[P_2]}{dt} &= k_{b3}[A_{11}][P_2] - (k_3 + k_{ub3})[A_{11}P_2] + a_{b4}[A_{10}][P_2] - (a_4 + a_{ub4})[A_{10}P_2]
\end{aligned} \tag{B.7}$$

The conservation condition for the substrate and each of the enzymes is shown below.

$$\begin{aligned}
A_{Total} &= [A_{00}] + [A_{10}] + [A_{01}] + [A_{11}] + [A_{00}K_1] + [A_{01}K_2] + [A_{11}P_2] + [A_{01}P_1] \\
&\qquad + [A_{00}K_2] + [A_{10}K_1] + [A_{11}P_1] + [A_{10}P_2] \\
K1_{Total} &= [A_{00}K_1] + [A_{10}K_1] + [K_1] \\
K2_{Total} &= [A_{01}K_2] + [A_{00}K_2] + [K_2] \\
P1_{Total} &= [A_{11}P_1] + [A_{01}P_1] + [P_1] \\
P2_{Total} &= [A_{10}P_2] + [A_{11}P_2] + [P_2]
\end{aligned} \tag{B.8}$$

### Mixed-Random DSP

The 'Mixed-Random' DSP which we study involves a combination of distributive phosphorylation and processive dephosphorylation, without any ordering. In chapter 2, multiple such mixed random modification networks (Mixed-Random 1,2 and 3 - refer Figure 2.1 and Figure 2.5) are analyzed from a specific viewpoint: their individual capacity to exhibit case 2 symmetry breaking. The goal there is to establish whether having processive modification results in any new insights. The model description for these systems is as follows.

**Mixed-Random 1 - Common Kinase Common Phosphatase**: Mixed-Random 1 has a common kinase that effects phosphorylation on both the modification sites. Using the

above nomenclature this system can be represented as a system of ODEs as follows.

$$
\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K] - a_{b1}[A_{00}][K] + k_{ub1}[A_{00}K_1] + a_{ub1}[A_{00}K_2] + k_4[A_{01}P] + a_4[A_{10}P] \\
\frac{d[A_{01}]}{dt} &= -k_{b2}[A_{01}][K] + k_1[(A_{00}K)_1] + k_{ub2}[A_{01}K] \\
\frac{d[A_{10}]}{dt} &= -a_{b2}[A_{10}][K] + a_1[(A_{00}K)_2] + a_{ub2}[A_{10}K] \\
\frac{d[A_{11}]}{dt} &= -k_{b3}[A_{11}][P] - a_{b3}[A_{11}][P] + k_2[A_{01}K] + a_2[A_{10}K] + k_{ub3}[(A_{11}P)_1] + a_{ub3}[A_{11}P_2] \\
\frac{d[(A_{00}K)_1]}{dt} &= k_{b1}[A_{00}][K] - (k_1 + k_{ub1})[(A_{00}K)_1] \\
\frac{d[A_{10}K]}{dt} &= a_{b2}[A_{10}][K] - (a_2 + a_{ub2})[A_{10}K] \\
\frac{d[(A_{00}K)_2]}{dt} &= a_{b1}[A_{00}][K] - (a_{ub1} + a_1)[(A_{00}K)_2] \\
\frac{d[A_{01}K]}{dt} &= k_{b2}[A_{01}][K] - (k_{ub2} + k_2)[A_{01}K] \\
\frac{d[(A_{11}P)_1]}{dt} &= k_{b3}[A_{11}]P - (k_{ub3} + k_3)[(A_{11}P)_1] \\
\frac{d[A_{10}P]}{dt} &= -[A_{10}P]a_4 + [A_{11}P_2]a_3 \\
\frac{d[A_{11}P_2]}{dt} &= a_{b3}[A_{11}][P] - (a_{ub3} + a_3)[A_{11}P_2] \\
\frac{d[A_{01}P]}{dt} &= -[A_{01}P]k_4 + [(A_{11}P)_1]k_3 \\
\frac{d[K]}{dt} &= -k_{b1}[A_{00}][K] + (k_1 + k_{ub1})[(A_{00}K)_1] - a_{b2}[A_{10}][K] + (a_2 + a_{ub2})[A_{10}K] - a_{b1}[A_{00}][K] \\
&\qquad + (a_{ub1} + a_1)[(A_{00}K)_2] - k_{b2}[A_{01}][K] + (k_{ub2} + k_2)[A_{01}K] \\
\frac{d[P]}{dt} &= -k_{b3}[A_{11}][P] + (k_{ub3} + k_3)[(A_{11}P)_1] + a_4[A_{10}P] - a_3[A_{11}P_2] \\
&\qquad - a_{b3}[A_{11}][P] + (a_{ub3} + a_3)[A_{11}P_2] + k_4[A_{01}P] - k_3[(A_{11}P)_1]
\end{aligned}
\tag{B.9}
$$

The total concentration of substrates and enzymes are conserved, and this is depicted below.

$$
\begin{aligned}
A_{Total} &= [A_{00}] + [A_{10}] + [A_{01}] + [A_{11}] + [A_{00}K_1] + [A_{01}K] + [A_{00}K_2] \\
&\qquad + [A_{10}K] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P] \\
K_{Total} &= [K_1] + [A_{00}K_1] + [A_{10}K] + [A_{01}K] + [(A_{00}K)_2] \\
P_{Total} &= [P] + [(A_{11}P)_1] + [A_{10}P] + [(A_{11}P)_2] + [A_{01}P]
\end{aligned}
\tag{B.10}
$$

**Mixed-Random 2 - Separate Kinase Common Phosphatase**: Mixed-Random 2: This modification network differs from the previous one in one respect: distinct kinases effect phosphorylation on each modification site, while a common phophatase performs the dephosphorylation in a processive manner, just as described above.This system can

be represented as a system of ODEs as follows,

$$\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K_1] - a_{b1}[A_{00}][K_2] + k_{ub1}[A_{00}K_1] + a_{ub1}[A_{00}K_2] + k_4[A_{01}P] + a_4[A_{10}P] \\
\frac{d[A_{01}]}{dt} &= -k_{b2}[A_{01}][K_2] + k_1[A_{00}K_1] + k_{ub2}[A_{01}K_2] \\
\frac{d[A_{10}]}{dt} &= -a_{b2}[A_{10}][K_1] + a_1[A_{00}K_2] + a_{ub2}[A_{10}K_1] \\
\frac{d[A_{11}]}{dt} &= -k_{b3}[A_{11}][P] - a_{b3}[A_{11}][P] + k_2[A_{01}K_2] + a_2[A_{10}K_1] + k_{ub3}[(A_{11}P)_1] + a_{ub3}[(A_{11}P)_2] \\
\frac{d[A_{00}K_1]}{dt} &= k_{b1}[A_{00}][K_1] - (k_1 + k_{ub1})[A_{00}K_1] \\
\frac{d[A_{10}K_1]}{dt} &= a_{b2}[A_{10}][K_1] - (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[A_{00}K_2]}{dt} &= a_{b1}[A_{00}][K_2] - (a_{ub1} + a_1)[A_{00}K_2] \\
\frac{d[A_{01}K_2]}{dt} &= k_{b2}[A_{01}][K_2] - (k_{ub2} + k_2)[A_{01}K_2] \\
\frac{d[A_{11}P_1]}{dt} &= k_{b3}[A_{11}][P] - (k_{ub3} + k_3)[(A_{11}P)_1] \\
\frac{d[A_{10}P]}{dt} &= -a_4[A_{10}P] + a_3[(A_{11}P)_2] \\
\frac{d[A_{11}P_2]}{dt} &= a_{b3}[A_{11}][P] - (a_{ub3} + a_3)[(A_{11}P)_2] \\
\frac{d[A_{01}P]}{dt} &= -k_4[A_{01}P] + k_3[(A_{11}P)_1] \\
\frac{d[K_1]}{dt} &= -k_{b1}[A_{00}][K_1] + (k_1 + k_{ub1})[A_{00}K_1] - a_{b2}[A_{10}][K_1] + (a_2 + a_{ub2})[A_{10}K_1] \\
\frac{d[K_2]}{dt} &= -a_{b1}[A_{00}][K_2] + (a_{ub1} + a_1)[A_{00}K_2] - k_{b2}[A_{01}][K_2] + (k_{ub2} + k_2)[A_{01}K_2]
\end{aligned} \tag{B.11}$$

The total concentrations of the substrate and respective enzymes are conserved as shown below.

$$\begin{aligned}
A_{Total} &= [A_{00}] + [A_{10}] + [A_{01}] + [A_{11}] + [A_{00}K_1] + [A_{01}K_2] + [A_{00}K_2] \\
&\qquad + [A_{10}K_1] + [(A_{11}P)_1] + [[A_{10}]P] + [(A_{11}P)_2] + [A_{01}P] \\
K1_{Total} &= [K_1] + [A_{00}K_1] + [A_{10}K_1] \\
K2_{Total} &= [K_2] + [A_{00}K_2] + [A_{01}K_2] \\
P_{Total} &= [P] + [(A_{11}P)_1] + [[A_{10}]P] + [(A_{11}P)_2] + [A_{01}P]
\end{aligned} \tag{B.12}$$

**Mixed-Random 3 - Separate Kinase Common Phosphatase (Unsaturated Phosphorylation)**: The model of Mixed-Random 3 is similar to that of the Mixed-Random 2 network, except that the dephosphorylation of the fully modified to unmodified form is described by a pair of linear reaction. This can happen when the catalytic constants for the dephosphorylation are significantly higher than the binding/unbinding of substrate to the phosphatase. The model is depicted below.

$$\begin{aligned}
\frac{d[A_{00}]}{dt} &= -k_{b1}[A_{00}][K_1] - a_{b1}[A_{00}][K_2] + k_{ub1}[A_{00}K_1] + a_{ub1}[A_{00}K_2] + k_3[A_{11}] + a_3[A_{11}] \\
\frac{d[A_{11}]}{dt} &= k_2[A_{01}K_2] + a_2[A_{10}K_1] - k_3[A_{11}] - a_3[A_{11}]
\end{aligned} \tag{B.13}$$

## Symmetry and symmetry breaking in multisite modification: Analytical proofs

This section contains a summary of the basic analytical arguments used to explore the feasibility of symmetry breaking and its associated features in the various MSP networks considered in this paper. This analysis is expanded on in detail in the appendix C.

The symmetry in the context of our models is established through a strict kinetic structure and enforcement of total enzyme concentrations which ensures that certain pairs of kinetic terms are equal (Refer Figure 2.1). Note that this requires, in general, the binding, unbinding and catalytic constants to be the same. This ensures that starting with symmetric initial conditions for the appropriate variables (substrates and associated enzymes and complexes), the system evolves maintaining this symmetry. In this context we point out that all three constants could affect the modification rate, and in fact studies [210, 211] show how even unbinding constants significantly affect efficacy of modification.

The equality of reaction rates results in symmetries in terms of concentrations of substrates at (one of) the steady state(s) of the system. The phenomenon of symmetry breaking allows for this symmetric state to lose stability giving rise to stable asymmetric steady states. In this section, we show the in feasibility of asymmetric states existing in some networks, thereby ruling out any symmetry breaking; in other instances we complement computational results showing symmetry breaking in other networks with analytical results, including necessary and sufficient conditions for symmetry breaking. This is done by solving the system of ODEs of the associated model at steady state and isolating asymmetric steady states therein (if they exist).

In each of the networks considered, we first isolate the substrate and enzyme pairs that share symmetry (symmetry breaking leading to asymmetric states thus naturally implies that this symmetric pairing is no longer maintained). For example, in case 1 symmetry of an ordered distributive DSP with common kinase and common phosphatase, the enzyme pairs $[K]$ & $[P]$ and substrate pairs $[A]$ & $[A_{pp}]$ are equal. Thus an asymmetric steady state is characterized by, $[K] \neq [P]$ and $[A] \neq [A_{pp}]$. This allows us to characterize the asymmetric state, and focus on that in the analysis.

Thus in this way, through a series of algebraic manipulations (to obtain reduced equations characterizing steady states, and asymmetric ones in particular) we isolate the necessary and sufficient conditions for the asymmetric state. However only a direct steady state determination is performed and the stability of these states is not discussed in the analytical work presented here. The following notation is used here on for the sake of brevity in analytical expressions,

$$c_i = \frac{k_{bi}}{k_i + k_{ubi}} \quad \text{and} \quad d_i = \frac{a_{bi}}{a_i + a_{ubi}}$$

**Ordered distributive DSP - common kinase and common phosphatase**

The ordered distributive DSP with common kinase and common phosphatase acting on both modification sites is capable of exhibiting case 1 symmetry. The kinetic constraints necessary to impose this symmetry is given in the schematic in Figure 2.2A; ($k_3 = k_1$, $k_4 = k_2$, $k_{b3} = k_{b1}$, $k_{b4} = k_{b2}$, $k_{ub3} = k_{ub1}$, $k_{ub4} = k_{ub2}$ and $P_{Total} = K_{Total}$). Under these constraints, case 1 symmetry is established resulting in equal concentrations of the substrates $[A]$ & $[A_{pp}]$ and equal concentrations of the enzymes $[K]$ & $[P]$. Imposing these constraints and evaluating the steady states of the system (by successively solving the for the steady states of the ODEs), we get the following equation involving the concentrations at steady state.

$$([K] - [P]) \left( \frac{[A_p](k_1 - k_2)c_2 + k_1}{k_1} \right) = 0 \tag{B.14}$$

From this expression, we can clearly see that the system accommodates asymmetric solutions (i.e. $[K] \neq [P]$) when the term $([A_p](k_1 - k_2)c_2 + k_1)$ is $0$. Using this information, we identify features of the asymmetric steady state. We show that at a given asymmetric state, the concentration of $[A_p]$ is fixed at a certain value (invariant), given only by key kinetic constants. Further, since the concentration of a substrate is always positive, this asymmetric state can only exist when $k_2 > k_1$, which gives us a necessary condition for the symmetry to break. The sufficiency of this condition follows by evaluating all species concentrations at this invariant value of $A_p$ and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. Thereby, we show the presence of asymmetric states for some suitable value of $A_{Total}$ (which is determined from the resulting steady state substrate concentrations). This is carried out in the appendix C.

- *Necessary and sufficient condition for symmetry breaking:*
  1. $k_2 > k_1$
- *Invariant:*
  1. $[A_p] = \frac{1}{c_2} \left[ \frac{k_1}{k_2 - k_1} \right]$

**Ordered distributive DSP - separate kinase and separate phosphatase**

The ordered distributive DSP with separate kinase and separate phosphatase acting on each modification site can present case 1 symmetry similar to the case of ordered distributive DSP with common kinase and common phosphatase acting on both modification sites. The symmetry is established with similar kinetic constraints (though now involving different pairs of enzymes; $K1_{Total} = P2_{Total}$ and $K2_{Total} = P1_{Total}$). At symmetric steady states this network can be shown to necessarily have the following symmetric enzyme pairing, $[K_1] = [P_2]$ and $[K_2] = [P_1]$ in addition to symmetry in the substrate pair $[A] = [A_{pp}]$. Through a simple algebraic analysis following solving

for steady states, we can establish that an asymmetric state violating these symmetric pairings of variables can never exist, thus ruling out symmetry breaking in the model (See Source code file (*Maple* file) for more detail).

Specifically by successively solving for the steady states of the ODEs, we can ascertain that the following equation is always true at steady state for the model, indicating that an asymmetric steady state is infeasible. $[A] = [A_{pp}] = [A_p]\left(\frac{[P1]}{[P2]}\right)\left(\frac{c_2 k_2}{c_1 k_1}\right)$

**Ordered distributive TSP - common kinase and common phosphatase**

The ordered distributive triple site phosphorylation (TSP) with common kinase common phosphatase is capable of exhibiting case 1 symmetry with $[A]$ = $[A_{ppp}]$ and $[A_p]$ = $[A_{pp}]$. The kinetic constraints necessary to impose this symmetry is given in the schematic in Appendix 2 - fig. 2.12; $k_4 = k_1$, $k_5 = k_2$, $k_6 = k_3$, $k_{b4} = k_{b1}$, $k_{b5} = k_{b2}$, $k_{b6} = k_{b3}$, $k_{ub4} = k_{ub1}$, $k_{ub5} = k_{ub2}$, $k_{ub6} = k_{ub3}$ and $P_{Total} = K_{Total}$. Under these constraints, case 1 symmetry is established with $[A]$ = $[A_{ppp}]$, $[A_p]$ = $[A_{pp}]$ along with $[K]$ = $[P]$. Solving the resulting system of equations for steady states we arrive at the following equation:

$$([K] - [P])\left(\frac{[A_p][[K] + [P]](k_1 - k_3)c_3 + k_1[P]}{[P]k_1}\right) = 0 \tag{B.15}$$

Hence for an asymmetric solution ($[K] \neq [P]$) to exist, the second term has to be $0$ and thus with this we find the features and necessary conditions of the asymmetric state. From this we show that at an asymmetric steady state, the sum of the concentration of the partially modified substrates ($[A_p]$ + $[A_{pp}]$) is fixed at a certain value (invariant), given only by key kinetic constants. This is an example (seen elsewhere) of the sum of concentrations of two species fixed at steady state. Further, since concentrations of substrates are strictly positive we can show that the asymmetric state can only exist when $k_3 > k_1$, which gives us the necessary conditions for the symmetry to break. The sufficiency of this condition follows by evaluating all species concentrations at values of $A_p$ and $A_{pp}$ obtained from this invariant and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. We show the presence of asymmetric states for some suitable value of $A_{Total}$. This is carried out in the appendix C.

- *Necessary and sufficient condition for symmetry breaking:*
  1. $k_3 > k_1$
- *Invariant:*
  1. $[A_p] + [A_{pp}] = \frac{1}{c_3}\left[\frac{k_1}{k_3 - k_1}\right]$

**Random System 1 - common kinase and common phosphatase**

The random distributive DSP with common kinase and common phosphatase acting on both modification sites (System 1) is capable of case 1, case 2 and case 3 symmetry. However only case 1 and case 3 symmetries can be broken. Here we present analytical arguments elucidating the presence of case 1 & case 3 symmetry breaking and its associated features. We also present the analytical arguments precluding case 2 symmetry breaking.

**Case 1 Symmetry:** Case 1 symmetry is established in the random DSP through the kinetic structure provided in Figure 2.1: ($k_3 = k_1$, $k_4 = k_2$, $k_{b3} = k_{b1}$, $k_{b4} = k_{b2}$, $k_{ub3} = k_{ub1}$, $k_{ub4} = k_{ub2}$, $a_3 = a_1$, $a_4 = a_2$, $a_{b3} = a_{b1}$, $a_{b4} = a_{b2}$, $a_{ub3} = a_{ub1}$, and $a_{ub4} = a_{ub2}$). In addition, the following constraint on enzyme total amounts also needs to be satisfied for exact symmetry to be present: $K_{Total} = P_{Total}$. Superimposing these kinetic constraints results in case 1 symmetry with $[A_{00}]$ = $[A_{11}]$ along with equality of the free enzyme concentrations $[K]$ = $[P]$. Similar to the analysis earlier on the ordered distributive DSP models, we solve the system for steady states in terms of fewer variables. This results in the following equation:

$$([K]-[P])\left(\frac{([A_{10}]((k_1-k_2)c_1-d_1k_2)c_2+c_1k_1)a_2+[A_{10}]c_2d_1k_2a_1)}{(c_1k_1a_2)}\right) \quad = \quad 0 \tag{B.16}$$

Thus for an asymmetric state ($[K] \neq [P]$) to exist, the second term needs to be equal to $0$. Setting this second term to zero reveals the features of the asymmetric steady state. Using this information, we can establish that the concentration of the partially modified substrates in the asymmetric steady states are both fixed at a constant value (invariants), determined by a few key kinetic constants. Since concentrations are always positive, we get the necessary conditions for symmetry breaking as shown below. The sufficiency of this condition follows by evaluating all species concentrations at this invariant value of $A_{01}$ and $A_{10}$ (inferring the other species concentrations) and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. From this, we can deduce that asymmetric states exist beyond a critical value of $A_{Total}$. This is carried out in the appendix C.

- *Necessary and sufficient condition for symmetry breaking:*
  1. $c_1a_2(k_2-k_1)+d_1k_2(a_2-a_1) > 0$
- *Invariants:*
  1. $[A_{01}] = \frac{-d_1k_2a_1}{d_2((d_1(a_1-a_2)-c_1a_2)k_2+c_1k_1a_2)}$
  2. $[A_{10}] = \frac{-c_1k_1a_2}{c_2(((k_1-k_2)c_1-d_1k_2)a_2+d_1k_2a_1)}$

**Case 2 symmetry:** Case 2 symmetry is likewise established in the random DSP through the kinetic structure provided in Figure 2.1 with constraints on pairs of parameters ($a_1 = k_1$, $a_2 = k_2$, $a_{b1} = k_{b1}$, $a_{b2} = k_{b2}$, $a_{ub1} = k_{ub1}$, $a_{ub2} = k_{ub2}$, $a_3 = k_3$, $a_4 = k_4$,

$a_{b3} = k_{b3}$, $a_{b4} = k_{b4}$, $a_{ub3} = k_{ub3}$, and $a_{ub4} = k_{ub4}$). However unlike other symmetries, it needs no additional constraint on total enzyme amounts for exact symmetry to be present. Under these conditions case 2 symmetry is established with $[A_{01}] = [A_{10}]$. Hence should symmetry break and an asymmetric steady state exist, it would be characterized by $[A_{01}] \neq [A_{10}]$. Solving for the steady states of these substrates in terms of the free enzyme concentrations and the fully unmodified substrate, we can ascertain that the concentrations of $[A_{01}]$ will always equal $[A_{10}]$ and thus that no asymmetric state can exist (See Source code file for more detail). In particular we can see that for any given steady state, $[A_{10}] = [A_{01}] = [A_{00}]\left(\frac{[K]}{[P]}\right)\left(\frac{c_1 k_1}{c_4 k_4}\right)$ making any asymmetric steady state impossible.

**Case 3:** Case 3 symmetry is established in the random DSP through the kinetic structure provided Figure 2.1 with constraints ($a_1 = k_3$, $a_2 = k_4$, $a_{b1} = k_{b3}$, $a_{b2} = k_{b4}$, $a_{ub1} = k_{ub3}$, $a_{ub2} = k_{ub4}$, $a_3 = k_1$, $a_4 = k_2$, $a_{b3} = k_{b1}$, $a_{b4} = k_{b2}$, $a_{ub3} = k_{ub1}$, and $a_{ub4} = k_{ub2}$). In addition, the following constraint on total enzyme concentrations needs to be satisfied for exact symmetry, $K_{Total} = P_{Total}$. Under these conditions, case 3 symmetry is established with $[A_{00}] = [A_{11}]$ and $[A_{01}] = [A_{10}]$, along with equality of the free enzyme concentrations $[K] = [P]$. Imposing these conditions and solving for the steady states of the system, we get the following correlation involving concentrations of the variables.

$$\begin{aligned}(\epsilon-1)(c_1c_2^2\epsilon k_1k_2^2 + (c_1k_1+c_3k_3)c_4(((\epsilon^2 - [A_{11}](c_1+c_3)\epsilon - [A_{11}] \quad ... \\ ... \quad (c_1+c_3))k_4 + [A_{11}]c_1k_1(\epsilon+1))k_2 + k_4[A_{11}]c_3k_3(\epsilon+1))c_2 + c_3c_4^2\epsilon k_3k_4^2) &= 0\end{aligned} \tag{B.17}$$

where $\epsilon = \frac{[K]}{[P]}$, Thus for an asymmetric state to exist ($[K] \neq [P]$, or $\epsilon \neq 1$), the second term needs to be equal to $0$. Using this information to further solve for the steady states of the ODEs reveals features of the asymmetric steady states. We find that the sum of the concentrations of the partially modified substrates are fixed at a constant value (invariant), determined by a few key kinetic constants. Since concentrations are always positive, we also get the necessary conditions for symmetry breaking as shown below. The sufficiency of this condition follows by evaluating all species concentrations at this invariant value of $A_{01} + A_{10}$ and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. This demonstrates that there exist values of $A_{Total}$ beyond which asymmetric steady states exist. This is carried out in appendix C.

- *Necessary and sufficient condition for symmetry breaking:*
  1. $c_3k_4(k_2 - k_3) - c_1k_2(k_1 - k_4) > 0$
- *Invariant:*
  1. $[A_{01}] + [A_{10}] = \frac{-c_1c_2k_1k_2 - c_3c_4k_3k_4}{(((-c_1-c_3)k_4 + c_1k_1)k_2 + c_3k_3k_4)c_4c_2}$

**Random System 2 - separate kinase and common phosphatase**

The random distributive DSP with separate kinase acting on each modification site and a common phosphatase effecting dephosphorylation (System 2) can only permit case 2 symmetry ($[A_{01}] = [A_{10}]$). This is due to the nature of independent enzymes effecting phosphorylation on each modification site while a common enzyme is responsible for dephosphorylation, which precludes case 1 and case 3 symmetry. In this subsection we present the analytical arguments precluding case 2 symmetry breaking in the network.

**Case 2:** Case 2 symmetry is established in the random ordered distributive DSP network through the kinetic structure provided Figure 2.1 with constraints ($a_1 = k_1$, $a_2 = k_2$, $a_{b1} = k_{b1}$, $a_{b2} = k_{b2}$, $a_{ub1} = k_{ub1}$, $a_{ub2} = k_{ub2}$, $a_3 = k_3$, $a_4 = k_4$, $a_{b3} = k_{b3}$, $a_{b4} = k_{b4}$, $a_{ub3} = k_{ub3}$, and $a_{ub4} = k_{ub4}$). In addition, the following constraints on total enzyme amounts need to be satisfied for exact symmetry: $K1_{Total} = K2_{Total}$. Under these kinetic constraints case 2 symmetry is established with $[A_{01}] = [A_{10}]$. Imposing these constraints and solving for the steady states of these substrates, we get the following correlation representing permissible steady states.

$$([K_1] - [K_2]) \left( \frac{[A_{00}]k_{b1} + k_1 + k_{ub1}}{k_1 + k_{ub1}} \right) = 0 \tag{B.18}$$

Since concentrations of the substrates and the kinetic constants are always positive, we can ascertain that the only steady state permitted by the system is when $[K_1] = [K_2]$, or $[A_{01}] = [A_{10}]$. Thus case 2 symmetry breaking is not possible in Random System 2 with separate kinase and common phosphatase. See appendix C.

**Random System 3 - separate kinase and separate phosphatase**

The random distributive DSP with separate kinase and separate phosphatase effecting modifications in each modification site (system 3) is capable of case 1, case 2 and case 3 symmetry, and each of these symmetries is capable of breaking. Here we derive the necessary and sufficient conditions for symmetry breaking, and the features of the symmetry broken state in that process.

**Case 1 Symmetry:** Case 1 symmetry is established in the random DSP through the kinetic structure provided Figure 2.1, with constraints ($k_3 = k_1$, $k_4 = k_2$, $k_{b3} = k_{b1}$, $k_{b4} = k_{b2}$, $k_{ub3} = k_{ub1}$, $k_{ub4} = k_{ub2}$, $a_3 = a_1$, $a_4 = a_2$, $a_{b3} = a_{b1}$, $a_{b4} = a_{b2}$, $a_{ub3} = a_{ub1}$, and $a_{ub4} = a_{ub2}$). In addition, the following constraint on total enzyme concentrations need to be satisfied for exact symmetry, $K1_{Total} = P2_{Total}$ and $K2_{Total} = P1_{Total}$. Under these conditions, case 1 symmetry is established with $[A_{00}] = [A_{11}]$ along with equality of the free enzyme concentrations $[K_1]$ & $[P_2]$ and $[K_2]$ & $[P_1]$. Imposing these kinetic constraints and evaluating the steady states of the system (by successively solving the steady states of ODEs for variables in terms of each other), we can get the

following equation representing permissible steady states of the model.

$$(\epsilon-1)\frac{[K1]^2\epsilon a_2c_1([A_{01}]c_2+1)k_1^2+k_2[P1]^2[A_{01}]a_1c_2d_1([A_{01}]c_2+1)k_1-[A_{01}]^2[P1]^2a_2c_2^2d_1k_2^2}{([K1]c_1\epsilon k_1^2a_2)} = 0 \tag{B.19}$$

where $\epsilon = \frac{[A_{00}]}{[A_{11}]}$. Thus for asymmetric solutions ($\epsilon \neq 1$) to exist, the second term in the expression needs to be equal to zero. Using this information and rearranging the terms, we can establish the following correlation between concentrations of the partially modified substrates $[A_{01}]$ and $[A_{10}]$,

$$\lambda_{[A_{01}]}\lambda_{[A_{10}]} = 1$$

where

$$\lambda_{[A_{01}]} = \frac{k_2c_2[A_{01}]}{k_1(c_2[A_{01}]+1)} \quad \text{and} \quad \lambda_{[A_{10}]} = \frac{a_2d_2[A_{10}]}{a_1(d_2[A_{10}]+1)} \tag{B.20}$$

This correlation represents asymmetric steady state solutions to the system of ODEs (for additional details, see appendix C). Simultaneously, by solving the steady state of the system, using the total individual enzyme conservation equations we get an alternative correlation between the partial substrate concentrations $[A_{01}]$ and $[A_{10}]$ as shown below, which is valid for all steady states,

$$\frac{P2_{Total}}{P1_{Total}} = \left[\frac{D\alpha\lambda_{[A_{01}]}+1}{D\alpha+\lambda_{[A_{10}]}}\right] \tag{B.21}$$

where

$$D = \frac{c_2[A_{01}]+1}{d_2[A_{10}]+1} \quad \text{and} \quad \alpha = \frac{[K2]+[P1]}{[K1]+[P2]} \tag{B.22}$$

Using the above correlations between concentrations of $[A_{00}]$ and $[A_{11}]$ (**??** representing the asymmetric solutions and **??** representing all feasible solutions) we can ascertain additional features of the asymmetric state. We find that in a symmetry broken state the concentration of the partially modified forms ($[A_{01}]$ and $[A_{10}]$) are fixed at constant values determined by key kinetic constants and total enzyme concentrations. Since concentrations cannot be negative, we can also thus isolate the necessary constraints for symmetry breaking in this network as shown below. The sufficiency of this condition follows by evaluating all species concentrations based on the invariant values of $A_{01}$ and $A_{10}$, and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. This indicates that there are values of $A_{Total}$ beyond which asymmetric steady states are guaranteed to exist. This is carried out in appendix C.

- *Necessary and sufficient conditions for symmetry breaking:*
  1. $k_2P1_{Total} > k_1P2_{Total}$
  2. $a_2P2_{Total} > a_1P1_{Total}$
- *Invariants:*
  1. $[A_{01}] = \frac{-k_1P2_{Total}}{c_2(k_1P2_{Total}-k_2P1_{Total})}$

2. $[A_{10}] = \frac{-a_1 P1_{Total}}{d_2(a_1 P1_{Total} - a_2 P2_{Total})}$

**Case 2 Symmetry:** Case 2 symmetry is likewise established in the random DSP through the kinetic structure provided in Figure 2.1, with constraints ($a_1 = k_1$, $a_2 = k_2$, $a_{b1} = k_{b1}$, $a_{b2} = k_{b2}$, $a_{ub1} = k_{ub1}$, $a_{ub2} = k_{ub2}$, $a_3 = k_3$, $a_4 = k_4$, $a_{b3} = k_{b3}$, $a_{b4} = k_{b4}$, $a_{ub3} = k_{ub3}$, and $a_{ub4} = k_{ub4}$). In addition to this the following constraints on total enzyme concentrations are also needed for exact symmetry, $K1_{Total} = K2_{Total} = K_{Total}$ and $P2_{Total} = P1_{Total} = P_{Total}$. Under these conditions, case 2 symmetry is established with $[A_{01}] = [A_{10}]$ along with equality of the free enzyme concentrations $[K_1] = [K_2]$ and $[P_1] = [P_2]$. Imposing these kinetic constraints and evaluating the steady states of the system, we can get the following equation representing permissible steady states of the model.

$$(\epsilon - 1)\frac{(-k_3\epsilon[P1]^2 c_4([A_{00}]c_1 + 1)k_4^2 - [K1]^2[A_{00}]c_1c_2k_1k_2([A_{00}]c_1 + 1)k_4 + [A_{00}]^2[K1]^2c_1^2c_2k_1^2k_3)[K2]}{([P1]c_4k_4^2k_3\epsilon)} = 0 \tag{B.23}$$

where $\epsilon = \frac{[A_{00}]}{[A_{11}]}$. Thus for asymmetric solutions ($\epsilon \neq 1$) to exist, the second term in the expression needs to be equal to zero. Using this information and rearranging the terms, we can establish the following correlation between concentrations of the fully modified and fully unmodified substrates $[A_{00}]$ and $[A_{11}]$,

$$\lambda_{[A_{00}]}\lambda_{[A_{11}]} = 1$$

where

$$\lambda_{[A_{00}]} = \frac{k_1c_1[A_{00}]}{k_4(c_1[A_{00}] + 1)} \quad \text{and} \quad \lambda_{[A_{11}]} = \frac{k_3c_3[A_{11}]}{k_2(c_3[A_{11}] + 1)} \tag{B.24}$$

This correlation represents a requirement for asymmetric steady state solutions to the system of ODEs. Simultaneously by solving the steady state of the system, using the total individual enzyme conservation equations, we get an alternative correlation between the partial substrate concentrations as shown below,

$$\frac{P_{Total}}{K_{Total}} = \alpha D\left[\frac{D\alpha\lambda_{[A_{00}]} + 1}{D\alpha + \lambda_{[A_{11}]}}\right] \tag{B.25}$$

where

$$D = \frac{c_1[A_{00}] + 1}{c_3[A_{11}] + 1} \quad \text{and} \quad \alpha = \frac{[K2] + [K1]}{[P1] + [P2]} \tag{B.26}$$

Using the above correlations between concentrations of $[A_{00}]$ and $[A_{11}]$ together (eq. (B.24) representing the asymmetric solutions and eq. (B.25) representing all feasible solutions), we can ascertain additional features of the asymmetric state. We find that in a symmetry broken state the concentrations of the completely modified and unmodified substrate forms ($[A_{00}]$ and $[A_{11}]$) are fixed at constant concentrations (invariant) given by key kinetic constants and total enzyme concentrations. Since concentrations cannot be negative, we can also thus isolate the necessary constraints for symmetry breaking in this network as shown below. The sufficiency of this condition follows by evaluating all species concentrations at this invariant value of $A_{01}$ and $A_{10}$ and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they

satisfy the system of ODEs at steady state and the conservation conditions. This then indicates that there exists finite positive values of $A_{Total}$ for which asymmetric states exist. This is carried out in appendix C.

- *Necessary and sufficient conditions for symmetry breaking:*
  1. $k_1 K_{Total} > k_4 P_{Total}$
  2. $k_3 P_{Total} > k_2 K_{Total}$
- *Invariants:*
  1. $[A_{00}] = \frac{k_4 P_{Total}}{c_1(k_1 K_{Total} - k_4 P_{Total})}$
  2. $[A_{11}] = \frac{-k_2 K_{Total}}{c_3(k_2 K_{Total} - k_3 P_{Total})}$

**Case 3:** Case 3 symmetry is established in the random DSP through the kinetic structure provided in Figure 2.1, with constraints ($a_1 = k_3$, $a_2 = k_4$, $a_{b1} = k_{b3}$, $a_{b2} = k_{b4}$, $a_{ub1} = k_{ub3}$, $a_{ub2} = k_{ub4}$, $a_3 = k_1$, $a_4 = k_2$, $a_{b3} = k_{b1}$, $a_{b4} = k_{b2}$, $a_{ub3} = k_{ub1}$, and $a_{ub4} = k_{ub2}$). In addition to this the following constraints on total enzyme concentrations are also needed for exact symmetry, $K1_{Total} = P1_{Total}$ and $K2_{Total} = P2_{Total}$. Under these conditions, case 3 symmetry is established with $[A_{01}] = [A_{10}]$ and $[A_{00}] = [A_{11}]$, along with equality of the free enzyme concentrations $[K_1] = [P_1]$ and $[P_2] = [K_2]$.

This symmetry can indeed break as shown computationally in the main text (Figure 2.3 and Appendix 2 - fig. 2.8). Similar to the approach used earlier, in order to obtain necessary and sufficient conditions for symmetry breaking, we solve for the steady states of the system of ODEs successively to obtain steady state correlations of variables in terms of each other. In so doing and by isolating asymmetric solutions of the type $[A_{00}] \neq [A_{11}]$ and $[A_{01}] \neq [A_{10}]$, we obtain the necessary conditions (see below) for symmetry breaking (Analysis not shown here - please refer to appendix C.

However unlike other classes of symmetries, the case 3 symmetry breaking in this model is not associated with a simple linear invariant in terms of concentrations of species. Moreover, the symmetry breaking in this instance can be of qualitatively different types, (i) where the asymmetry in the symmetry-broken states is more pronounced in the partial phosphoforms $A_{01}$ and $A_{10}$ or (ii) where the asymmetry is more pronounced in the fully modified/unmodified forms $A_{00}$ and $A_{11}$.

These come with contrasting qualitative implications with regard to the symmetry broken steady states. As shown earlier in Figure 2.3B (Panel 2) and Appendix 2 - fig. 2.8, depending on the nature of symmetry breaking outlined above, either the sum of the partially modified substrates, or the sum of the completely modified and unmodified substrates, can exhibit an approximate robustness in concentration relative to the other pair. This qualitative difference can be traced to the underlying kinetics. In particular, in our analytical work (appendix C), we have shown that the choice of kinetics can dictate the the nature of symmetry breaking.

This approximate concentration robustness exhibited by either pair of substrates can be traced to the fact that in an asymmetric steady state, one pair of substrate concentrations ([$A_{01}$] and [$A_{10}$] or [$A_{00}$] and [$A_{11}$]) are bounded and asymptotically reach a certain value as total substrate concentration approaches infinity, while the other pair can vary in an unbounded manner with varying total substrate concentration. This is summarized below.

- *Necessary and sufficient conditions for symmetry breaking*
  1. $k_1 > k_4$ and $k_3 > k_2$

     OR

  2. $k_4 > k_1$ and $k_2 > k_3$
- *Features of asymmetric steady states*
  1. IF $k_1 > k_4$ and $k_3 > k_2$
     - Concentrations of [$A_{00}$] and [$A_{11}$] exhibit approximate robustness.
       * Asymptotic concentration of [$A_{00}$] at infinite $A_{Total}$ = $\frac{k_4}{c_1(k_1-k_4)}$ or $\frac{k_2}{c_3(k_3-k_2)}$
       * Asymptotic concentration of [$A_{11}$] at infinite $A_{Total}$ = $\frac{k_2}{c_3(k_3-k_2)}$ or $\frac{k_4}{c_1(k_1-k_4)}$
  2. IF $k_4 > k_1$ and $k_2 > k_3$
     - Concentration of [$A_{01}$] and [$A_{10}$] exhibit approximate robustness.
       * Asymptotic concentration of [$A_{01}$] at infinite $A_{Total}$ = $\frac{k_3}{c_2(k_2-k_3)}$ or $\frac{k_1}{c_4(k_4-k_1)}$
       * Asymptotic concentration of [$A_{10}$] at infinite $A_{Total}$ = $\frac{k_1}{c_4(k_4-k_1)}$ or $\frac{k_3}{c_2(k_2-k_3)}$

These asymptotes obtained analytically have been cross validated with bifurcation analysis and the specific cross validation for the figures in chapter 2 are provided in the file (Read_Me.mw of the Source code file and appendix C).

**Mixed Random 1 - common kinase and common phosphatase**

The mixed-random ordered DSP with a common kinase effecting distributive phosphorylation and a common phosphatase effecting processive dephosphorylation is capable of case 2 symmetry. However this symmetry cannot be broken. In this subsection we present the analytical arguments precluding case 2 symmetry breaking in this network.

**Case 2 Symmetry:** Case 2 symmetry is established in the mixed random DSP through the kinetic structure provided in Figure 2.2E with constraints $a_1 = k_1$, $a_{b1} = k_{b1}$, $a_{ub1} = k_{ub1}$, $a_2 = k_2$, $a_{b2} = k_{b2}$, $a_{ub2} = k_{ub2}$, $a_3 = k_3$, $a_{b3} = k_{b3}$, $a_{ub3} = k_{ub3}$ and $k_4 = a_4$.

Under these conditions case 2 symmetry is established with $[A_{01}] = [A_{10}]$. Imposing these constraints and solving for the steady states of these substrates in terms of the free enzyme concentrations and the fully unmodified substrate, we can ascertain that $[A_{01}]$ always equals $[A_{10}]$. In particular $[A_{01}] = [A_{10}] = [A_{00}]\left(\frac{c_1 k_1}{c_2 k_2}\right)$ is always true at steady state, making any asymmetric steady state infeasible (See appendix C for more details).

**Mixed Random 2 - separate kinase and common phosphatase**

The mixed random DSP with separate kinases acting on each site effecting distributive phosphorylation and a common phosphatase effecting processive dephosphorylation is capable of case 2 symmetry and it is possible for this symmetry to break. In this subsection, we show the necessary conditions for symmetry breaking and features of the asymmetric state.

**Case 2:** Case 2 symmetry is established in the mixed random DSP through the kinetic structure provided Figure 2.2E with constraints $a_1 = k_1$, $a_{b1} = k_{b1}$, $a_{ub1} = k_{ub1}$, $a_2 = k_2$, $a_{b2} = k_{b2}$, $a_{ub2} = k_{ub2}$, $a_3 = k_3$, $a_{b3} = k_{b3}$ and $a_{ub3} = k_{ub3}$ and $k_4 = a_4$. In addition the following constraints on total enzyme concentrations needs to be satisfied for exact symmetry: $K1_{Total} = K2_{Total}$. Under these conditions case 2 symmetry is established with $[A_{01}] = [A_{10}]$ along with equality of the free enzyme concentrations $[K_1] = [K_2]$. Imposing these kinetic constraints and evaluating the steady states of the system, we get the following equation representing permissible steady states of the model.

$$(\epsilon - 1)(-k_1\epsilon + [A_{01}]c_2(k_1 - k_2)) \quad = \quad 0 \tag{B.27}$$

where $\epsilon = \frac{[K_1]}{[K_2]}$. Thus we can see that the system accommodates an asymmetric steady state where $\epsilon \neq 1$, provided the second term in the equation is *0*. From this we can ascertain the features of asymmetric state by isolating correlations for $A_{01}$ from the above equation and using it to simplify the steady states of system of ODEs. We find that in a symmetry broken state the concentrations of the completely modified and unmodified substrate forms ($[A_{00}]$ and $[A_{11}]$) are fixed at constant values (invariant) given by key kinetic constants and total enzyme concentrations. Since concentrations cannot be negative, we can also thus isolate the necessary constraints for symmetry breaking in this network as shown below. The sufficiency of this condition follows by evaluating all species concentrations at this invariant value of $A_{00}$ and $A_{11}$ and proving that (i) they are positive when the necessary condition is satisfied, and (ii) that they satisfy the system of ODEs at steady state and the conservation conditions. This ensures that there exists values of $A_{Total}$ for which asymmetric states exist. This is carried out in the appendix C.

- *Necessary and sufficient conditions for symmetry breaking:*

  1. $k_2 < k_1$

2. $k_2 K1_{Total}(k_3 + k_4) < P_{Total} k_3 k_4$

- *Invariant(s):*

1. $[A_{00}] = \frac{1}{c_1}\left[\frac{k_2}{(k_1 - k_2)}\right]$
2. $[A_{11}] = \frac{-k_2 k_4 K1_{Total}}{(2(k_2 K1_{Total}(k_3+k_4) - P_{Total} k_3 k_4))c_3}$

In the limit where dephosphorylation is acting in the unsaturated limit (Mixed Random 2A), as described in the models section earlier, the dephosphorylation is replaced by a single linear reaction from $A_{11}$ to $A_{00}$ (with a rate constant which is denoted by $k_3$ which implicitly contains the effect of the total phosphatase concentration: in fact it corresponds to $k_{b3} P_{Total}$ in the previous model). In this case the symmetry established can still break. The necessary and sufficient conditions, and the features of the asymmetric state are as follows (the analysis is the same as above - see appendix C for more details),

- *Necessary and sufficient condition for symmetry breaking:*

1. $k_2 < k_1$

- *Invariant(s):*

1. $[A_{00}] = \frac{1}{c_1}\left[\frac{k_2}{k_1 - k_2}\right]$
2. $[A_{11}] = K1_{Total}\left[\frac{k_2}{2k_3}\right]$

**Sufficiency of necessary conditions**

We have established necessary conditions for the presence of symmetry breaking in various classes of symmetries and networks. We have also further shown that these necessary conditions are also sufficient for symmetry breaking to occur at some positive total substrate concentration. The sufficiency argument has been briefly discussed above at the appropriate sections and the approach is consolidated below. For more details please refer to appendix C. The feasibility of any steady state for the models described above relies on the concentrations of variables at steady state satisfying three separate constraints.

1. Variable values satisfy ODE description of the model (for a given set of kinetic parameters)
2. Conservation equations (for the enzymes and total substrate concentrations) should be satisfied.
3. All variable values (concentrations) must be positive.

We obtained the necessary conditions for symmetry breaking by suitably leveraging points 1, 2 and 3 above. By solving for the steady states of the system of ODEs, we obtained correlations between concentrations of various variables in terms of kinetic

parameters and total enzyme concentrations. Further, by isolating asymmetric solutions, we established the features of the asymmetric state. Then by requiring that the concentrations of key substrates (invariants) be positive, we obtained the necessary conditions.

We extended this argument to ensure sufficiency of these conditions by evaluating concentrations of all variables in the asymmetric state and showing that they are indeed positive if the necessary conditions (in terms of kinetic constants and total enzyme concentrations) are satisfied for some total substrate concentration (note here that the bifurcation diagram is along $A_{Total}$). We also verify that in each case the total enzyme conservations are satisfied. In this manner we have shown that the necessary conditions are indeed sufficient for symmetry breaking to occur for some total enzyme and substrate concentration.

In addition - the analysis provided in the appendix C) also evaluates and predicts the position of symmetry breaking along $A_{Total}$ (not shown here in the appendix). Note that the intersection of the symmetric steady state and the asymmetric steady state denotes a Pitchfork bifurcation and thus the position of symmetry breaking. In each case, along with the invariants we also evaluate the total substrate concentration at which the pitchfork bifurcation occurs by imposing features of symmetry in terms of concentration of various variables in the asymmetric steady state invariants and correlations. This is presented in the Source code file (*Maple* file) and has been cross validated with bifurcation analysis for all plots provided chapter 2. This prediction and cross validation can be found in (Read_Me.mw) and appendix C.

**Origins of absolute concentration robustness**

In each instance of symmetry breaking encountered in the above models, we observe that either the concentration of specific individual substrates or sum of concentrations of specific substrates exhibit strict absolute concentration robustness in the asymmetric branches (with case 3 symmetry breaking in the separate kinase separate phosphatase network being an exception, where *approximate* concentration robustness is exhibited by sums of concentrations of pairs of substrates). Overall, following symmetry breaking in these networks, the asymmetric steady states are characterized by the concentration of the specific substrates (or sums thereof) being (exactly) fixed and in fact this remains so indefinitely along increasing $A_{Total}$ concentrations.

To understand the origins of ACR and its dependence on symmetry and symmetry breaking, we use the ordered DSP model as a basis of exploration.

Note that ACR in this instance is defined to mean the concentration of some substrate form being exactly maintained at some fixed concentration (along a steady state branch) for a range of total concentrations of either the substrate or the enzymes. In this definition, we make no assumptions on the kinetic regime of modification/demodification of substrates. Note that it is possible for the concentration of a substrate species to be

*approximately* constant (something encountered in limiting regimes of enzymatic action) and this will be described subsequently.

We begin by analyzing the ordered distributive double site phosphorylation network with common kinase and common phosphatase (refer fig. 2.1 for the schematic) in the absence of any symmetry in either the kinetic parameters or the total concentrations of enzymes/substrate. This network has three substrate forms, $A$, $A_p$ and $A_{pp}$ that can potentially exhibit ACR. Note that earlier in our analysis of case 1 symmetry breaking in the model, we observed $A_p$ exhibiting ACR with increasing concentration of $A_{Total}$, in the asymmetric branches following symmetry breaking.

The analysis carried out here, is structured to answer four key questions regarding the phenomenon of ACR.

- Which substrates in this network are capable of exhibiting ACR?
- Is ACR (where possible) only exhibited with respect to the (change of) total substrate concentration? Is ACR possible with changing total enzyme concentration?
- Are there any additional constraints on the kinetic parameters required to observe ACR (where possible)?
- What associated features (if any) are exhibited when ACR is observed in this network?

The mathematical analysis is described in detail in the attached *Maple* document (5.1). The results are summarized here following which the key analytical arguments used are briefly presented.

- Our analysis reveals that ACR can only be exhibited by the partially modified substrate form ($A_p$)
- ACR in $A_p$ is only possible with changing concentration of total substrate.
- Further analysis of the ACR in $A_p$ reveals necessary (and sufficient) constraints on the kinetics and total concentrations of enzymes.

$$\frac{k_3 P_{Total}}{c_2\left(k_2 K_{Total} - k_3 P_{Total}\right)} \quad = \quad \frac{k_1 K_{Total}}{c_4\left(k_4 P_{Total} - k_1 K_{Total}\right)} \quad > \quad 0$$

- Our analysis was able to further characterize associated features of this network when $A_p$ exhibits ACR. In particular,
  - If this system exhibits ACR in $A_p$ for a range of total substrate concentration, it is necessarily multistable within that range with two steady states exhibiting ACR.
  - There necessarily exists another steady state branch for all positive total substrate concentrations. This branch intersects one of the two ACR branches

which at some $A_{Total}$ value (computationally found to be a transcritical bifurcation, refer Appendix 2 - fig. 2.13)

– This non-ACR branch state is characterized by a fixed ratio of free (unbound) Kinase to free (unbound) Phosphatase concentration for all positive total substrate concentrations, given by

$$\frac{[K]}{[P]} = \frac{K_{Total}}{P_{Total}}$$

This analysis allows us to summarise a number of key insights, specifically in the context of symmetry and symmetry breaking.

Complemented by Appendix 2 - fig. 2.13, our analysis reveals that a strict case 1 symmetry is not a prerequisite for encountering ACR in the network. However there is a kinetic requirement (which is sufficient) for the network to exhibit ACR (mathematically exact) (see above and in eq. (B.31)). It is noteworthy to observe that this requirement is satisfied trivially by an assumption of case 1 symmetry in the network, but is a weaker condition. Similarly the non-ACR branch of steady states becomes the symmetric branch with the assumption of case 1 symmetry and the transcritical bifurcation (at the intersection of the non-ACR branch and an ACR branch) as seen in Appendix 2 - fig. 2.13 becomes a pitchfork bifurcation.

Thus while case 1 symmetry is not a strict pre-requisite for obtaining ACR in the model, it is a suitable vantage point to analyze the phenomenon, additionally serving to reduce the parametric complexity of the network and highlighting multi-site specific characteristics such as directionality of modifications.

In the following subsection - we provide the key insight used (refer to attached *Maple* document (Source code file, section 5.1) to obtain proofs for the above conclusions. We also provide a sketch of the proof establishing the presence of ACR in $A_p$ with increasing (changing) $A_{Total}$ here.

**Approach used for determining absence or presence of ACR in the ordered DSP with common kinase and common phosphatase**:

In the absence of any imposed symmetry, the DSP model is a set of 9 ODEs involving 15 parameters, 9 variables and 3 conservation conditions. However at steady state the ODEs reduce to a system of non-linear equations for the variable concentrations. The solution of these equations in sequence, allows for eliminating many of the variables (by writing them in terms of a smaller set of variables). This systematic algebraic reduction of the system of equations allows for the system at steady state to be represented as a set of two coupled polynomial equations in two variables. In each proof this reduction is in terms of the following two variables, namely, variable of the species being investigated for exhibiting the ACR and the variable, which is the ratio of free kinase to free phosphatase concentrations (denoted by $\epsilon$). Note that the (feasible) solutions of

these coupled polynomials determine the steady state concentrations, as concentration of all other variables of the system can be obtained as functions of these variables.

As a consequence of ACR in the variable of interest, with changing total concentration (of substrate/enzyme), both equations need to be satisfied with only $\epsilon$ being allowed to vary. Thus the resulting two polynomials should accommodate a common root for $\epsilon$ for changing total amounts of either substrate/the enzymes (as pertinent to the proof) if ACR is to be exhibited in the substrate form of interest. This insight allows us to rule out ACR by contradiction in cases where the resulting polynomials do not provide this flexibility. Further in the case of ACR in $A_p$ with total amount of substrate - it allows us to elucidate the necessary and sufficient conditions for obtaining ACR in the network.

In the next section we show exactly how this is pursued by providing a sketch of the proof for the presence of ACR in $A_p$ with changing total amounts of substrate.

**Proof: Partially modified substrate exhibits ACR with changing total substrate concentration**

As mentioned earlier - at steady the DSP model can be simplified as a set of coupled polynomials in two variables; namely, the substrate form being investigated for exhibiting ACR and the ratio of free (unbound) kinase and the free (unbound) phosphatase. In the context of this specific proof the two variables are $A_p$ and $\epsilon$. This results in the following two coupled polynomials whose feasible solutions determine the steady state concentrations of the system.

$$0 = [A_p c_2(k_2 K_{Total} - k_3 P_{Total}) - k_3 P_{Total}]\, k_1 \epsilon + [A_p c_4(k_1 K_{Total} - k_4 P_{Total}) + k_1 K_{Total}]\, k_3 \tag{B.28}$$

$$\begin{aligned} 0 = {} & \left[-(c_3(c_1\epsilon * k_1 + c_4 k_4)k_3 + c_1 c_2 \epsilon^2 k_1 k_2)(c_2 \epsilon k_2 + c_4 k_3)\right] A_p^2 + \\ & \Big[k_3(c_2 k_1 c_1((-P_{Total}k_3 + k_2(A_{Total} - P_{Total}))c_3 - k_2)\epsilon^2 + \\ & k_3 c_1 c_3((c_4 A_{Total} - c_4 P_{Total} - 1)k_1 - P_{Total} c_4 k_4)\epsilon - c_3 c_4 k_3 k_4)A_p + \\ & c_1 c_3 \epsilon k_1 k_3^2 A_{Total} \end{aligned} \tag{B.29}$$

where the concentrations of the other variables can be obtained from these as shown below.

$$\begin{aligned} [A] &= \left[\frac{A_p}{\epsilon}\right]\left[\frac{c_4 k_4}{c_1 k_1}\right] & & & [A_{pp}] &= [A_p][\epsilon]\left[\frac{c_2 k_2}{c_3 k_3}\right] \\ [AK] &= \frac{k_3 c_4 k_4 P_{Total}[A_p]}{k_1([\epsilon][A_p]c_2 k_2 + k_3 c_4 [A_p] + k_3)} & & [A_p K] &= \frac{\epsilon P_{Total} k_3 c_2 [A_p]}{k_3 + (c_2 k_2 \epsilon + c_4 k_3)[A_p]} \\ [A_{pp}P] &= \frac{k_2 c_2 P_{Total}[A_p][\epsilon]}{k_3 + k_3 c_4 [A_p] + k_2 c_2 [A_p][\epsilon]} & & [A_p P] &= \frac{k_3 c_4 P_{Total}[A_p]}{k_3 + (k_3 c_4 + k_2 c_2)[A_p]} \\ [P] &= \frac{P_{Total} k_3}{k_3 + k_3 c_4 [A_p] + c_2 k_2 [A_p][\epsilon]} \end{aligned} \tag{B.30}$$

Now assuming $A_p$ exhibits ACR for changing $A_{Total}$, both equations need to be satisfied with only epsilon allowed to vary.

The two polynomials have the following structure. 1. The polynomial in eq. (B.29) is a cubic polynomial in $\epsilon$ whose coefficients include the parameter $A_{Total}$ (and $[A_p]$). 2.

The polynomial in eq. (B.28) is a linear expression in $\epsilon$ and its coefficients do not involve total substrate concentration parameter $A_{Total}$ (but includes $[A_p]$).

It then follows that as $A_{Total}$ changes the roots of the cubic polynomial eq. (B.29) (for $\epsilon$) change. However, the polynomial in eq. (B.28) cannot accommodate a changing $\epsilon$ (being independent of $A_{Total}$), unless the presence of $\epsilon$ in this equation is eliminated by a suitable parameter choice. In this instance, equation eq. (B.28) can be satisfied independent of $\epsilon$ (and $A_{Total}$).

This can only be accomplished if the coefficient of $\epsilon$ and the constant term in the polynomial (eq. (B.28)) are both simultaneously zero. Thus we get kinetic constraints that permit the possibility of ACR. Since these expressions both include $A_p$, this also provides the ACR concentration of $A_p$ as shown below,

$$[A_p] \quad = \quad \frac{k_3 P_{Total}}{c_2\left(k_2 K_{Total} - k_3 P_{Total}\right)} \quad = \quad \frac{k_1 K_{Total}}{c_4\left(k_4 P_{Total} - k_1 K_{Total}\right)} \quad > \quad 0 \tag{B.31}$$

Hence, the above simultaneously establishes the ACR concentration of $A_p$, while providing us the kinetic constraints required for ACR. As mentioned earlier - it is also noteworthy to observe that a case 1 symmetry in kinetics and enzyme concentrations trivially transforms this kinetic constraint to that observed earlier in the symmetric instance.

Thus, under conditions in, eq. (B.31), the polynomial eq. (B.28) is satisfied independent of $A_{Total}$ or $\epsilon$. The other polynomial (eq. (B.29)) can be rewritten as a polynomial in $\epsilon$ (with $A_p$ assuming the fixed value given in eq. (B.31)) as shown below in eq. (B.32).

$$\begin{aligned}
0 \quad = \quad & \left[c_1 c_2^2 k_1 k_2^2 [A_p]^2\right] \epsilon^3 + \\
& [c_1((([A_p] + P_{Total} - A_{Total})c_3 + [A_p]c_4 + 1)k_2 + P_{Total} c_3 k_3)[A_p]k_3 c_2 k_1]\, \epsilon^2 + \\
& \Big[(c_1([A_p]^2 c_4 k_1 + (((-A_{Total} + P_{Total})k_1 + P_{Total} k_4)c_4 + k_1)[A_p] - A_{Total} k_1)k_3 + [A_p]^2 c_2 c_4 k_2 k_4)k_3 c_3\Big]\, \epsilon + \\
& \Big[[A_p]c_3 c_4 k_3^2 k_4([A_p]c_4 + 1)\Big]
\end{aligned} \tag{B.32}$$

Note that the coefficient of $\epsilon^3$ and the constant term in this polynomial are both positive indicating the presence of one negative real root. This also implies that the product of the three roots is negative. Thus, if there exists an ACR branch (i.e. positive real roots in $\epsilon$) it necessarily implies that there is exists another positive real root for $\epsilon$ indicating the presence of another second ACR branch.

*Sufficiency*: For large enough $A_{Total}$, the sum of the roots (the ratio of the coefficient of the $\epsilon^2$ and the leading coefficient) is positive and it can be shown that the discriminant of this polynomial is also positive (refer appendix C). This together guarantees the presence of two positive real roots in $\epsilon$, proving the sufficiency of kinetic constraints in eq. (B.32) to obtain ACR in $A_p$ at some finite range of concentrations of $A_{Total}$.

Extension of this proof to ascertain associated features of ACR networks in DSP is provided in detail in appendix C.

**Approximate concentration robustness**

We now turn to the case of approximate ACR wherein a substrate may exhibit concentration robustness approximately in some limiting regime of enzymatic action. We will show that such behavior can be readily obtained in a different scenarios.

We begin our analysis by focussing on the ordered double site phosphorylation network. We assume that the concentration of the phosphatase is much higher than the concentration of the kinase or the substrate. Thus the dephosphorylation of substrates, $A_p$ and $A_{pp}$, can be approximated by simple first order reactions.

*Approximate concentration robustness in the partial substrate ($A_p$)*: We assume that the phosphorylation of $A_p$ is in the unsaturated limit and the phosphorylation of $A$ is saturating. In this case the concentration of the $A_pK$ complex is negligible and further $K \approx \frac{K_{Total}}{1+\alpha A} \approx \frac{K_{Total}}{\alpha A}$. This implies that the flux of phosphorylation of $A_p$ is a zeroth order reaction, $\frac{K_{Total}\beta}{\alpha}$. Since steady state in this network involves pairwise equilibrium, we have,

$$Rate_{Phosphorylation-of-A} = Rate_{DePhosphorylation-of-A_p} \implies \frac{\beta K_{Total}}{\alpha} \approx A_p\gamma$$

where $\alpha$, $\beta$, $\gamma$ are kinetic constants of the network

Thus the concentration of $A_p$ is approximately given by $A_p \approx \frac{\beta K_{Total}}{\alpha\gamma}$ for a range of $A_{Total}$.

*Approximate concentration robustness in the fully modified substrate ($A_{pp}$)*:

Here similar to the logic above, if the phosphorylation of $A_p$ is in the saturated limit and the phosphorylation of $A$ is in the unsaturated limit, then (as a consequence of equilibrium between $A_p$ and $A_{pp}$) we find that $A_{pp}$ exhibits approximate ACR.

A similar proof assuming an excess of kinase in the network can be used to show the feasibility of approximate concentration robustness in $A$.

*Comment about different kinases and different phosphatases*: The above logic can readily be used in the case of different kinases or phosphatases. In fact having different enzymes only removes constraints on kinetic regimes/enzyme amounts.

*Random modification networks*: The above approach can be employed to establish approximate ACR for one species in the random modification network. This is facilitated by having separate kinases and separate phosphatases. Here we simply need to ensure that the two production reactions for a substrate are acting in the saturated regime (these are associated with different enzymes) which can be assumed to perform other modifications in the unsaturated limit. Further enzymes involved in the removal of the substrate are assumed to be in large amounts relative to the substrate concentration.

## B.2 Multi-level complexity of cells and populations

**Mathematical description of the various intra-cellular circuits considered**

As mentioned earlier these models have been adapted from Tyson et al., [52]. The modular description of the Single population and two population master models mean that for the most part these models are considered as black boxes with input Signal ($S$, which is manipulated by various interactions and the other levels), and output $c$. In order to ensure that the circuit description is modular and isn't affected by the interactions or dilution of the response into or out of the cell, the output $c$, is made to affect the production of the specifically defined intra-cellular species $e_1$, which then diffuses out as $e_2$. Please see master model description in the chapter for a more elaborate description of this setup.

Here we describe the mathematical description of these systems and the parameter values chosen for them. These parameter values are rarely disturbed, thus allowing us to get a clear sense of the expected response from the system.

**Covalent modification system**
*Expected behavior:* Sigmoidal response with saturation.

$$\begin{aligned} \frac{dc}{dt} &= \frac{k_1 S R_p}{K_{m1} + R_p} - \frac{k_2 c}{K_{m2} + c} \\ \frac{dR_p}{dt} &= \frac{k_2 c}{K_{m2} + c} - \frac{k_1 S R_p}{K_{m1} + R_p} \end{aligned} \tag{B.33}$$

*Parameters*: $k_1$ = 1; $k_2$ = 1; $R_p + c$ = 1, $K_{m1}$ = 0.05; $K_{m2}$ = 0.05;

**Positive feedback motif** *Expected behavior:* Mutual activation model that is capable of presenting with bistability (irreversible).

$$\begin{aligned} \frac{dc}{dt} &= k_0 E_p + k_1 S - k_2 X c \\ \frac{dE}{dt} &= \frac{k_3 c E_p}{K_3 + E_p} - \frac{k_4 E}{K_4 + E} \\ \frac{dE_p}{dt} &= \frac{k_4 E}{K_4 + E} - \frac{k_3 c E_p}{K_3 + E_p} \end{aligned} \tag{B.34}$$

*Parameters*: $k_0$ = 0.4; $k_1$ = 0.01; $k_2$ = 1; $k_3$ = 1; $k_4$ = 0.2; $K_3$ = 0.05; $K_4$ = 0.05;

**Negative feedback motif**
*Expected Behavior:* Homeostatic response.

$$\begin{aligned} \frac{dc}{dt} &= k_0 E c - k_2 S c \\ \frac{dE}{dt} &= \frac{k_3 E_p}{K_3 + E_p} - \frac{k_4 c E}{K_4 + E} \\ \frac{dE_p}{dt} &= \frac{k_4 c E}{K_4 + E} - \frac{k_3 E_p}{K_3 + E_p} \end{aligned} \tag{B.35}$$

*Parameters*: $k_0$ = 1; $k_2$ = 1; $k_3$ = 0.5; $k_4$ = 1; $K_3$ = 0.01; $K_4$ = 0.01;

**Negative feedback oscillator**

*Expected Behavior:* Oscillatory response (through a super-critical Hopf point)

$$\begin{aligned}
\frac{dX}{dt} &= k_0 + k_1 S - k_2 X + k_{21} c X \\
\frac{dY_p}{dt} &= \frac{k_3 X Y}{K_3 + Y} - \frac{k_4 c Y_p}{K_4 + Y_p} \\
\frac{dY}{dt} &= \frac{k_4 c Y_p}{K_4 + Y_p} - \frac{k_3 X Y}{K_3 + Y} \\
\frac{dR}{dt} &= \frac{k_6 c}{K_6 + c} - \frac{k_5 Y_p R}{K_5 + R} - \\
\frac{dc}{dt} &= \frac{k_5 Y_p R}{K_5 + R} - \frac{k_6 c}{K_6 + c}
\end{aligned} \tag{B.36}$$

*Parameters*: $k_0$=0, $k_1$=1, $k_2$=0.01, $k_{21}$=10, $k_3$=0.1, $k_4$=0.2, $k_5$=0.1, $k_6$=0.05, $Y + Y_p$ = 1; $R + R_p$=1; $K_3$ = 0.01; $K_4$ = 0.01; $K_5$ = 0.01; $K_6$ = 0.01;

**Substrate depletion oscillator**

*Expected Behavior:* Oscillatory response (through a sub-critical Hopf point)

$$\begin{aligned}
\frac{dX}{dt} &= k_1 S - (k_{01} + k_0 * E_p) X \\
\frac{dc}{dt} &= (k_{01} + k_0 * E_p) X - k_1 c \\
\frac{dE_p}{dt} &= \frac{k_3 R E}{K_3 + E} - \frac{k_4 E_p}{K_4 + E_p} \\
\frac{dE}{dt} &= \frac{k_4 E_p}{K_4 + E_p} - \frac{k_3 R E}{K_3 + E} -
\end{aligned} \tag{B.37}$$

*Parameters*: $k_{01}$ = 0.01; $k_0$ = 0.4; $k_1$ = 1; $k_2$ = 1; $k_3$ = 1; $k_4$ = 0.3; $K_3$ = 0.05; $K_4$ = 0.05;

**Positive feedback motif (reversible bistability - exemplar)**

*Expected behavior:* Mutual inhibition model that is capable of presenting with bistability (irreversible).

$$\begin{aligned}
\frac{dc}{dt} &= k_0 + k_1 S - k_2 c - k_{21} E_p c \\
\frac{dE}{dt} &= \frac{k_3 E_p}{K_3 + E_p} - \frac{k_4 c E}{K_4 + E} \\
\frac{dE_p}{dt} &= \frac{k_4 c E}{K_4 + E} - \frac{k_3 E_p}{K_3 + E_p}
\end{aligned} \tag{B.38}$$

*Parameters*: $k_0$ = 0; $k_1$ = 0.05; $k_2$ = 0.1; $k_{21}$ = 0.5; $k_3$ = 1; $k_4$ = 0.2; $K_3$ = 0.05; $K_4$ = 0.05;

The expected circuit response from these networks is given in fig. 5.3.

# Bibliography


[1] Thapanar Suwanmajo. "Modelling and systems approaches for elucidating information processing in multi-site phosphorylation systems". 2017.

[2] Friedrich Marks, Ursula Klingmüller, and Karin Müller-Decker. *Cellular Signal Processing*. Second edition. — New York, NY: Garland Science, 2017.: Garland Science, May 2017. ISBN: 9781315165479. DOI: 10.4324/9781315165479. URL: https://www.taylorfrancis.com/books/9781351677226.

[3] Alba Jiménez et al. "A spectrum of modularity in multi-functional gene circuits". In: *Molecular Systems Biology* 13.4 (2017), p. 925. ISSN: 1744-4292. DOI: 10.15252/msb.20167347.

[4] Kee Myoung Nam et al. "Robustness and parameter geography in post-translational modification systems". In: *PLoS Computational Biology* 16.5 (2020), pp. 1–52. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1007573.

[5] Bruce. Alberts. *Molecular biology of the cell*. 5th ed. New York: Garland Science, 2008. ISBN: 9780815341062.

[6] Fei Sun et al. "Revealing various coupling of electron transfer and proton pumping in mitochondrial respiratory chain". In: *Current Opinion in Structural Biology* 23.4 (2013), pp. 526–538. ISSN: 0959440X. DOI: 10.1016/j.sbi.2013.06.013. URL: http://dx.doi.org/10.1016/j.sbi.2013.06.013.

[7] Nika N. Danial and Stanley J. Korsmeyer. "Cell Death: Critical Control Points". In: *Cell* 116.2 (2004), pp. 205–219. ISSN: 00928674. DOI: 10.1016/S0092-8674(04)00046-7.

[8] Maik Hüttemann et al. "The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis". In: *Mitochondrion* 11.3 (2011), pp. 369–381. ISSN: 15677249. DOI: 10.1016/j.mito.2011.01.010.

[9] Vsevolod V. Gurevich and Eugenia V. Gurevich. "Analyzing the roles of multifunctional proteins in cells: The case of arrestins and GRKs". In: *Critical Reviews in Biochemistry and Molecular Biology* 50.5 (2015), pp. 440–452. ISSN: 15497798. DOI: 10.3109/10409238.2015.1067185.

[10] Constance J. Jeffery. "What is Protein Moonlighting and Why is it Important?" In: *Moonlighting Proteins: Novel Virulence Factors in Bacterial Infections* figure 1 (2016), pp. 1–19. DOI: 10.1002/9781118951149.ch1.

[11] Constance J. Jeffery. "An introduction to protein moonlighting". In: *Biochemical Society Transactions* 42.6 (2014), pp. 1679–1683. ISSN: 14708752. DOI: 10.1042/BST20140226.

[12] W. Xu et al. "The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase". In: *Blood* 87.11 (1996), pp. 4502–4506. ISSN: 00064971. DOI: 10.1182/blood.v87.11.4502.bloodjournal87114502.

[13] Nelli G. Markova, Adriana Pinkas-Sarafova, and Marcia Simon. "A metabolic enzyme of the short-chain dehydrogenase/reductase superfamily may moonlight in the nucleus as a repressor of promoter activity". In: *Journal of Investigative Dermatology* 126.9 (2006), pp. 2019–2031. ISSN: 0022202X. DOI: 10.1038/sj.jid.5700347.

[14] Christopher Walsh. *Posttranslational Modification of Proteins. Expanding Nature's Inventory.* First Edit. New York: Roberts & Co, 2005. ISBN: 0974707732.

[15] Etienne Hebert-Chatelain. "Src kinases are important regulators of mitochondrial functions". In: *International Journal of Biochemistry and Cell Biology* 45.1 (2013), pp. 90–98. ISSN: 18785875. DOI: 10.1016/j.biocel.2012.08.014. URL: http://dx.doi.org/10.1016/j.biocel.2012.08.014.

[16] Emine C. Koc and Hasan Koc. "Regulation of mammalian mitochondrial translation by post-translational modifications". In: *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1819.9-10 (2012), pp. 1055–1066. ISSN: 18749399. DOI: 10.1016/j.bbagrm.2012.03.003. URL: http://dx.doi.org/10.1016/j.bbagrm.2012.03.003.

[17] Shimin Zhao et al. "Regulation of Cellular Metabolism by Protein Lysine Acetylation". In: *Science* 327.5968 (Feb. 2010), pp. 1000–1004. ISSN: 0036-8075. DOI: 10.1126/science.1179689. URL: https://www.science.org/doi/10.1126/science.1179689.

[18] Yan Xu and Kuo-Chen Chou. "Recent Progress in Predicting Posttranslational Modification Sites in Proteins". In: *Current Topics in Medicinal Chemistry* 16.6 (2015), pp. 591–603. ISSN: 15680266. DOI: 10.2174/1568026615666150819110421.

[19] Shahin Ramazi and Javad Zahiri. "Post-translational modifications in proteins: Resources, tools and prediction methods". In: *Database* 2021.7 (2021), pp. 1–20. ISSN: 17580463. DOI: 10.1093/database/baab012.

[20] George A. Khoury, Richard C. Baliban, and Christodoulos A. Floudas. "Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database". In: *Scientific Reports* 1 (2011), pp. 1–5. ISSN: 20452322. DOI: 10.1038/srep00090.

[21] Xiang Jiao Yang. "Multisite protein modification and intramolecular signaling". In: *Oncogene* 24.10 (2005), pp. 1653–1662. ISSN: 09509232. DOI: 10.1038/sj.onc.1208173.

[22] David W. Meek and Carl W. Anderson. "Posttranslational modification of p53: cooperative integrators of function." In: *Cold Spring Harbor perspectives in biology* 1.6 (2009), pp. 1–16. ISSN: 19430264. DOI: 10.1101/cshperspect.a000950.

[23] Diane P. Hanger et al. "Novel phosphorylation sites in Tau from Alzheimer brain support a role for casein kinase 1 in disease pathogenesis". In: *Journal of Biological Chemistry* 282.32 (2007), pp. 23645–23654. ISSN: 00219258. DOI: 10.1074/jbc.M703269200. URL: http://dx.doi.org/10.1074/jbc.M703269200.

[24] Alexander Stepanov et al. "A mathematical model of multisite phosphorylation of tau protein". In: *PLoS ONE* 13.2 (2018), pp. 1–17. ISSN: 19326203. DOI: 10.1371/journal.pone.0192519.

[25] Carolina Alquezar, Shruti Arya, and Aimee W. Kao. "Tau Post-translational Modifications: Dynamic Transformers of Tau Function, Degradation, and Aggregation". In: *Frontiers in Neurology* 11.January (2021), pp. 1–24. ISSN: 16642295. DOI: 10.3389/fneur.2020.595532.

[26] Thapanar Suwanmajo and J. Krishnan. "Mixed mechanisms of multi-site phosphorylation". In: *Journal of the Royal Society Interface* 12.107 (2015). ISSN: 17425662. DOI: 10.1098/rsif.2014.1405.

[27] Nick I. Markevich, Jan B. Hoek, and Boris N. Kholodenko. "Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades". In: *Journal of Cell Biology* 164.3 (2004), pp. 353–359. ISSN: 00219525. DOI: 10.1083/jcb.200308060.

[28] Carla M. Kumbale, Eberhard O. Voit, and Qiang Zhang. "Emergence and enhancement of ultrasensitivity through posttranslational modulation of protein stability". In: *Biomolecules* 11.11 (2021). ISSN: 2218273X. DOI: 10.3390/biom11111741.

[29] Matthew Thomson and Jeremy Gunawardena. "Unlimited multistability in multisite phosphorylation systems". In: *Nature* 460.7252 (2009), pp. 274–277. ISSN: 00280836. DOI: 10.1038/nature08102. URL: http://dx.doi.org/10.1038/nature08102.

[30] M. J. Rust et al. "Ordered phosphorylation governs oscillation of a three-protein orcadian clock". In: *Chemtracts* 20.10 (2007), pp. 410–411. ISSN: 14319268.

[31] Carsten Conradi and Anne Shiu. "Dynamics of Posttranslational Modification Systems: Recent Progress and Future Directions". In: *Biophysical Journal* 114.3 (2018), pp. 507–515. ISSN: 15420086. DOI: 10.1016/j.bpj.2017.11.3787. URL: https://doi.org/10.1016/j.bpj.2017.11.3787.

[32] Robert Borotkanics and Harold Lehmann. "Network motifs that recur across species, including gene regulatory and protein–protein interaction networks". In: *Archives of Toxicology* 89.4 (2015), pp. 489–499. ISSN: 14320738. DOI: 10.1007/s00204-014-1274-y.

[33] Esti Yeger-Lotem et al. “Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction”. In: *Proceedings of the National Academy of Sciences of the United States of America* 101.16 (2004), pp. 5934–5939. ISSN: 00278424. DOI: 10.1073/pnas.0306752101.

[34] Oren Shoval and Uri Alon. “SnapShot: Network Motifs”. In: *Cell* 143.2 (2010), pp. 326–326. ISSN: 00928674. DOI: 10.1016/j.cell.2010.09.050. URL: http://dx.doi.org/10.1016/j.cell.2010.09.050.

[35] Uri Alon. “Network motifs: Theory and experimental approaches”. In: *Nature Reviews Genetics* 8.6 (2007), pp. 450–461. ISSN: 14710056. DOI: 10.1038/nrg2102.

[36] Alex K. Shalek et al. “Single-cell RNA-seq reveals dynamic paracrine control of cellular variation”. In: *Nature* 510.7505 (2014), pp. 363–369. ISSN: 14764687. DOI: 10.1038/nature13437.

[37] Wuming Gong et al. “Inferring dynamic gene regulatory networks in cardiac differentiation through the integration of multi-dimensional data”. In: *BMC Bioinformatics* 16.1 (2015), pp. 1–16. ISSN: 14712105. DOI: 10.1186/s12859-015-0460-0.

[38] John J. Tyson and Béla Novák. “Functional Motifs in Biochemical Reaction Networks”. In: *Annual Review of Physical Chemistry* 61.1 (Mar. 2010), pp. 219–240. ISSN: 0066-426X. DOI: 10.1146/annurev.physchem.012809.103457. URL: https://www.annualreviews.org/doi/10.1146/annurev.physchem.012809.103457.

[39] Lorenza Trabalzini and John M Walker. *Ras Signaling IN Series Editor*. ISBN: 978-1-62703-790-7. DOI: 10.1007/978-1-62703-791-4.

[40] Zhong Lin Yang et al. “Inositol polyphosphate multikinase IPMK-1 regulates development through IP3/calcium signaling in Caenorhabditis elegans”. In: *Cell Calcium* 93.December 2020 (2021), p. 102327. ISSN: 15321991. DOI: 10.1016/j.ceca.2020.102327. URL: https://doi.org/10.1016/j.ceca.2020.102327.

[41] Evren U. Azeloglu and Ravi Iyengar. “Signaling networks: Information flow, computation, and decision making”. In: *Cold Spring Harbor Perspectives in Biology* 7.4 (2015), pp. 1–13. ISSN: 19430264. DOI: 10.1101/cshperspect.a005934.

[42] Gábor Balázsi, Alexander van Oudenaarden, and James J. Collins. “Cellular Decision Making and Biological Noise: From Microbes to Mammals”. In: *Cell* 144.6 (Mar. 2011), pp. 910–925. ISSN: 00928674. DOI: 10.1016/j.cell.2011.01.030. URL: https://linkinghub.elsevier.com/retrieve/pii/S0092867411000699.

[43] Abhishek Mallela, Maulik K. Nariya, and Eric J. Deeds. “Crosstalk and ultrasensitivity in protein degradation pathways”. In: *PLoS Computational Biology* 16.12 December (2020), pp. 1–18. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1008492. URL: http://dx.doi.org/10.1371/journal.pcbi.1008492.

[44] Govind Menon, Chinedu Okeke, and J. Krishnan. “Modelling compartmentalization towards elucidation and engineering of spatial organization in biochem-

ical pathways". In: *Scientific Reports* 7.1 (2017), pp. 1–18. ISSN: 20452322. DOI: 10.1038/s41598-017-11081-8. URL: http://dx.doi.org/10.1038/s41598-017-11081-8.

[45] Wooyoung Kim et al. "Biological network motif detection and evaluation". In: *BMC Systems Biology* 5.SUPPL. 3 (2011). ISSN: 17520509. DOI: 10.1186/1752-0509-5-S3-S5.

[46] Arun S. Konagurthu and Arthur M. Lesk. "On the origin of distribution patterns of motifs in biological networks". In: *BMC Systems Biology* 2 (2008), pp. 7–11. ISSN: 17520509. DOI: 10.1186/1752-0509-2-73.

[47] Arun S Konagurthu and Arthur M Lesk. "Single and multiple input modules in regulatory networks". In: *Proteins: Structure, Function, and Bioinformatics* 73.2 (Apr. 2008), pp. 320–324. ISSN: 08873585. DOI: 10.1002/prot.22053. URL: https://onlinelibrary.wiley.com/doi/10.1002/prot.22053.

[48] Piers J. Ingram, Michael P.H. Stumpf, and Jaroslav Stark. "Network motifs: Structure does not determine function". In: *BMC Genomics* 7 (2006), pp. 1–12. ISSN: 14712164. DOI: 10.1186/1471-2164-7-108.

[49] Sandra L. Harris and Arnold J. Levine. "The p53 pathway: Positive and negative feedback loops". In: *Oncogene* 24.17 (2005), pp. 2899–2908. ISSN: 09509232. DOI: 10.1038/sj.onc.1208615.

[50] Benjamin Pfeuty and Kunihiko Kaneko. "The combination of positive and negative feedback loops confers exquisite flexibility to biochemical switches". In: *Physical Biology* 6.4 (2009). ISSN: 14783975. DOI: 10.1088/1478-3975/6/4/046013.

[51] Huizhong Hu et al. "Feedforward and feedback regulation of the MAPK and PI3K oscillatory circuit in breast cancer". In: *Cellular Signalling* 25.1 (2013), pp. 26–32. ISSN: 08986568. DOI: 10.1016/j.cellsig.2012.09.014. URL: http://dx.doi.org/10.1016/j.cellsig.2012.09.014.

[52] John J. Tyson, Katherine C. Chen, and Bela Novak. "Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell". In: *Current Opinion in Cell Biology* 15.2 (2003), pp. 221–231. ISSN: 09550674. DOI: 10.1016/S0955-0674(03)00017-6.

[53] Qiang Zhang et al. "Molecular Signaling Network Motifs Provide a Mechanistic Basis for Cellular Threshold Responses". In: *Environmental Health Perspectives* 122.12 (Dec. 2014), pp. 1261–1270. ISSN: 0091-6765. DOI: 10.1289/ehp.1408244. URL: https://ehp.niehs.nih.gov/doi/10.1289/ehp.1408244.

[54] Wenying Shou, Sri Ram, and Jose M.G. Vilar. "Synthetic cooperation in engineered yeast populations". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.6 (2007), pp. 1877–1882. ISSN: 00278424. DOI: 10.1073/pnas.0610575104.

[55] Taliman Afroz and Chase L. Beisel. "Understanding and exploiting feedback in synthetic biology". In: *Chemical Engineering Science* 103 (2013), pp. 79–90. ISSN:

00092509. DOI: 10.1016/j.ces.2013.02.017. URL: http://dx.doi.org/10.1016/j.ces.2013.02.017.

[56] Lingchong You et al. "Programmed population control by cell-cell communication and regulated killing". In: *Nature* 428.6985 (2004), pp. 868–871. ISSN: 00280836. DOI: 10.1038/nature02491.

[57] Frederick K. Balagaddé et al. "A synthetic Escherichia coli predator-prey ecosystem". In: *Molecular Systems Biology* 4.187 (2008), pp. 1–8. ISSN: 17444292. DOI: 10.1038/msb.2008.24.

[58] Sarah M. Webster et al. "Programmed population controlby cell–cell communicationand regulated killing". In: *Nature* 428.6985 (2004), pp. 868–871. ISSN: 00280836. DOI: 10.1038/nature02468.1..

[59] Hao Song et al. "Programming microbial population dynamics by engineered cell-cell communication". In: *Biotechnology Journal* 6.7 (2011), pp. 837–849. ISSN: 18606768. DOI: 10.1002/biot.201100132.

[60] William Bacchus et al. "Synthetic two-way communication between mammalian cells". In: *Nature Biotechnology* 30.10 (2012), pp. 991–996. ISSN: 10870156. DOI: 10.1038/nbt.2351.

[61] Wilfried Weber et al. "A synthetic metabolite-based mammalian inter-cell signaling system". In: *Molecular BioSystems* 5.7 (2009), pp. 757–763. ISSN: 1742206X. DOI: 10.1039/b902070p.

[62] Yitong Ma et al. "Synthetic mammalian signaling circuits for robust cell population control". In: *Cell* 185.6 (2022), pp. 967–979. ISSN: 10974172. DOI: 10.1016/j.cell.2022.01.026.

[63] Paweł Kocieniewski, James R. Faeder, and Tomasz Lipniacki. "The interplay of double phosphorylation and scaffolding in MAPK pathways". In: *Journal of Theoretical Biology* 295 (2012), pp. 116–124. ISSN: 00225193. DOI: 10.1016/j.jtbi.2011.11.014.

[64] Rainis Venta et al. "A processive phosphorylation circuit with multiple kinase inputs and mutually diversional routes controls G1/S decision". In: *Nature Communications* 11.1 (2020). ISSN: 20411723. DOI: 10.1038/s41467-020-15685-z.

[65] Evin Valk et al. "Multistep phosphorylation systems: Tunable components of biological signaling circuits". In: *Molecular Biology of the Cell* 25.22 (2014), pp. 3456–3460. ISSN: 19394586. DOI: 10.1091/mbc.E14-02-0774.

[66] Philip Lössl et al. "Deciphering the interplay among multisite phosphorylation, interaction dynamics, and conformational transitions in a tripartite protein system". In: *ACS Central Science* 2.7 (2016), pp. 445–455. ISSN: 23747951. DOI: 10.1021/acscentsci.6b00053.

[67] Andre Levchenko, Jehoshua Bruck, and Paul W. Sternberg. "Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties". In: *Proceedings of the National Academy of*

*Sciences of the United States of America* 97.11 (2000), pp. 5818–5823. ISSN: 00278424. DOI: 10.1073/pnas.97.11.5818.

[68] Jan Philipp Reising et al. "Prostaglandin E2 Exerts Biphasic Dose Response on the PreBötzinger Complex Respiratory-Related Rhythm". In: *Frontiers in Neural Circuits* 16.May (2022), pp. 1–15. ISSN: 16625110. DOI: 10.3389/fncir.2022.826497.

[69] Stefan Hoops et al. "COPASI - A COmplex PAthway SImulator". In: *Bioinformatics* 22.24 (2006), pp. 3067–3074. ISSN: 13674803. DOI: 10.1093/bioinformatics/btl485.

[70] Lawrence F Shampine and Mark W Reichelt. "The matlab ode suite". In: *SIAM journal on scientific computing* 18.1 (1997), pp. 1–22.

[71] A. Dhooge, W. Govaerts, and Yu A. Kuznetsov. "MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs". In: *ACM Transactions on Mathematical Software* 29.2 (2003), pp. 141–164. ISSN: 00983500. DOI: 10.1145/779359.779362.

[72] W Govaerts et al. "MATCONT: Continuation toolbox for ODEs in Matlab". In: August (2017), p. 111.

[73] Maplesoft - a division of Waterloo Maple Inc. *Maple*. 2018. URL: https://www.maplesoft.com/support/help/maple/view.aspx?path=author.

[74] Carsten Conradi and Maya Mincheva. "Catalytic constants enable the emergence of bistability in dual phosphorylation". In: *Journal of the Royal Society Interface* 11.95 (2014). ISSN: 17425662. DOI: 10.1098/rsif.2014.0158.

[75] Boris Y. Rubinstein et al. "Long-Term dynamics of multisite phosphorylation". In: *Molecular Biology of the Cell* 27.14 (2016), pp. 2331–2340. ISSN: 19394586. DOI: 10.1091/mbc.E16-03-0137.

[76] Thapanar Suwanmajo, Vaidhiswaran Ramesh, and J. Krishnan. "Exploring cyclic networks of multisite modification reveals origins of information processing characteristics". In: *Scientific Reports* 10.1 (2020), pp. 1–18. ISSN: 20452322. DOI: 10.1038/s41598-020-73045-9. URL: https://doi.org/10.1038/s41598-020-73045-9.

[77] Thapanar Suwanmajo and J. Krishnan. "Biphasic responses in multi-site phosphorylation systems". In: *Journal of the Royal Society Interface* 10.89 (2013). ISSN: 17425662. DOI: 10.1098/rsif.2013.0742.

[78] Thapanar Suwanmajo and J. Krishnan. "Exploring the intrinsic behaviour of multisite phosphorylation systems as part of signalling pathways". In: *Journal of the Royal Society Interface* 15.143 (2018). ISSN: 17425662. DOI: 10.1098/rsif.2018.0109.

[79] Carsten Conradi et al. "Identifying parameter regions for multistationarity". In: *PLoS Computational Biology* 13.10 (2017), pp. 1–27. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1005751.

[80] Mitchell Eithun and Anne Shiu. "An all-encompassing global convergence result for processive multisite phosphorylation systems". In: *Mathematical Bio-*

*sciences* 291 (2017), pp. 1–9. ISSN: 18793134. DOI: 10.1016/j.mbs.2017.05.006. URL: http://dx.doi.org/10.1016/j.mbs.2017.05.006.

[81] Hwai Ray Tung. "Precluding oscillations in Michaelis–Menten approximations of dual-site phosphorylation systems". In: *Mathematical Biosciences* 306.August (2018), pp. 56–59. ISSN: 18793134. DOI: 10.1016/j.mbs.2018.10.008. URL: https://doi.org/10.1016/j.mbs.2018.10.008.

[82] Nicholas A. Lyons et al. "Sequential primed kinases create a damage-responsive phosphodegron on Eco1". In: *Nature Structural and Molecular Biology* 20.2 (2013), pp. 194–201. ISSN: 15459993. DOI: 10.1038/nsmb.2478.

[83] Ellen C. OShaughnessy et al. "Tunable signal processing in synthetic MAP kinase cascades". In: *Cell* 144.1 (2011), pp. 119–131. ISSN: 00928674. DOI: 10.1016/j.cell.2010.12.014. URL: http://dx.doi.org/10.1016/j.cell.2010.12.014.

[84] Oliver R. Maguire and Wilhelm T.S. Huck. "On the importance of reaction networks for synthetic living systems". In: *Emerging Topics in Life Sciences* 3.5 (2019), pp. 517–527. ISSN: 2397-8554. DOI: 10.1042/etls20190016.

[85] Jeremy B.A. Green and James Sharpe. "Positional information and reaction-diffusion: Two big ideas in developmental biology combine". In: *Development (Cambridge)* 142.7 (2015), pp. 1203–1211. ISSN: 14779129. DOI: 10.1242/dev.114991.

[86] Philip K. Maini et al. "Turing's model for biological pattern formation and the robustness problem". In: *Interface Focus* 2.4 (2012), pp. 487–496. ISSN: 20428901. DOI: 10.1098/rsfs.2011.0113.

[87] Koh Onimaru et al. "The fin-to-limb transition as the re-organization of a Turing pattern". In: *Nature Communications* 7.May (2016). ISSN: 20411723. DOI: 10.1038/ncomms11582.

[88] Anum Khan, Molly McQuilken, and Amy S. Gladfelter. "Septins and Generation of Asymmetries in Fungal Cells". In: *Annual Review of Microbiology* 69.1 (2015), pp. 487–503. ISSN: 15453251. DOI: 10.1146/annurev-micro-091014-104250.

[89] Fei Wang. "The signaling mechanisms underlying cell polarity and chemotaxis." In: *Cold Spring Harbor perspectives in biology* 1.4 (2009), pp. 1–16. ISSN: 19430264. DOI: 10.1101/cshperspect.a002980.

[90] Jean Pierre Changeux. "Allostery and the monod-wyman-changeux model after 50 years". In: *Annual Review of Biophysics* 41.1 (2012), pp. 103–133. ISSN: 1936122X. DOI: 10.1146/annurev-biophys-050511-102222.

[91] István Hargittai and Magdolna Hargittai. *Symmetry: A unifying concept*. Shelter Publications, Inc., 1994.

[92] David Hochberg et al. "Stoichiometric network analysis of spontaneous mirror symmetry breaking in chemical reactions". In: *Physical Chemistry Chemical Physics* 19.27 (2017), pp. 17618–17636. ISSN: 1463-9076. DOI: 10.1039/C7CP02159C. URL: http://xlink.rsc.org/?DOI=C7CP02159C.

[93] Josep M. Ribó et al. "Spontaneous mirror symmetry breaking and origin of biological homochirality". In: *Journal of the Royal Society Interface* 14.137 (2017). ISSN: 17425662. DOI: 10.1098/rsif.2017.0699.

[94] Uri Hananel et al. "Spontaneous and directed symmetry breaking in the formation of chiral nanocrystals". In: *Proceedings of the National Academy of Sciences of the United States of America* 166.23 (2019), pp. 11159–11164. ISSN: 10916490. DOI: 10.1073/pnas.1821923116.

[95] Yo Kushida et al. "Deterministic and Stochastic Chiral Symmetry Breaking Exhibited by Racemic Aminomethylenehelicene Oligomers". In: *Chemistry - A European Journal* 23.2 (2017), pp. 327–333. ISSN: 15213765. DOI: 10.1002/chem.201601792.

[96] Kenso Soai et al. "Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule". In: *Nature* 378.6559 (Dec. 1995), pp. 767–768. ISSN: 0028-0836. DOI: 10.1038/378767a0. URL: http://www.nature.com/articles/378767a0.

[97] Yong Chen and Wentao Ma. "The origin of biological homochirality along with the origin of life". In: *PLoS Computational Biology* 16.1 (2020), pp. 1–20. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1007592.

[98] Donna G. Blackmond. "Autocatalytic Models for the Origin of Biological Homochirality". In: *Chemical Reviews* 120.11 (2020), pp. 4831–4847. ISSN: 15206890. DOI: 10.1021/acs.chemrev.9b00557.

[99] Craig C. Jolley, Koji L. Ode, and Hiroki R. Ueda. "A Design Principle for a Posttranslational Biochemical Oscillator". In: *Cell Reports* 2.4 (2012), pp. 938–950. ISSN: 22111247. DOI: 10.1016/j.celrep.2012.09.006. URL: http://dx.doi.org/10.1016/j.celrep.2012.09.006.

[100] Paweł Kocieniewski, James R. Faeder, and Tomasz Lipniacki. "The interplay of double phosphorylation and scaffolding in MAPK pathways". In: *Journal of Theoretical Biology* 295 (2012), pp. 116–124. ISSN: 00225193. DOI: 10.1016/j.jtbi.2011.11.014.

[101] Sina Ghaemmaghami et al. "Global analysis of protein expression in yeast". In: *Nature* 425.6959 (2003), pp. 737–741. ISSN: 00280836. DOI: 10.1038/nature02046.

[102] Koji L. Ode and Hiroki R. Ueda. "Design principles of phosphorylation-dependent timekeeping in eukaryotic circadian clocks". In: *Cold Spring Harbor Perspectives in Biology* 10.8 (2018). ISSN: 19430264. DOI: 10.1101/cshperspect.a028357.

[103] Yuta Shinohara et al. "Temperature-Sensitive Substrate and Product Binding Underlie Temperature-Compensated Phosphorylation in the Clock". In: *Molecular Cell* 67.5 (2017), pp. 783–798. ISSN: 10974164. DOI: 10.1016/j.molcel.2017.08.009. URL: http://dx.doi.org/10.1016/j.molcel.2017.08.009.

[104] Tetsuhiro S. Hatakeyama and Kunihiko Kaneko. "Generic temperature compensation of biological clocks by autonomous regulation of catalyst concentration". In: *Proceedings of the National Academy of Sciences of the United States of*

*America* 109.21 (2012), pp. 8109–8114. ISSN: 00278424. DOI: 10.1073/pnas.1120711109.

[105] Shuhei S. Sugai, Koji L. Ode, and Hiroki R. Ueda. "A Design Principle for an Autonomous Post-translational Pattern Formation". In: *Cell Reports* 19.4 (2017), pp. 863–874. ISSN: 22111247. DOI: 10.1016/j.celrep.2017.03.081. URL: http://dx.doi.org/10.1016/j.celrep.2017.03.081.

[106] Steven Henry Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. eng. 2. print. Cambridge, Mass: Perseus Books, 2001, p. 498. ISBN: 9780738204536.

[107] Kazuhiro Aoki et al. "Processive phosphorylation of ERK MAP kinase in mammalian cells". In: *Proceedings of the National Academy of Sciences* 108.31 (Aug. 2011), pp. 12675–12680. ISSN: 0027-8424. DOI: 10.1073/pnas.1104030108. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.1104030108.

[108] Kazuhiro Aoki et al. "A Quantitative Model of ERK MAP Kinase Phosphorylation in Crowded Media". In: *Scientific Reports* 3.1 (Dec. 2013), p. 1541. ISSN: 2045-2322. DOI: 10.1038/srep01541. URL: http://www.nature.com/articles/srep01541.

[109] Ildar I. Sadreev et al. "A Systems Model of Phosphorylation for Inflammatory Signaling Events". In: *PLoS ONE* 9.10 (Oct. 2014). Ed. by Gautam Sethi, e110913. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0110913. URL: https://dx.plos.org/10.1371/journal.pone.0110913.

[110] German A. Enciso and Shane Ryerson. "The effect of site-to-site variability in ultrasensitive dose responses". In: *Journal of Mathematical Biology* 74.1-2 (2017), pp. 23–41. ISSN: 14321416. DOI: 10.1007/s00285-016-1013-5.

[111] Chidambaram Ramachandran et al. "Sequential Dephosphorylation of a Multiply Phosphorylated Insulin Receptor Peptide by Protein Tyrosine Phosphatases". In: *Biochemistry* 31.17 (1992), pp. 4232–4238. ISSN: 15204995. DOI: 10.1021/bi00132a012.

[112] Mardo Kõivomägi et al. "Multisite phosphorylation networks as signal processors for Cdk1". In: *Nature Structural and Molecular Biology* 20.12 (2013), pp. 1415–1424. ISSN: 15459993. DOI: 10.1038/nsmb.2706.

[113] Eric Batchelor and Mark Goulian. "Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system". In: *Proceedings of the National Academy of Sciences* 100.2 (Jan. 2003), pp. 691–696. ISSN: 0027-8424. DOI: 10.1073/pnas.0234782100. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.0234782100.

[114] J. Krishnan, Lingjun Lu, and Aiman Alam Nazki. "The interplay of spatial organization and biochemistry in building blocks of cellular signalling pathways". In: *Journal of the Royal Society Interface* 17.166 (2020). ISSN: 17425662. DOI: 10.1098/rsif.2020.0251.

[115] Guy Shinar and Martin Feinberg. "Design principles for robust biochemical reaction networks: What works, what cannot work, and what might almost

work". In: *Mathematical Biosciences* 231.1 (May 2011), pp. 39–48. ISSN: 00255564. DOI: 10.1016/j.mbs.2011.02.012. URL: http://dx.doi.org/10.1016/j.mbs.2011.02.012%20https://linkinghub.elsevier.com/retrieve/pii/S0025556411000307.

[116] J. Krishnan and Ioannis Floros. "Adaptive information processing of network modules to dynamic and spatial stimuli". In: *BMC Systems Biology* (2019). ISSN: 17520509. DOI: 10.1186/s12918-019-0703-1.

[117] Alejandra C. Ventura et al. "Signaling properties of a covalent modification cycle are altered by a downstream target". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.22 (2010), pp. 10032–10037. ISSN: 00278424. DOI: 10.1073/pnas.0913815107.

[118] Domitilla Del Vecchio, Alexander J. Ninfa, and Eduardo D. Sontag. "Modular cell biology: Retroactivity and insulation". In: *Molecular Systems Biology* 4.161 (2008). ISSN: 17444292. DOI: 10.1038/msb4100204.

[119] Supriya Krishnamurthy et al. "The stochastic behavior of a molecular switching circuit with feedback". In: *Biology Direct* 2 (2007). ISSN: 17456150. DOI: 10.1186/1745-6150-2-13.

[120] Jeroen S. Van Zon et al. "An allosteric model of circadian KaiC phosphorylation". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.18 (2007), pp. 7420–7425. ISSN: 00278424. DOI: 10.1073/pnas.0608665104.

[121] Bruce L Martin. "Regulation by Covalent Modification". In: *eLS*. Wiley, Aug. 2014. ISBN: 9780470015902. DOI: 10.1002/9780470015902.a0000866.pub3. URL: https://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0000866.pub3.

[122] Thapanar Suwanmajo and J. Krishnan. "Biphasic responses in multi-site phosphorylation systems". In: *Journal of the Royal Society Interface* 10.89 (2013). ISSN: 17425662. DOI: 10.1098/rsif.2013.0742.

[123] Franziska Witzel and Nils Blüthgen. "When More Is Less: Dual Phosphorylation Protects Signaling Off State against Overexpression". In: *Biophysical Journal* 115.7 (2018), pp. 1383–1392. ISSN: 15420086. DOI: 10.1016/j.bpj.2018.08.019.

[124] Julia Kamenz, Lendert Gelens, and James E. Ferrell. "Bistable, Biphasic Regulation of PP2A-B55 Accounts for the Dynamics of Mitotic Substrate Phosphorylation". In: *Current Biology* 31.4 (2021), pp. 794–808. ISSN: 18790445. DOI: 10.1016/j.cub.2020.11.058. URL: https://doi.org/10.1016/j.cub.2020.11.058.

[125] Barbara Szomolay and Vahid Shahrezaei. "Bell-shaped and ultrasensitive dose-response in phosphorylation-dephosphorylation cycles: the role of kinase-phosphatase complex formation". In: *BMC Systems Biology* 6 (2012). ISSN: 17520509. DOI: 10.1186/1752-0509-6-26.

[126] Dongsan Kim, Yung Keun Kwon, and Kwang Hyun Cho. "The biphasic behavior of incoherent feed-forward loops in biomolecular regulatory networks". In:

*BioEssays* 30.11-12 (2008), pp. 1204–1211. ISSN: 02659247. DOI: 10.1002/bies.20839.

[127] Liu Yang and Pablo A. Iglesias. "Positive feedback may cause the biphasic response observed in the chemoattractant-induced response of Dictyostelium cells". In: *Systems and Control Letters* 55.4 (2006), pp. 329–337. ISSN: 01676911. DOI: 10.1016/j.sysconle.2005.08.009.

[128] Thawfeek M. Varusai et al. "Protein-protein interactions generate hidden feedback and feed-forward loops to trigger bistable switches, oscillations and biphasic dose-responses". In: *Molecular BioSystems* 11.10 (2015), pp. 2750–2762. ISSN: 17422051. DOI: 10.1039/c5mb00385g.

[129] Stefan Hoops et al. "COPASI–a COmplex PAthway SImulator". In: *Bioinformatics* 22.24 (Dec. 2006), pp. 3067–3074. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btl485. URL: https://dx.doi.org/10.1093/bioinformatics/btl485%20https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btl485.

[130] A. Harvey Millar et al. "The Scope, Functions, and Dynamics of Posttranslational Protein Modifications". In: *Annual Review of Plant Biology* 70 (2019), pp. 119–151. ISSN: 15452123. DOI: 10.1146/annurev-arplant-050718-100211.

[131] Mardo Köivomägi et al. "Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase". In: *Nature* 480.7375 (2011), pp. 128–131. ISSN: 00280836. DOI: 10.1038/nature10560.

[132] Jorge E. Allende and Catherine C. Allende. "Protein kinase CK2: an enzyme with multiple substrates and a puzzling regulation". In: *The FASEB Journal* 9.5 (1995), pp. 313–323. ISSN: 0892-6638. DOI: 10.1096/fasebj.9.5.7896000.

[133] Milad Ghomlaghi et al. "Feedback, crosstalk and competition: Ingredients for emergent non-linear behaviour in the pi3k/mtor signalling network". In: *International Journal of Molecular Sciences* 22.13 (2021). ISSN: 14220067. DOI: 10.3390/ijms22136944.

[134] Erin M. Shockley et al. "Signal integration and information transfer in an allosterically regulated network". In: *npj Systems Biology and Applications* 5.1 (2019). ISSN: 20567189. DOI: 10.1038/s41540-019-0100-9. URL: http://dx.doi.org/10.1038/s41540-019-0100-9.

[135] Daniel D. Seaton and J. Krishnan. "The coupling of pathways and processes through shared components". In: *BMC Systems Biology* 5 (2011). ISSN: 17520509. DOI: 10.1186/1752-0509-5-103.

[136] James E. Ferrell et al. "Simple, realistic models of complex biological processes: Positive feedback and bistability in a cell fate switch and a cell cycle oscillator". In: *FEBS Letters* 583.24 (2009), pp. 3999–4005. ISSN: 00145793. DOI: 10.1016/j.febslet.2009.10.068.

[137] H. Frederik Nijhout, Janet A. Best, and Michael C. Reed. "Systems biology of robustness and homeostatic mechanisms". In: *Wiley Interdisciplinary Reviews: Sys-*

*tems Biology and Medicine* 11.3 (2019), pp. 1–23. ISSN: 1939005X. DOI: 10.1002/wsbm.1440.

[138] Subhayu Basu et al. "Spatiotemporal control of gene expression with pulse-generating networks". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.17 (2004), pp. 6355–6360. ISSN: 00278424. DOI: 10.1073/pnas.0307571101.

[139] Britta Helwig et al. "Bottom-Up Construction of an Adaptive Enzymatic Reaction Network". In: *Angewandte Chemie - International Edition* 57.43 (2018), pp. 14065–14069. ISSN: 15213773. DOI: 10.1002/anie.201806944.

[140] Evgeny Katz. *Molecular and Supramolecular Information Processing*. Ed. by Evgeny Katz. Wiley, July 2012. ISBN: 9783527331956. DOI: 10.1002/9783527645442. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9783527645442.

[141] Supriya Krishnamurthy et al. "The stochastic behavior of a molecular switching circuit with feedback". In: *Biology Direct* 2 (2007). ISSN: 17456150. DOI: 10.1186/1745-6150-2-13.

[142] Siobhan S. McMahon et al. "Information processing by simple molecular motifs and susceptibility to noise". In: *Journal of the Royal Society Interface* 12.110 (2015). ISSN: 17425662. DOI: 10.1098/rsif.2015.0597.

[143] Amruta Sahoo and Sebastian Pechmann. "Functional network motifs defined through integration of protein-protein and genetic interactions". In: (2021), pp. 1–35.

[144] Carlos Salazar and Thomas Höfer. "Multisite protein phosphorylation - From molecular mechanisms to kinetic models". In: *FEBS Journal* 276.12 (2009), pp. 3177–3198. ISSN: 1742464X. DOI: 10.1111/j.1742-4658.2009.07027.x.

[145] Andre Levchenko, Jehoshua Bruck, and Paul W. Sternberg. "Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties". In: *Proceedings of the National Academy of Sciences of the United States of America* 97.11 (2000), pp. 5818–5823. ISSN: 00278424. DOI: 10.1073/pnas.97.11.5818.

[146] Marek Kochańczyk et al. "Relaxation oscillations and hierarchy of feedbacks in MAPK signaling". In: *Scientific Reports* 7.January (2017), pp. 1–15. ISSN: 20452322. DOI: 10.1038/srep38244. URL: http://dx.doi.org/10.1038/srep38244.

[147] Jeremy Gunawardena. "Multisite protein phosphorylation makes a good threshold but can be a poor switch". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.41 (2005), pp. 14617–14622. ISSN: 00278424. DOI: 10.1073/pnas.0507322102.

[148] Vaidhiswaran Ramesh and J. Krishnan. "Symmetry breaking meets multisite modification". In: *eLife* 10 (2021), pp. 1–53. ISSN: 2050084X. DOI: 10.7554/eLife.65358.

[149] Liming Wang and Eduardo D. Sontag. "On the number of steady states in a multiple futile cycle". In: *Journal of Mathematical Biology* 57.1 (2008), pp. 29–52. ISSN: 03036812. DOI: 10.1007/s00285-007-0145-z.

[150] Stuart R. Jefferys and Morgan C. Giddings. "Baking a mass-spectrometry data PIE with McMC and simulated annealing: Predicting protein post-translational modifications from integrated top-down and bottom-up data". In: *Bioinformatics* 27.6 (2011), pp. 844–852. ISSN: 13674803. DOI: 10.1093/bioinformatics/btr027.

[151] Philippe P. Roux and Pierre Thibault. "The coming of age of phosphoproteomics-from large data sets to inference of protein functions". In: *Molecular and Cellular Proteomics* 12.12 (2013), pp. 3453–3464. ISSN: 15359476. DOI: 10.1074/mcp.R113.032862.

[152] Xiaojing Yang et al. "Design Principles of the Yeast G1/S Switch". In: *PLoS Biology* 11.10 (2013). ISSN: 15449173. DOI: 10.1371/journal.pbio.1001673.

[153] David W. Meek. "Multisite phosphorylation and the integration of stress signals at p53". In: *Cellular Signalling* 10.3 (1998), pp. 159–166. ISSN: 08986568. DOI: 10.1016/S0898-6568(97)00119-8.

[154] David W. Meek and Ted R. Hupp. "The regulation of MDM2 by multisite phosphorylation-Opportunities for molecular-based intervention to target tumours?" In: *Seminars in Cancer Biology* 20.1 (2010), pp. 19–28. ISSN: 1044579X. DOI: 10.1016/j.semcancer.2009.10.005. URL: http://dx.doi.org/10.1016/j.semcancer.2009.10.005.

[155] Anand Ponnuswamy, Ted Hupp, and Robin Fåhraeus. "Concepts in MDM2 Signaling: Allosteric Regulation and Feedback Loops". In: *Genes and Cancer* 3.3-4 (2012), pp. 291–297. ISSN: 19476027. DOI: 10.1177/1947601912454140.

[156] Ishrat Ahmed et al. "Huntington's disease: Neural dysfunction linked to inositol polyphosphate multikinase". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.31 (2015), pp. 9751–9756. ISSN: 10916490. DOI: 10.1073/pnas.1511810112.

[157] Ishrat Ahmed. "THE ROLE OF INOSITOL POLYPHOSPHATE MULTIKINASE IN HUNTINGTON ' S DISEASE by". In: May (2015).

[158] Zhilin Qu, James N. Weiss, and W. Robb MacLellan. "Regulation of the mammalian cell cycle: A model of the G1-to-S transition". In: *American Journal of Physiology - Cell Physiology* 284.2 53-2 (2003). ISSN: 03636143. DOI: 10.1152/ajpcell.00066.2002.

[159] Luyang Sun et al. "Substrate phosphorylation and feedback regulation in JFK-promoted p53 destabilization". In: *Journal of Biological Chemistry* 286.6 (2011), pp. 4226–4235. ISSN: 00219258. DOI: 10.1074/jbc.M110.195115. URL: http://dx.doi.org/10.1074/jbc.M110.195115.

[160] David W. Meek. "Regulation of the p53 response and its relationship to cancer". In: *Biochemical Journal* 469.3 (2015), pp. 325–346. ISSN: 14708728. DOI: 10.1042/BJ20150517.

[161] Amandine Chaix, Amir Zarrinpar, and Satchidananda Panda. "The circadian coordination of cell biology". In: *Journal of Cell Biology* 215.1 (2016), pp. 15–25. ISSN: 15408140. DOI: 10.1083/jcb.201603076.

[162] Carsten Conradi et al. "Identifying parameter regions for multistationarity". In: *PLoS Computational Biology* 13.10 (2017). ISSN: 15537358. DOI: 10.1371/journal.pcbi.1005751.

[163] Elisenda Feliu and Carsten Wiuf. "Enzyme-sharing as a cause of multi-stationarity in signalling systems". In: *Journal of the Royal Society Interface* 9.71 (2012), pp. 1224–1232. ISSN: 17425662. DOI: 10.1098/rsif.2011.0664.

[164] James E. Ferrell. "Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability". In: *Current Opinion in Cell Biology* 14.2 (2002), pp. 140–148. ISSN: 09550674. DOI: 10.1016/S0955-0674(02)00314-9.

[165] Hao Song et al. "Bifurcation and singularity analysis of a molecular network for the induction of long-term memory". In: *Biophysical Journal* 90.7 (2006), pp. 2309–2325. ISSN: 00063495. DOI: 10.1529/biophysj.105.074500. URL: http://dx.doi.org/10.1529/biophysj.105.074500.

[166] K. Sriram, Maria Rodriguez-Fernandez, and Francis J. Doyle. "A detailed modular analysis of heat-shock protein dynamics under acute and chronic stress and its implication in anxiety disorders". In: *PLoS ONE* 7.8 (2012). ISSN: 19326203. DOI: 10.1371/journal.pone.0042958.

[167] Sudhakaran Prabakaran, Jeremy Gunawardena, and Eduardo Sontag. "Paradoxical results in perturbation-based signaling network reconstruction". In: *Biophysical Journal* 106.12 (2014), pp. 2720–2728. ISSN: 15420086. DOI: 10.1016/j.bpj.2014.04.031. URL: http://dx.doi.org/10.1016/j.bpj.2014.04.031.

[168] Koji L Ode and Hiroki R Ueda. " Lost in clocks: non-canonical circadian oscillation discovered in Drosophila cells ". In: *Molecular Systems Biology* 14.9 (2018), pp. 2–4. ISSN: 1744-4292. DOI: 10.15252/msb.20188567.

[169] Serdar Özsezen et al. "Inference of the High-Level Interaction Topology between the Metabolic and Cell-Cycle Oscillators from Single-Cell Dynamics". In: *Cell Systems* 9.4 (2019), pp. 354–365. ISSN: 24054720. DOI: 10.1016/j.cels.2019.09.003.

[170] François Jacob. "Evolution and tinkering". In: *Science* 196.4295 (1977), pp. 1161–1166. ISSN: 00368075. DOI: 10.1126/science.860134.

[171] Abderahmane Derouiche et al. "Evolution and tinkering: what do a protein kinase, a transcriptional regulator and chromosome segregation/cell division proteins have in common?" In: *Current Genetics* 62.1 (2016), pp. 67–70. ISSN: 14320983. DOI: 10.1007/s00294-015-0513-y.

[172] Maaruthy Yelleswarapu et al. "Sigma Factor-Mediated Tuning of Bacterial Cell-Free Synthetic Genetic Oscillators". In: *ACS Synthetic Biology* 7.12 (2018), pp. 2879–2887. ISSN: 21615063. DOI: 10.1021/acssynbio.8b00300.

[173] Ronghui Zhu et al. “Synthetic multistability in mammalian cells”. In: *Science* 375.6578 (Jan. 2022). ISSN: 0036-8075. DOI: 10.1126/science.abg9765. URL: https://www.science.org/doi/10.1126/science.abg9765.

[174] Erick Armingol et al. “Deciphering cell–cell interactions and communication from gene expression”. In: *Nature Reviews Genetics* 22.2 (2021), pp. 71–88. ISSN: 14710064. DOI: 10.1038/s41576-020-00292-x. URL: http://dx.doi.org/10.1038/s41576-020-00292-x.

[175] M. Ricarda Wilson, Troy W. Close, and James E. Trosko. “Cell population dynamics (apoptosis, mitosis, and cell-cell communication) during disruption of homeostasis”. In: *Experimental Cell Research* 254.2 (2000), pp. 257–268. ISSN: 00144827. DOI: 10.1006/excr.1999.4771.

[176] Katja Kolar et al. “A synthetic mammalian network to compute population borders based on engineered reciprocal cell-cell communication”. In: *BMC Systems Biology* 9.1 (2015), pp. 1–8. ISSN: 17520509. DOI: 10.1186/s12918-015-0252-1. URL: http://dx.doi.org/10.1186/s12918-015-0252-1.

[177] William Bacchus and Martin Fussenegger. “Engineering of synthetic intercellular communication systems”. In: *Metabolic Engineering* 16.1 (2013), pp. 33–41. ISSN: 10967176. DOI: 10.1016/j.ymben.2012.12.001. URL: http://dx.doi.org/10.1016/j.ymben.2012.12.001.

[178] Andreia Carvalho et al. “Genetically encoded sender-receiver system in 3D mammalian cell culture”. In: *ACS Synthetic Biology* 3.5 (2014), pp. 264–272. ISSN: 21615063. DOI: 10.1021/sb400053b.

[179] Wei Dong Wang et al. “Construction of an artificial intercellular communication network using the nitric oxide signaling elements in mammalian cells”. In: *Experimental Cell Research* 314.4 (2008), pp. 699–706. ISSN: 00144827. DOI: 10.1016/j.yexcr.2007.11.023.

[180] Hao Song et al. “Synthetic microbial consortia: From systematic analysis to construction and applications”. In: *Chemical Society Reviews* 43.20 (2014), pp. 6954–6981. ISSN: 14604744. DOI: 10.1039/c4cs00114a.

[181] Ming Tang Chen and Ron Weiss. “Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana”. In: *Nature Biotechnology* 23.12 (2005), pp. 1551–1555. ISSN: 10870156. DOI: 10.1038/nbt1162.

[182] Jeffrey J. Tabor et al. “A Synthetic Genetic Edge Detection Program”. In: *Cell* 137.7 (2009), pp. 1272–1281. ISSN: 00928674. DOI: 10.1016/j.cell.2009.04.048. URL: http://dx.doi.org/10.1016/j.cell.2009.04.048.

[183] Robert D. Lovchik et al. “A microfluidic device for depositing and addressing two cell populations with intercellular population communication capability”. In: *Biomedical Microdevices* 12.2 (Apr. 2010), pp. 275–282. ISSN: 1387-2176. DOI: 10.1007/s10544-009-9382-9. URL: http://link.springer.com/10.1007/s10544-009-9382-9.

[184] Subhayu Basu et al. "A synthetic multicellular system for programmed pattern formation". In: *Nature* 434.7037 (2005), pp. 1130–1134. ISSN: 00280836. DOI: 10.1038/nature03461.

[185] Daniella M. Schwartz et al. "Type I/II cytokines, JAKs, and new strategies for treating autoimmune diseases". In: *Nature Reviews Rheumatology* 12.1 (2016), pp. 25–36. ISSN: 17594804. DOI: 10.1038/nrrheum.2015.167.

[186] K. G. Guruharsha, Mark W. Kankel, and Spyros Artavanis-Tsakonas. "The Notch signalling system: Recent insights into the complexity of a conserved pathway". In: *Nature Reviews Genetics* 13.9 (2012), pp. 654–666. ISSN: 14710056. DOI: 10.1038/nrg3272.

[187] Graça Raposo and Willem Stoorvogel. "Extracellular vesicles: Exosomes, microvesicles, and friends". In: *Journal of Cell Biology* 200.4 (2013), pp. 373–383. ISSN: 00219525. DOI: 10.1083/jcb.201211138.

[188] Gerd R. Burmester, Eugen Feist, and Thomas Dörner. "Emerging cell and cytokine targets in rheumatoid arthritis". In: *Nature Reviews Rheumatology* 10.2 (2014), pp. 77–88. ISSN: 17594790. DOI: 10.1038/nrrheum.2013.168.

[189] Sarah B. Schrier et al. "Synergistic Communication between CD4+ T Cells and Monocytes Impacts the Cytokine Environment". In: *Scientific Reports* 6.June (2016), pp. 1–11. ISSN: 20452322. DOI: 10.1038/srep34942.

[190] Qiong Xue et al. "Analysis of single-cell cytokine secretion reveals a role for paracrine signaling in coordinating macrophage responses to TLR4 stimulation". In: *Science Signaling* 8.381 (June 2015). ISSN: 1945-0877. DOI: 10.1126/scisignal.aaa2155. URL: https://www.science.org/doi/10.1126/scisignal.aaa2155.

[191] Frances R. Balkwill, Melania Capasso, and Thorsten Hagemann. "The tumor microenvironment at a glance". In: *Journal of Cell Science* 125.23 (2012), pp. 5591–5596. ISSN: 00219533. DOI: 10.1242/jcs.116392.

[192] Marilyn G. Farquhar and George E. Palade. "JUNCTIONAL COMPLEXES IN VARIOUS EPITHELIA". In: *Journal of Cell Biology* 17.2 (May 1963), pp. 375–412. ISSN: 1540-8140. DOI: 10.1083/jcb.17.2.375. URL: https://rupress.org/jcb/article/17/2/375/1226/JUNCTIONAL-COMPLEXES-IN-VARIOUS-EPITHELIA.

[193] M. Phipps et al. "Carcinogenic implications of the neighborhood coherence principle (NCP)". In: *Medical Hypotheses* 31.4 (1990), pp. 289–301. ISSN: 03069877. DOI: 10.1016/0306-9877(90)90021-6.

[194] Kevin Thurley, Lani F. Wu, and Steven J. Altschuler. "Modeling Cell-to-Cell Communication Networks Using Response-Time Distributions". In: *Cell Systems* 6.3 (2018), pp. 355–367. ISSN: 24054720. DOI: 10.1016/j.cels.2018.01.016. URL: http://dx.doi.org/10.1016/j.cels.2018.01.016.

[195] Tobias Großkopf and Orkun S. Soyer. "Synthetic microbial communities". In: *Current Opinion in Microbiology* 18.1 (2014), pp. 72–77. ISSN: 18790364. DOI: 10.

1016/j.mib.2014.02.002. URL: http://dx.doi.org/10.1016/j.mib.2014.02.002.

[196] Matteo Cavaliere et al. "Cooperation in microbial communities and their biotechnological applications". In: *Environmental Microbiology* 19.8 (2017), pp. 2949–2963. ISSN: 14622920. DOI: 10.1111/1462-2920.13767.

[197] J. Cremer et al. "Cooperation in Microbial Populations: Theory and Experimental Model Systems". In: *Journal of Molecular Biology* 431.23 (2019), pp. 4599–4644. ISSN: 10898638. DOI: 10.1016/j.jmb.2019.09.023. URL: https://doi.org/10.1016/j.jmb.2019.09.023.

[198] Jieun Choi and Yong Jung Kim. "Predator–prey equations with constant harvesting and planting". In: *Journal of Theoretical Biology* 458 (2018), pp. 47–57. ISSN: 10958541. DOI: 10.1016/j.jtbi.2018.08.044. URL: https://doi.org/10.1016/j.jtbi.2018.08.044.

[199] Vicente Aboites et al. "Interpretation and Dynamics of the Lotka–Volterra Model in the Description of a Three-Level Laser". In: *Photonics* 9.1 (2022). ISSN: 23046732. DOI: 10.3390/photonics9010016.

[200] Sougata Roy and Thomas B. Kornberg. "Paracrine signaling mediated at cell-cell contacts". In: *BioEssays* 37.1 (Jan. 2015), pp. 25–33. ISSN: 02659247. DOI: 10.1002/bies.201400122. URL: http://doi.wiley.com/10.1002/bies.201400122.

[201] Yiteng Dang, Douwe A.J. Grundel, and Hyun Youk. "Cellular Dialogues: Cell-Cell Communication through Diffusible Molecules Yields Dynamic Spatial Patterns". In: *Cell Systems* 10.1 (2020), pp. 82–98. ISSN: 24054720. DOI: 10.1016/j.cels.2019.12.001. URL: https://doi.org/10.1016/j.cels.2019.12.001.

[202] Govind Menon and J. Krishnan. "Design Principles for Compartmentalization and Spatial Organization of Synthetic Genetic Circuits". In: *ACS Synthetic Biology* 8.7 (2019), pp. 1601–1619. ISSN: 21615063. DOI: 10.1021/acssynbio.8b00522.

[203] James S. Clark et al. "Post-translational Modifications of the p53 Protein and the Impact in Alzheimer's Disease: A Review of the Literature". In: *Frontiers in Aging Neuroscience* 14.April (2022), pp. 1–11. ISSN: 16634365. DOI: 10.3389/fnagi.2022.835288.

[204] Marie Cargnello and Philippe P. Roux. "Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases". In: *Microbiology and Molecular Biology Reviews* 75.1 (2011), pp. 50–83. ISSN: 1092-2172. DOI: 10.1128/mmbr.00031-10.

[205] David L. Moyes et al. "A biphasic innate immune MAPK response discriminates between the yeast and hyphal forms of candida albicans in epithelial cells". In: *Cell Host and Microbe* 8.3 (2010), pp. 225–235. ISSN: 19313128. DOI: 10.1016/j.chom.2010.08.002. URL: http://dx.doi.org/10.1016/j.chom.2010.08.002.

[206] Min Liu et al. “Protein modifications in Hedgehog signaling: Cross talk and feedback regulation confer divergent Hedgehog signaling activity”. In: *BioEssays* 43.12 (2021). ISSN: 15211878. DOI: 10.1002/bies.202100153.

[207] Ryan Suderman et al. “Fundamental trade-offs between information flow in single cells and cellular populations”. In: *Proceedings of the National Academy of Sciences of the United States of America* 114.22 (2017), pp. 5755–5760. ISSN: 10916490. DOI: 10.1073/pnas.1615660114.

[208] Thomas Pfeiffer and Sebastian Bonhoeffer. “Evolution of cross-feeding in microbial populations.” In: *The American naturalist* 163.6 (2004). ISSN: 15375323. DOI: 10.1086/383593.

[209] Salva Duran-Nebreda and Ricard V. Solé. “Toward Synthetic Spatial Patterns in Engineered Cell Populations with Chemotaxis”. In: *ACS Synthetic Biology* 5.7 (2016), pp. 654–661. ISSN: 21615063. DOI: 10.1021/acssynbio.5b00254.

[210] Tetsuhiro S Hatakeyama and Kunihiko Kaneko. “Kinetic Memory Based on the Enzyme-Limited Competition”. In: *PLoS Computational Biology* 10.8 (Aug. 2014). Ed. by Gennady M. Verkhivker, e1003784. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1003784. URL: https://dx.plos.org/10.1371/journal.pcbi.1003784.

[211] Tetsuhiro S Hatakeyama and Kunihiko Kaneko. “Transition in relaxation paths in allosteric molecules: Enzymatic kinetically constrained model”. In: *Physical Review Research* 2.1 (Jan. 2020), p. 012005. ISSN: 2643-1564. DOI: 10.1103/PhysRevResearch.2.012005. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.012005.

# Appendix C

# Extended Appendix

## PDF versions of *Maple* Code (analytical proofs & models)

# Symmetry breaking meets multisite modification - Source code file

**Vaidhiswaran Ramesh[1], J Krishnan[1,2]**

**Affiliations:**
1 - Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College London, London, SW7 2AZ, UK
2 - Institute for Systems and Synthetic Biology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

**Analytical proofs exploring the presence and features of symmetry and symmetry breaking in various multisite phosphorylation networks**

This text details the analytical arguments used to ascertain features associated with symmetry breaking and the impossibility of the behavior in various MSP networks considered in the main text.

**Models**
The networks described in the main text are modelled as a system of ODEs (Refer to Fig 1 in the main text for reaction scfhematics and Appendix 2 Figure 10 for kinetic nomenclature of elementary reactions). We use mass action kinetic description for reaction rates and the overall model is constructed as a combination of such descriptions for the elementary reactions.

**Cross Validation and Parameters**
The models provided here in Maple have been analyzed and cross validated with results from Matlab simulations and bifurcation computations using the Matlab package MatCont. Matlab and bifurcation computations have in turn been cross validated with simulations from the software COPASI, which self generates the ODE expressions based on a reaction schematic. The parameters used for generating the figures are presented here (in addition to the being present in appendix 2). These parameters have been used to perform an additional cross validation of the analytical work by mapping the presence and features of symmetry breaking, as predicted by the analytical work and obtained in the computational analysis.

**Organization of Results**
The general organization of this Maple workbook is as follows. There are a total of 10 maple worksheets organized across 5 folders.

Read_Me (This file) : Contains the parameters used for generating the figures, and further cross validation of the invariants and necessary conditions predicted by the analytical work.

Folder 2: Ordered distributive double site phosphorylation (DSP) models and triple site phosphorylation (TSP) (2.1 - 2.3)
Folder 3: Random DSP (3.1 - 3.3)
Folder 4: Mixed-Random DSP (4.1 - 4.3)
Folder 5: Absolute Concentration Robustness (5.1)

Each of the folders contain the discussions and analytical work pertaining to the various models in that category. For example, in Random folder: The three different models, corresponding to common kinase-common phosphatase, separate kinase-common phosphatase and separate kinase-separate phosphatase enzyme configurations, are presented in separate Maple documents labelled System_1, System_2 and System_3. Within each of these models, the three different classes of symmetries, where applicable are presented.

**Necessary and Sufficient Conditions.**
In this supplementary material, we provide analytical arguments regarding the presence and absence of various cases of symmetry breaking in multiple DSP networks. In networks permitting symmetry breaking, we obtain a necessary and sufficient condition for the behavior to manifest at some total substrate concentrations, in terms of the kinetic parameters (and in some cases total enzyme concentrations).

These constraints act as sufficient conditions to ensure symmetry breaking behavior is seen for some finite value of total substrate concentration ($A_{Total}$). The choice of bifurcation along $A_{Total}$ is due to its flexibility in accommodating different classes of symmetry while changing. However the symmetry breaking can be encountered upon bifurcation along any parameter of interest (kinetic or enzyme concentrations) provided there is sufficient total substrate concentration.

Please note: This entire document is also presented in a PDF format and is available as Supplementary file 1.

## Parameters

The parameters used in generating the figures are presented below. The kinetic nomenclature pertaining to the specific system and class of symmetry is used below (refer main text, Fig 1D). As discussed earlier the bifurcation is performed along $A_{Total}$ in these figures. A cross verification of the analytical results showing that the necessary conditions for symmetry breaking are met is presented. The features of the symmetry broken state as predicted by the analytical work is also presented.

### Figure 2

#### A. Case 1 - Double site phosphorylation

*restart* :

$k_1 := 0.1 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 0.5 : k_{b2} := 1 : k_{ub2} := 1 :$

$P_{Total} := 0.1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} :$$

##### Necessary and sufficient condition

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $k_2 > k_1$

$$0.1 < 0.5 \qquad \textbf{(1.1.1.1)}$$

##### Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)

The invariant in the asymmetric steady states in ordered distributive DSP post case 1 symmetry breaking is the concentration of the partially modified substrate [Ap].
This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$Ap = -\frac{k_1}{c_2 \left(k_1 - k_2\right)}$$

$$Ap = 0.3750000000 \qquad \textbf{(1.1.2.1)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
Below, the position of this pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \frac{\left(-c_1 c_2 P_{Total} - 2 c_2\right) k_2^{\,2} - k_1 c_1 k_2 + P_{Total} k_1^{\,2} c_1 c_2}{\left(k_1 - k_2\right) c_1 c_2 k_2}$$

$$A_{Total} = 3.245000000 \qquad \textbf{(1.1.2.2)}$$

**B. Case 1 - System 1 Random DSP**

*restart* :

$k_1 := 0.25 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 0.4 : k_{b2} := 1 : k_{ub2} := 1 :$

$a_1 := 0.1 : a_{b1} := 1 : a_{ub1} := 1 :$
$a_2 := 1 : a_{b2} := 1 : a_{ub2} := 1 :$

$P_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} :$$
$$d_1 := \frac{a_{b1}}{a_1 + a_{ub1}} : d_2 := \frac{a_{b2}}{a_2 + a_{ub2}} :$$

**Necessary Constraint**

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $c_1 a_2\left(k_2 - k_1\right) + d_1 k_2\left(a_2 - a_1\right) > 0$

$$0 < 1.163636364 \qquad \textbf{(1.2.1.1)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$**

The invariant in the asymmetric steady states in System 1 Random ordered DSP post case 1 symmetry breaking is the concentration of the partially modified substrates [A01] and [A10]. This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A01 := -\frac{c_1 k_1 a_2}{c_2 \left(\left(\left(k_1 - k_2\right) c_1 - d_1 k_2\right) a_2 + d_1 k_2 a_1\right)}$$

$$A01 := 0.6260162603 \qquad \textbf{(1.2.2.1)}$$

$$A10 := -\frac{d_1 k_2 a_1}{d_2 \left(\left(d_1 \left(a_1 - a_2\right) - c_1 a_2\right) k_2 + c_1 k_1 a_2\right)}$$

$$A10 := 0.1626016260 \qquad \textbf{(1.2.2.2)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
Below, the position of this pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \left(d_2 \left(-c_2 \left(c_1 + d_1\right) \left(c_1 P_{Total} + d_1 P_{Total} + 2\right) k_2^{\,2} - c_1 k_1 \left(c_1 + d_1\right) k_2 + c_1^{\,2} c_2 k_1^{\,2} P_{Total}\right) a_2^{\,2} + 2 d_1 \left(\left(-\frac{c_1}{2} - \frac{d_1}{2}\right) k_2 + c_1 d_2 k_1 P_{Total}\right) c_2 k_2 a_1 a_2 + a_1^{\,2} c_2 d_1^{\,2} d_2 k_2^{\,2} P_{Total}\right) \Big/ \left(\left(\left(\left(-c_1 - d_1\right) k_2 + c_1 k_1\right) a_2 + a_1 d_1 k_2\right) a_2 \left(c_1 + d_1\right) c_2 d_2 k_2\right)$$

$$A_{Total} = 3.922980455 \qquad \textbf{(1.2.2.3)}$$

## C. Case 1 - System 3 Random DSP

*restart* :

$k_1 := 0.5 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1.5 : k_{b2} := 1 : k_{ub2} := 1 :$

$a_1 := 0.1 : a_{b1} := 1 : a_{ub1} := 1 :$
$a_2 := 1 : a_{b2} := 1 : a_{ub2} := 1 :$

$PI_{Total} := 1 :$
$P2_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} :$$
$$d_1 := \frac{a_{b1}}{a_1 + a_{ub1}} : d_2 := \frac{a_{b2}}{a_2 + a_{ub2}} :$$

### Necessary Constraint

Here we show that the necessary and sufficient conditions for symmetry breaking (refer to

analytical work) is satisfied for the parameters used.

1. $k_2 \cdot P1_{Total} > k_1 \cdot P2_{Total}$

$$0.5 < 1.5 \qquad \textbf{(1.3.1.1)}$$

2. $a_2 \cdot P2_{Total} > a_1 \cdot P1_{Total}$

$$0.1 < 2. \qquad \textbf{(1.3.1.2)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in System 3 Random ordered DSP post case 1 symmetry breaking is the concentration of the partially modified substrates [A01] and [A10]. This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A01 := -\frac{k_1\, P2_{Total}}{c_2\left(k_1\, P2_{Total} - k_2\, P1_{Total}\right)}$$

$$A01 := 1.250000000 \qquad \textbf{(1.3.2.1)}$$

$$A10 := -\frac{a_1\, P1_{Total}}{d_2\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)}$$

$$A10 := 0.2222222222 \qquad \textbf{(1.3.2.2)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.

First we evaluate the value of [K2] and [P2] using the various correlations between the substrates and free enzyme concentrations at steady state. Below we use the *solve* command in Maple to evaluate the value for [K2] from the correlation obtained between [A01] and [A10] earlier. We use additional arguments (use assumptions) to only obtain solutions for [K2] that permit positive concentrations of free enzyme concentrations.

$$K2 := solve\left(A10 = \frac{A01\, K2^2\, a_1\, c_2\, d_1\, k_2}{P2^2\, c_1\, k_1\, a_2\, d_2}, K2, useassumptions\right) \text{assuming } P2 > 0, K2 > 0$$

$$K2 := 0.7370277311\, P2 \qquad \textbf{(1.3.2.3)}$$

This gives us the value of $\alpha$ (the ratio of [K2]/[P2]). Using this in the expression for [P2] obtained earlier from solving P2Con at the symmetric state, we get the concentration of [P2] to be,

$$P2 := eval\left(-\frac{\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)\left(k_1\, P2_{Total} - k_2\, P1_{Total}\right)}{k_2\left(a_1\,\alpha - a_2\right) P1_{Total} + a_2\, P2_{Total}\left(-\alpha\, k_2 + k_1\right)}, \alpha = 1.356801051\right)$$

$$P2 := 0.3178323642 \qquad \textbf{(1.3.2.4)}$$

$$A11 := \frac{K2\ A01\ c_2\ k_2}{P2\ c_1\ k_1}$$

$$A11 := 1.658312396 \qquad \textbf{(1.3.2.5)}$$

$$A00 := \frac{K2\ A01\ c_2\ k_2}{P2\ c_1\ k_1}$$

$$A00 := 1.658312396 \qquad \textbf{(1.3.2.6)}$$

Using these values, the position of the pitch fork bifurcation is evaluated as shown below.

$$A_{Total} = \left( -\left( \left( -2\ a_2 \left( \left( P2\ \left( k_2 + k_1 \right) c_1 + k_2\ \left( K2\ d_1 + 1 \right) \right) K2\ c_2 + \frac{P2\ c_1\ k_1}{2} \right) d_2\ {P2_{Total}}^2 + 2\ {PI_{Total}} \left( \left( \left( P2\ \left( k_1\ P2 + K2\ \left( k_2 + k_1 \right) \right) \right) c_1 + K2\ k_2\ \left( K2\ d_1 + 1 \right) \right) d_2 + \frac{P2\ c_1\ k_1}{2} \right) c_2 + \frac{P2\ c_1\ d_2\ k_1}{2} \right) a_1\ {P2_{Total}} - 2\ {PI_{Total}}^2\ P2\ c_1 \left( P2\ d_2 + \frac{1}{2} \right) a_1\ k_2\ c_2 \right) \right) \right) \Big/ \left( \left( k_1\ {P2_{Total}} - k_2\ {PI_{Total}} \right) c_1\ P2\ d_2\ \left( a_1\ {PI_{Total}} - a_2\ {P2_{Total}} \right) c_2 \right)$$

$$A_{Total} = 6.502775632 \qquad \textbf{(1.3.2.7)}$$

## D. Case 2 - System 3 Random DSP

*restart* :

$k_1 := 2.35 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 0.46 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 1.86 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 1.1 : k_{b4} := 1 : k_{ub4} := 1 :$

$P_{Total} := 1 :$
$K_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

### Necessary Conditions

Here we show that the necessary and sufficient conditions for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $k_1 \cdot K_{Total} > k_4 \cdot P_{Total}$

$$1.1 < 2.35 \quad \textbf{(1.4.1.1)}$$

2. $k_3 \cdot P_{Total} > k_2 \cdot K_{Total}$

$$0.46 < 3.72 \quad \textbf{(1.4.1.2)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in System 3 Random ordered DSP post case 2 symmetry breaking is the concentration of the partially modified substrate [A00] and [A11]. This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A00 := \frac{k_4 P_{Total}}{c_1 \left(k_1 K_{Total} - k_4 P_{Total}\right)}$$

$$A00 := 2.948000000 \quad \textbf{(1.4.2.1)}$$

$$A11 := -\frac{k_2 K_{Total}}{c_3 \left(k_2 K_{Total} - k_3 P_{Total}\right)}$$

$$A11 := 0.9397142857 \quad \textbf{(1.4.2.2)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.

First we evaluate the value of [K2] and [P2] using the various correlations between the substrates and free enzyme concentrations at steady state. Below we use the *solve* command in Maple to evaluate the value for [K2] from the correlation obtained between [A01] and [A10] earlier. We use additional arguments (use assumptions) to only obtain solutions for [K2] that permit positive concentrations of free enzyme concentrations.

$$K2 := solve\left(A11 = \frac{A00\, K2^2\, c_1\, c_2\, k_1\, k_2}{P2^2\, c_4\, k_4\, c_3\, k_3}, K2, useassumptions\right) \text{assuming } P2 > 0, K2 > 0$$

$$K2 := 0.7009391379\, P2 \quad \textbf{(1.4.2.3)}$$

This gives us the value of $\epsilon$ (the ratio of [K2]/[P2]). Using this in the expression for [P2] obtained earlier from solving P2Con at the symmetric state, we get the concentration of [P2] to be,

$$P2 := eval\left(\frac{\left(k_2 K_{Total} - k_3 P_{Total}\right)\left(k_1 K_{Total} - P_{Total} k_4\right)}{k_1 \left(\epsilon k_2 - k_3\right) K_{Total} - k_3 P_{Total} \left(\epsilon k_1 - k_4\right)}, \epsilon = 0.7009391379\right)$$

$$P2 := 0.3778808204 \quad \textbf{(1.4.2.4)}$$

$$A01 := \frac{K2\ A00\ c_1\ k_1}{P2\ c_4\ k_4}$$

$$A01 := 2.767307716 \quad \textbf{(1.4.2.5)}$$

$$A10 := \frac{K2\ A00\ c_1\ k_1}{P2\ c_4\ k_4}$$

$$A10 := 2.767307716 \quad \textbf{(1.4.2.6)}$$

Using these values, the position of the pitch fork bifurcation is evaluated as shown below.

$$A_{Total} = \frac{1}{P2^2\ c_4\ k_4\ c_3\ k_3}\left(2\left(P2\left(\left(\left(K2\left(k_4+k_1\right)c_1+\frac{k_4}{2}\right)c_4\ k_3\ P2+\left(\left(K2\ c_2+1\right)k_3\right.\right.\right.\right.$$
$$\left.\left.\left.\left.+K2\ c_2\ k_2\right)c_1\ k_1\ K2\right)c_3+\frac{K2^2\ c_1\ c_2\ k_1\ k_2}{2}\right)A00\right)$$

$$A_{Total} = 12.13682517 \quad \textbf{(1.4.2.7)}$$

### E. Case 2 - Mixed-Random 2 DSP

*restart* :

$k_1 := 2 : k_{b1} := 1 : k_{ub1} := 1$ :
$k_2 := 0.1 : k_{b2} := 1 : k_{ub2} := 1$ :
$k_3 := 0.75 : k_{b3} := 1 : k_{ub3} := 1$ :
$k_4 := 1$ :

$P_{Total} := 0.2 : KI_{Total} := 0.1$ :

$$c_1 := \frac{k_{b1}}{k_1+k_{ub1}} : c_2 := \frac{k_{b2}}{k_2+k_{ub2}} : c_3 := \frac{k_{b3}}{k_3+k_{ub3}} :$$

#### Necessary Conditions

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $k_2 < k_1$

$$0.1 < 2 \quad \textbf{(1.5.1.1)}$$

2. $k_2\ KI_{Total}\left(k_3+k_4\right) < P_{Total}\ k_3\ k_4$

$$0.0350 < 0.150 \quad \textbf{(1.5.1.2)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in Mixed Random 2 ordered DSP post case 2 symmetry breaking is the concentration of the partially modified substrate [A00] and [A11]. This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A00 = \frac{k_2}{\left(k_1 - k_2\right) c_1}$$

$$A00 = 0.1578947368 \qquad \textbf{(1.5.2.1)}$$

$$A11 = -\frac{k_2\, k_4\, KI_{Total}}{2\, c_3 \left(k_2\, KI_{Total} \left(k_3 + k_4\right) - P_{Total}\, k_3\, k_4\right)}$$

$$A11 = 0.06603773585 \qquad \textbf{(1.5.2.2)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$\begin{aligned} A_{Total} := {} & \Big(-2\, c_3 \left(k_3 + k_4\right) \left(\left(k_1 + k_4\right) k_3 + k_1\, k_4\right) c_1\, c_2\, KI_{Total}^{\,2}\, k_2^{\,3} \\ & + 2\, c_2 \Big(c_3 \left(c_1\, P_{Total}\, k_4^{\,2} + k_1 \left(c_1\, P_{Total} + 1\right) k_4 + c_1\, k_1^{\,2}\, KI_{Total}\right) k_3^{\,2} \\ & + 2 \left(\left(\frac{1}{2}\, c_1\, c_3\, P_{Total} + \frac{1}{4}\, c_1 + \frac{1}{2}\, c_3\right) k_4 + c_1\, c_3\, k_1\, KI_{Total}\right) k_4\, k_1\, k_3 \\ & + c_1\, c_3\, k_1^{\,2}\, k_4^{\,2}\, KI_{Total}\Big)\, KI_{Total}\, k_2^{\,2} - 2\, k_3 \Big(c_3 \left(c_2\, P_{Total}\, k_4 + c_1\, k_1\, KI_{Total} \left(\right.\right. \\ & \left.\left. -KI_{Total}\, c_2 + c_2\, P_{Total} - 2\right)\right) k_3 + c_1\, k_1\, k_4\, KI_{Total} \Big(c_2\, c_3\, P_{Total} - c_2\, c_3\, KI_{Total} \\ & + \frac{1}{2}\, c_2 - 2\, c_3\Big)\Big)\, k_4\, k_1\, k_2 - 2\, c_1\, c_3\, k_1^{\,2}\, k_3^{\,2}\, k_4^{\,2}\, P_{Total} \left(KI_{Total}\, c_2 + 2\right)\Big) \Big/ \left(2 \left(k_1 \right.\right. \\ & \left.\left. - k_2\right) c_3\, k_3 \left(k_2 \left(k_3 + k_4\right) KI_{Total} - P_{Total}\, k_3\, k_4\right) k_4\, c_1\, c_2\, k_1\right) \end{aligned}$$

$$A_{Total} := 2.668055280 \qquad \textbf{(1.5.2.3)}$$

## Figure 3

### A. Case 3 - System 1 Random DSP (Hopf and Pitchfork)

*restart* :

$k_1 := 100 : k_{b1} := 100 : k_{ub1} := 1 :$
$k_2 := 2 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 0.01 : k_{b3} := 100 : k_{ub3} := 1 :$
$k_4 := 20 : k_{b4} := 0.1 : k_{ub4} := 1 :$

$P_{Total} := 1.25 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

**Necessary conditions**

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $c_3 k_4 (k_2 - k_3) - c_1 k_2 (k_1 - k_4) > 0$

$$0 < 3782.178218 \qquad \textbf{(2.1.1.1)}$$

**Symmetry Breaking**

The invariant in the asymmetric steady states in System 1 Random ordered DSP post case 3 symmetry breaking is the sum of the concentrations of the partially modified substrates [A01] and [A10].
This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A01 + A10 = \frac{-c_1 c_2 k_1 k_2 - c_3 c_4 k_3 k_4}{(((-c_1 - c_3) k_4 + c_1 k_1) k_2 + c_3 k_3 k_4) c_4 c_2}$$

$$A01 + A10 = 11.01047120 \qquad \textbf{(2.1.2.1)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
Below, the position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \left(2 (c_1 c_2 k_1 k_2 + c_3 c_4 k_3 k_4) \left(-\frac{c_1 k_2^2 k_1 ((k_1 + k_2) c_1 + c_3 (k_2 + k_3)) c_2^3}{2}\right.\right.$$
$$+ \left(\left(k_1 k_2 P_{Total} (k_1 - k_4) (k_1 + k_2) c_1^3 + \left(((-2 k_1 - k_3) k_4 + k_1 (k_1 + k_3)) k_2^2\right.\right.\right.$$

$$+\left(-k_1\left(k_1+k_3\right)k_4+2\,{k_1}^2\,k_3\right)k_2+{k_1}^2\,k_3\,k_4\Big)\,P_{Total}\,c_3$$

$$+\frac{k_2\left(\left(k_1-4\,k_4\right)k_2-k_4\,k_1\right)k_1}{2}\Bigg)\,{c_1}^2+c_3\left(\left(\left(\left(-k_1-2\,k_3\right)k_4+k_1\,k_3\right){k_2}^2\right.\right.$$

$$\left.+k_1\,k_3\left(k_3-k_4\right)k_2+2\,k_1\,{k_3}^2\,k_4\right)P_{Total}\,c_3+\frac{{k_2}^2\left(\left(-4\,k_1-3\,k_3\right)k_4+k_1\,k_3\right)}{2}\Bigg)\,c_1$$

$$-k_4\left(\left({k_2}^2\,P_{Total}-{k_3}^2\,P_{Total}\right)c_3+\frac{3\,k_2\left(k_2-\frac{k_3}{3}\right)}{2}\right){c_3}^2\,k_3\Bigg)\,c_4$$

$$-\frac{c_1\,k_2\,k_1\left(\left(k_1+k_2\right)c_1+c_3\left(k_2+k_3\right)\right)\left(c_1\,k_1+c_3\,k_3\right)}{2}\Bigg)\,{c_2}^2+\Bigg(\Bigg(\Bigg(k_1\,k_2\,P_{Total}\left(k_1\right.$$

$$\left.-k_4\right)\left(k_1+k_4\right){c_1}^3+\Bigg(2\left(\left(\left(-k_1-\frac{k_3}{2}\right){k_4}^2-\frac{k_3\,k_4\,k_1}{2}+{k_1}^2\,k_3\right)k_2\right.$$

$$\left.+\frac{k_1\,k_3\,k_4\left(k_1+k_4\right)}{2}\right)P_{Total}\,c_3+\frac{k_1\,k_2\,k_4\left(k_1-3\,k_4\right)}{2}\Bigg)\,{c_1}^2+c_3\Bigg(\left(\left(-k_1\right.\right.$$

$$\left.\left.-2\,k_3\right){k_4}^2-k_3\left(k_1+k_3\right)k_4+k_1\,{k_3}^2\right)k_2+k_3\left(k_1+k_3\right){k_4}^2+2\,k_1\,{k_3}^2\,k_4\Big)\,P_{Total}\,c_3$$

$$-\frac{3\,{k_4}^2\left(\left(k_1+\frac{4\,k_3}{3}\right)k_2-\frac{k_1\,k_3}{3}\right)}{2}\Bigg)\,c_1-k_4\Bigg(P_{Total}\left(k_3+k_4\right)\left(k_2-k_3\right)c_3$$

$$+\left(\frac{k_3}{2}+2\,k_4\right)k_2-\frac{k_3\,k_4}{2}\Bigg)\,{c_3}^2\,k_3\Bigg)\,c_4$$

$$+\frac{\left(k_2\left(k_1-3\,k_4\right)c_1-3\,k_4\,c_3\left(k_2-\frac{k_3}{3}\right)\right)\left(c_1\,k_1+c_3\,k_3\right)^2}{2}\Bigg)\,c_4\,c_2$$

$$-\frac{k_4\left(\left(k_1+k_4\right)c_1+c_3\left(k_3+k_4\right)\right)\left(c_1\,k_1+c_3\,k_3+c_4\,k_4\right)c_3\,k_3\,{c_4}^2}{2}\Bigg)\Bigg)\Bigg/$$

$$\Bigg(\Bigg(c_1\,k_2\,k_1\left(\left(k_1+k_2\right)c_1+c_3\left(k_2+k_3\right)\right){c_2}^2-\Bigg(k_2\left(k_1-3\,k_4\right)c_1-3\,k_4\,c_3\Bigg(k_2$$

$$-\frac{k_3}{3}\Bigg)\Bigg)\left(c_1\,k_1+c_3\,k_3\right)c_4\,c_2+k_4\left(\left(k_1+k_4\right)c_1+c_3\left(k_3+k_4\right)\right)c_3\,k_3\,{c_4}^2\Bigg)\left(\left(k_1\right.\right.$$

$$-k_4)\,k_2\,c_1-k_4\,c_3\,(k_2-k_3))\,(c_1\,k_1+c_3\,k_3)\,c_4\,c_2\Big)$$

$A_{Total} = 13.24004002$

**A. Case 3 - System 1 Random DSP (Tristability)**

*restart* :

$k_1 := 100 : k_{b1} := 100 : k_{ub1} := 1 :$
$k_2 := 2 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 0.01 : k_{b3} := 100 : k_{ub3} := 1 :$
$k_4 := 20 : k_{b4} := 0.1 : k_{ub4} := 1 :$

$P_{Total} := 10 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

**Necessary conditions**

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $c_3\,k_4\,(k_2-k_3)-c_1\,k_2\,(k_1-k_4)>0$

$0 < 3782.178218$ **(2.2.1.1)**

**Symmetry Breaking**

The invariant in the asymmetric steady states in System 1 Random ordered DSP case 3 post symmetry breaking is the sum of the concentrations of the partially modified substrates [A01] and [A10].
This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A01 + A10 = \frac{-c_1\,c_2\,k_1\,k_2 - c_3\,c_4\,k_3\,k_4}{\left(\left(\left(-c_1-c_3\right)k_4 + c_1\,k_1\right)k_2 + c_3\,k_3\,k_4\right)c_4\,c_2}$$

$A01 + A10 = 11.01047120$ **(2.2.2.1)**

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates

the presence of a pitch fork bifurcation.
The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$
\begin{aligned}
A_{Total} = {} & \left( 2 \left( c_1 c_2 k_1 k_2 + c_3 c_4 k_3 k_4 \right) \left( - \frac{c_1 k_2^{\,2} k_1 \left( \left( k_1 + k_2 \right) c_1 + c_3 \left( k_2 + k_3 \right) \right) c_2^{\,3}}{2} \right. \right. \\
& + \left( \left( k_1 k_2 P_{Total} \left( k_1 - k_4 \right) \left( k_1 + k_2 \right) c_1^{\,3} + \left( \left( \left( -2 k_1 - k_3 \right) k_4 + k_1 \left( k_1 + k_3 \right) \right) k_2^{\,2} \right. \right. \right. \\
& \left. + \left( -k_1 \left( k_1 + k_3 \right) k_4 + 2 k_1^{\,2} k_3 \right) k_2 + k_1^{\,2} k_3 k_4 \right) P_{Total} c_3 \\
& \left. + \frac{k_2 \left( \left( k_1 - 4 k_4 \right) k_2 - k_4 k_1 \right) k_1}{2} \right) c_1^{\,2} + c_3 \left( \left( \left( \left( -k_1 - 2 k_3 \right) k_4 + k_1 k_3 \right) k_2^{\,2} \right. \right. \\
& \left. + k_1 k_3 \left( k_3 - k_4 \right) k_2 + 2 k_1 k_3^{\,2} k_4 \right) P_{Total} c_3 + \frac{k_2^{\,2} \left( \left( -4 k_1 - 3 k_3 \right) k_4 + k_1 k_3 \right)}{2} \Bigg) c_1 \\
& - k_4 \left( \left( k_2^{\,2} P_{Total} - k_3^{\,2} P_{Total} \right) c_3 + \frac{3 k_2 \left( k_2 - \frac{k_3}{3} \right)}{2} \right) c_3^{\,2} k_3 \Bigg) c_4 \\
& \left. - \frac{c_1 k_2 k_1 \left( \left( k_1 + k_2 \right) c_1 + c_3 \left( k_2 + k_3 \right) \right) \left( c_1 k_1 + c_3 k_3 \right)}{2} \right) c_2^{\,2} + \left( \left( k_1 k_2 P_{Total} \left( k_1 \right. \right. \right. \\
& \left. - k_4 \right) \left( k_1 + k_4 \right) c_1^{\,3} + \left( 2 \left( \left( \left( -k_1 - \frac{k_3}{2} \right) k_4^{\,2} - \frac{k_3 k_4 k_1}{2} + k_1^{\,2} k_3 \right) k_2 \right. \right. \\
& \left. + \frac{k_1 k_3 k_4 \left( k_1 + k_4 \right)}{2} \right) P_{Total} c_3 + \frac{k_1 k_2 k_4 \left( k_1 - 3 k_4 \right)}{2} \Bigg) c_1^{\,2} + c_3 \left( \left( \left( -k_1 \right. \right. \right. \\
& \left. - 2 k_3 \right) k_4^{\,2} - k_3 \left( k_1 + k_3 \right) k_4 + k_1 k_3^{\,2} \right) k_2 + k_3 \left( k_1 + k_3 \right) k_4^{\,2} + 2 k_1 k_3^{\,2} k_4 \Big) P_{Total} c_3 \\
& \left. - \frac{3 k_4^{\,2} \left( \left( k_1 + \frac{4 k_3}{3} \right) k_2 - \frac{k_1 k_3}{3} \right)}{2} \right) c_1 - k_4 \left( P_{Total} \left( k_3 + k_4 \right) \left( k_2 - k_3 \right) c_3 \right. \\
& \left. + \left( \frac{k_3}{2} + 2 k_4 \right) k_2 - \frac{k_3 k_4}{2} \right) c_3^{\,2} k_3 \Bigg) c_4
\end{aligned}
$$

$$+\frac{\left(k_2\left(k_1-3\,k_4\right)c_1-3\,k_4\,c_3\left(k_2-\frac{k_3}{3}\right)\right)\left(c_1\,k_1+c_3\,k_3\right)^2}{2}\Bigg)c_4\,c_2$$

$$-\frac{k_4\left(\left(k_1+k_4\right)c_1+c_3\left(k_3+k_4\right)\right)\left(c_1\,k_1+c_3\,k_3+c_4\,k_4\right)c_3\,k_3\,{c_4}^2}{2}\Bigg)\Bigg)\Bigg/$$

$$\Bigg(\Bigg(c_1\,k_2\,k_1\left(\left(k_1+k_2\right)c_1+c_3\left(k_2+k_3\right)\right){c_2}^2-\Bigg(k_2\left(k_1-3\,k_4\right)c_1-3\,k_4\,c_3\left(k_2-\frac{k_3}{3}\right)\Bigg)\left(c_1\,k_1+c_3\,k_3\right)c_4\,c_2+k_4\left(\left(k_1+k_4\right)c_1+c_3\left(k_3+k_4\right)\right)c_3\,k_3\,{c_4}^2\Bigg)\left(\left(k_1-k_4\right)k_2\,c_1-k_4\,c_3\left(k_2-k_3\right)\right)\left(c_1\,k_1+c_3\,k_3\right)c_4\,c_2\Bigg)$$

$$A_{Total}=28.25979662 \qquad \textbf{(2.2.2.2)}$$

**B. Case 3 - System 3 Random DSP (Hopf)**

*restart* :

$k_1 := 150 : k_{b1} := 100 : k_{ub1} := 1 :$
$k_2 := 50 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 1 : k_{b3} := 0.01 : k_{ub3} := 1 :$
$k_4 := 10 : k_{b4} := 500 : k_{ub4} := 1 :$

$K1_{Total} := 1 : P1_{Total} := 1 :$
$K2_{Total} := 1 : P2_{Total} := 1 :$

**B. Case 3 - System 3 Random DSP (PitchFork - Approximate robustness in A00 + A11)**

*restart* :

$k_1 := 10 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 5 : k_{b4} := 1 : k_{ub4} := 1 :$

$K1_{Total} := 1 : P1_{Total} := 1 :$

## Appendix 2 Figure 1

### B. Case 2 - Mixed-Random 3 DSP

*restart* :

$k_1 := 0.9 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 0.8 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 1 :$

$KI_{Total} := 0.1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} :$$

#### Necessary Conditions

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $k_2 < k_1$

$$0.8 < 0.9 \qquad \textbf{(3.1.1.1)}$$

#### Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)

The invariant in the asymmetric steady states in Mixed Random 3 ordered DSP case 2 post symmetry breaking is the concentrations of the partially modified substrates [A00] and [A11]. This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A11 = \frac{KI_{Total}\, k_2}{2\, k_3}$$

$$A11 = 0.02000000000 \qquad \textbf{(3.1.2.1)}$$

$$A00 = \frac{k_2}{(k_1 - k_2)\, c_1}$$

$$A00 = 15.20000000 \qquad \textbf{(3.1.2.2)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.

The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \frac{1}{2\left(k_1 - k_2\right) c_1 c_2 k_3 k_1}\left(\left(KI_{Total}\left(k_2 + 2k_3\right) c_2 + 4k_3\right) c_1 {k_1}^2 - \left(KI_{Total} k_2 c_1 - 2k_3\right) c_2 k_2 k_1 - 2\,KI_{Total} {k_2}^2 c_1 c_2 k_3\right)$$

$$= A_{Total} = 47.80888889 \qquad \textbf{(3.1.2.3)}$$

## Appendix 2 Figure 2

### A. Case 3 - System 1 Random DSP (Hopf)

*restart* :

$k_1 := 100 : k_{b1} := 100 : k_{ub1} := 1 :$
$k_2 := 20 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 1 : k_{b3} := 100 : k_{ub3} := 1 :$
$k_4 := 2 : k_{b4} := 1 : k_{ub4} := 1 :$

$P_{Total} := 1 :$

### A. Case 3 - System 1 Random DSP (Oscillations - Dynamic response)

*restart* :

$A_{Total} := 12.95 :$

### A. Case 3 - System 1 Random DSP (Pitchfork)

*restart* :

$k_1 := 0.1 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 5 : k_{b4} := 1 : k_{ub4} := 1 :$

$P_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

**Necessary conditions**

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $c_3 k_4 (k_2 - k_3) - c_1 k_2 (k_1 - k_4) > 0$

$$0 < 2.787878788 \qquad \textbf{(4.3.1.1)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in System Random 1 ordered distributive DSP post case 3 symmetry breaking is the sum of the concentrations of the partially modified substrates [A00] and [A11].
This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A01 + A10 = \frac{-c_1 c_2 k_1 k_2 - c_3 c_4 k_3 k_4}{(((-c_1 - c_3) k_4 + c_1 k_1) k_2 + c_3 k_3 k_4) c_4 c_2}$$

$$A01 + A10 = 2.586956524 \qquad \textbf{(4.3.2.1)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \left(2 (c_1 c_2 k_1 k_2 + c_3 c_4 k_3 k_4) \left(-\frac{c_1 k_2^2 k_1 ((k_1 + k_2) c_1 + c_3 (k_2 + k_3)) c_2^3}{2} + \left(\left(k_1 k_2 P_{Total} (k_1 - k_4) (k_1 + k_2) c_1^3 + \left(((-2 k_1 - k_3) k_4 + k_1 (k_1 + k_3)) k_2^2 + (-k_1 (k_1 + k_3) k_4 + 2 k_1^2 k_3) k_2 + k_1^2 k_3 k_4\right) P_{Total} c_3 + \frac{k_2 ((k_1 - 4 k_4) k_2 - k_4 k_1) k_1}{2}\right) c_1^2 + c_3 \left(((-k_1 - 2 k_3) k_4 + k_1 k_3) k_2^2 + k_1 k_3 (k_3 - k_4) k_2 + 2 k_1 k_3^2 k_4\right) P_{Total} c_3 + \frac{k_2^2 ((-4 k_1 - 3 k_3) k_4 + k_1 k_3)}{2}\right) c_1\right)\right.$$

$$
\begin{aligned}
&- k_4 \left( \left( {k_2}^2 P_{Total} - {k_3}^2 P_{Total} \right) c_3 + \frac{3 k_2 \left( k_2 - \frac{k_3}{3} \right)}{2} \right) {c_3}^2 k_3 \Bigg) c_4 \\
&- \frac{c_1 k_2 k_1 \left( \left( k_1 + k_2 \right) c_1 + c_3 \left( k_2 + k_3 \right) \right) \left( c_1 k_1 + c_3 k_3 \right)}{2} \Bigg) {c_2}^2 + \Bigg( \Bigg( k_1 k_2 P_{Total} \left( k_1 \right. \\
&\left. - k_4 \right) \left( k_1 + k_4 \right) {c_1}^3 + \Bigg( 2 \Bigg( \Bigg( \left( -k_1 - \frac{k_3}{2} \right) {k_4}^2 - \frac{k_3 k_4 k_1}{2} + {k_1}^2 k_3 \Bigg) k_2 \\
&+ \frac{k_1 k_3 k_4 \left( k_1 + k_4 \right)}{2} \Bigg) P_{Total} c_3 + \frac{k_1 k_2 k_4 \left( k_1 - 3 k_4 \right)}{2} \Bigg) {c_1}^2 + c_3 \Bigg( \Big( \Big( -k_1 \\
&- 2 k_3 \Big) {k_4}^2 - k_3 \left( k_1 + k_3 \right) k_4 + k_1 {k_3}^2 \Big) k_2 + k_3 \left( k_1 + k_3 \right) {k_4}^2 + 2 k_1 {k_3}^2 k_4 \Big) P_{Total} c_3 \\
&- \frac{3 {k_4}^2 \left( \left( k_1 + \frac{4 k_3}{3} \right) k_2 - \frac{k_1 k_3}{3} \right)}{2} \Bigg) c_1 - k_4 \Bigg( P_{Total} \left( k_3 + k_4 \right) \left( k_2 - k_3 \right) c_3 \\
&+ \left( \frac{k_3}{2} + 2 k_4 \right) k_2 - \frac{k_3 k_4}{2} \Bigg) {c_3}^2 k_3 \Bigg) c_4 \\
&+ \frac{\left( k_2 \left( k_1 - 3 k_4 \right) c_1 - 3 k_4 c_3 \left( k_2 - \frac{k_3}{3} \right) \right) \left( c_1 k_1 + c_3 k_3 \right)^2}{2} \Bigg) c_4 c_2 \\
&- \frac{k_4 \left( \left( k_1 + k_4 \right) c_1 + c_3 \left( k_3 + k_4 \right) \right) \left( c_1 k_1 + c_3 k_3 + c_4 k_4 \right) c_3 k_3 {c_4}^2}{2} \Bigg) \Bigg) \Bigg/ \\
&\Bigg( \Bigg( \Bigg( c_1 k_2 k_1 \left( \left( k_1 + k_2 \right) c_1 + c_3 \left( k_2 + k_3 \right) \right) {c_2}^2 - \Bigg( k_2 \left( k_1 - 3 k_4 \right) c_1 - 3 k_4 c_3 \Bigg( k_2 \\
&- \frac{k_3}{3} \Bigg) \Bigg) \left( c_1 k_1 + c_3 k_3 \right) c_4 c_2 + k_4 \left( \left( k_1 + k_4 \right) c_1 + c_3 \left( k_3 + k_4 \right) \right) c_3 k_3 {c_4}^2 \Bigg) \left( \left( k_1 \right. \right. \\
&\left. \left. - k_4 \right) k_2 c_1 - k_4 c_3 \left( k_2 - k_3 \right) \right) \left( c_1 k_1 + c_3 k_3 \right) c_4 c_2 \Bigg)
\end{aligned}
$$

$$A_{Total} = 8.713626644 \qquad \textbf{(4.3.2.2)}$$

## Appendix 2 Figure 3

### Case 1 - System 1 Random DSP (Independent leg bifurcation)

*restart* :

$k_1 := 1 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1 : k_{b2} := 1 : k_{ub2} := 1 :$
$P_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} :$$

$a_1 := 0.1 : a_{b1} := 1 : a_{ub1} := 1 :$
$a_2 := 1 : a_{b2} := 1 : a_{ub2} := 1 :$

$$d_1 := \frac{a_{b1}}{a_1 + a_{ub1}} : d_2 := \frac{a_{b2}}{a_2 + a_{ub2}} :$$

**Necessary Constraint**

Here we show that the necessary and sufficient conditions for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $c_1 a_2 (k_2 - k_1) + a_1 k_2 (a_2 - a_1) > 0$

$0 < 0.6000000000$ **(5.1.1.1)**

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in System Random 1 ordered distributive DSP post case 1 symmetry breaking is the concentrations of the partially modified substrates [A01] and [A10].
This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A10 := -\frac{d_1 k_2 a_1}{d_2 \left( \left( d_1 \left( a_1 - a_2 \right) - c_1 a_2 \right) k_2 + c_1 k_1 a_2 \right)}$$

$A10 := 0.2222222222$ **(5.1.2.1)**

$$A01 := -\frac{c_1 k_1 a_2}{c_2 \left( \left( \left( k_1 - k_2 \right) c_1 - d_1 k_2 \right) a_2 + d_1 k_2 a_1 \right)}$$

$A01 := 1.222222222$ **(5.1.2.2)**

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.

The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \left( d_2 \left( -c_2 \left( c_1 + d_1 \right) \left( c_1 P_{Total} + d_1 P_{Total} + 2 \right) k_2^{\,2} - c_1 k_1 \left( c_1 + d_1 \right) k_2 + c_1^{\,2} c_2 k_1^{\,2} P_{Total} \right) a_2^{\,2} + 2 d_1 \left( \left( -\frac{c_1}{2} - \frac{d_1}{2} \right) k_2 + c_1 d_2 k_1 P_{Total} \right) c_2 k_2 a_1 a_2 + a_1^{\,2} c_2 d_1^{\,2} d_2 k_2^{\,2} P_{Total} \right) \Big/ \left( \left( \left( \left( -c_1 - d_1 \right) k_2 + c_1 k_1 \right) a_2 + a_1 d_1 k_2 \right) a_2 \left( c_1 + d_1 \right) c_2 d_2 k_2 \right)$$

$$A_{Total} = 5.308243724 \qquad \textbf{(5.1.2.3)}$$

## Appendix 2 Figure 4

### Case 3 - System 3 Random DSP (PitchFork - Approximate robustness in A01 + A10)

*restart* :

$k_1 := 1 : k_{b1} := 10 : k_{ub1} := 1 :$
$k_2 := 2 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 1 : k_{b3} := 10 : k_{ub3} := 1 :$
$k_4 := 20 : k_{b4} := 1 : k_{ub4} := 1 :$

$K1_{Total} := 5 : P1_{Total} := 5 :$
$K2_{Total} := 5 : P2_{Total} := 5 :$

## Appendix 2 Figure 5

### Case 2 - System 3 Random DSP (Approximate robustness in non exact symmetric state)

*restart* :

$k_1 := 1.25 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1.1 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2.5 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 0.4 : k_{b4} := 1 : k_{ub4} := 1 :$

$a_1 := 1.25 : a_{b1} := 1 : a_{ub1} := 1 :$
$a_{b2} := 1 : a_{ub2} := 1 :$

$a_3 := 2.5 : a_{b3} := 1 : a_{ub3} := 1 :$
$a_4 := 0.4 : a_{b4} := 1 : a_{ub4} := 1 :$

$KI_{Total} := 1 : PI_{Total} := 1 :$

**Exact symmetry used for comparison**

$k_1 := 1.25 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1.1 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2.5 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 0.4 : k_{b4} := 1 : k_{ub4} := 1 :$

$KI_{Total} := 1 : PI_{Total} := 1 :$

## Appendix 2 Figure 6

**Approximate concentration robustness in near symmetric systems (Ordered DSP with common kinase common phosphatase)**

*restart* :

$k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 0.5 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 0.1 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 0.5 : k_{b2} := 1 : k_{ub2} := 1 :$

$P_{Total} := 0.1 : K_{Total} := 0.1 :$

k1 is varied between 50% to 150% from it's symmetric value (0.1). All other parameters remain the same as those used in Fig2A.

## Appendix 2 Figure 7

**A. Case 3 - System 1 Random DSP**

*restart* :

$k_1 := 30 : k_{b1} := 100 : k_{ub1} := 1 :$
$k_2 := 2 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 0.3 : k_{b3} := 100 : k_{ub3} := 1 :$

$k_4 := 20 : k_{b4} := 0.1 : k_{ub4} := 1 :$

$KI_{Total} := 1 : PI_{Total} := 1 :$

**B. Case 3 - System 1 Random DSP**

Same kinetic parameters as above

$KI_{Total} := 20 : PI_{Total} := 20 :$

## Appendix 2 Figure 8

**Case 1 - Triple site phosphorylation**

$restart :$

$k_1 := 0.1 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 1.5 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 2 : k_{b3} := 1 : k_{ub3} := 1 :$

$P_{Total} := 0.1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} :$$

**Necessary Constraint**

Here we show that the necessary and sufficient condition for symmetry breaking (refer to analytical work) is satisfied for the parameters used.

1. $k_3 > k_1$

$$0.1 < 2. \qquad \textbf{(10.1.1.1)}$$

**Symmetry Breaking (Invariants and position of symmetry in the bifurcation diagram along $A_{Total}$)**

The invariant in the asymmetric steady states in distributive ordered TSP post case 1 symmetry breaking is the sum of the concentrations of the partially modified substrates [Ap] and [App].

This is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$Ap + App = -\frac{k_1}{c_3\left(k_1 - k_3\right)}$$

$$Ap + App = 0.1578947368 \qquad \textbf{(10.1.2.1)}$$

The intersection of the symmetric and asymmetric steady states on the bifurcation plot indicates the presence of a pitch fork bifurcation.
The position of the pitch fork bifurcation is evaluated (based on the expression from the analytical work) for the kinetic parameters used in generating the plot.

$$A_{Total} = \frac{\left(2\,c_1\,c_2\,c_3\,k_1^{\,2} - 2\,c_1\,c_2\,c_3\,k_1\,k_3 + 2\,c_1\,c_3^{\,2}\,k_1^{\,2} - 2\,c_1\,c_3^{\,2}\,k_3^{\,2}\right) P_{Total}}{\left(k_1 - k_3\right) c_1\,c_3\left(c_2\,k_1 - c_3\,k_1 + 3\,c_3\,k_3\right)} + \frac{-c_1\,c_2\,k_1^{\,2} + c_1\,c_3\,k_1^{\,2} - 3\,c_1\,c_3\,k_1\,k_3 - c_2\,c_3\,k_1\,k_3 + c_3^{\,2}\,k_1\,k_3 - 3\,c_3^{\,2}\,k_3^{\,2}}{\left(k_1 - k_3\right) c_1\,c_3\left(c_2\,k_1 - c_3\,k_1 + 3\,c_3\,k_3\right)} =$$

$$A_{Total} = 1.389543626 \qquad \textbf{(10.1.2.2)}$$

## Appendix 2 Figure 9

### Ordered DSP (Common kinase common phosphatase) - ACR with non-symmetric kinetics along changing $A_{Total}$

*restart* :

$k_1 := 1 : k_{b1} := 1 : k_{ub1} := 1 :$
$k_2 := 2 : k_{b2} := 1 : k_{ub2} := 1 :$
$k_3 := 0.5 : k_{b3} := 1 : k_{ub3} := 1 :$
$k_4 := 1.2 : k_{b4} := 11 : k_{ub4} := 1 :$

$K_{Total} := 1 : P_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

#### *Necessary and sufficient condition*

Here we show that the necssary and sufficient condition for ACR in Ap is satisfied by the

parameters used.

1. $$\frac{k_3 P_{Total}}{c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)} = -\frac{k_1 K_{Total}}{c_4 \left(K_{Total} k_1 - k_4 P_{Total}\right)}$$
$$1.000000000 = 1.000000000 \quad \textbf{(11.1.1.1)}$$

### *ACR concentration of Ap*

The ACR concentration of Ap (as seen in the analytical work in section 5.1) is thus given by,

$$Ap = \frac{k_3 P_{Total}}{c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)}$$
$$Ap = 1.000000000 \quad \textbf{(11.1.2.1)}$$

This is verified in the compuatational result in the figure.

# Double Site Ordered Distributive Phosphorylation : Common Kinase Common Phosphatase Case 1 Symmetry - Present and Breaks

In this Maple file we analytically show the presence of symmetry breaking in ordered distributive DSP network with common kinase and common phosphatase effecting phosphorylation and dephoshorylation respectively. We do this by first describing the model as a system of ODEs and imposing the kinetic constraints pertaining to case 1 symmetry. We further describe the enzyme and substrate conservations associated with the model. By solving for the steady states of the system of ODEs we obtain relations between substrate variables (concentrations) in terms of each other and the free enzyme concentrations. In ordered distributive DSP (with common kinase common phosphatase), $K_{Total} = P_{Total}$ is required for exact case 1 symmetry to be present. We use this information to get a further simplified expression describing all possible steady states of the system. With case 1 symmetry in the ordered distributive MSP, the free enzyme kinase and phosphatase share a strict symmetry in the symmetric steady state ([K] = [P]). Thus by isolating steady states not of this type from the earlier correlation, we ascertain the features of the asymmetric states emerging from symmetry breaking. This procedure is carried out in detail below using built in Maple commands.

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer to Appendix 2 figure 10). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.**

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot ApK + k_{ub4} \cdot ApP - k_{b2} \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - \left(k_{ub3} + k_3\right) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dP := -k_{b3} \cdot App \cdot P + \left(k_{ub3} + k_3\right) \cdot AppP - k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

**The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A - Ap - App - AK - ApK - AppP - ApP :$$
$$PCon := P_{Total} - P - AppP - ApP :$$
$$KCon := K_{Total} - K - AK - ApK :$$

**Kinetic constraints for case 1 symmetry (refer main text) are imposed on the original model.**

$$k_3 := k_1 : k_{ub3} := k_{ub1} : k_{b3} := k_{b1} :$$
$$k_4 := k_2 : k_{ub4} := k_{ub2} : k_{b4} := k_{b2} :$$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase also need to be equal for exact case 1 symmetry to be present. This is imposed as shown below.**

$$K_{Total} := P_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$ and $c_2$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ([K] & [P]) and the concentration of the partially modified substrate ([Ap]). In order to do this, we use the Maple command *solve*, which solves the supplied equation for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [AK]) is given below in detail.**

**The differential equation of [AK] is given by,**

$$\frac{d\,[AK]}{dt} = dAK$$

$$\frac{d\,[AK]}{dt} = c_1 \left(k_{ub1} + k_1\right) A\,K - \left(k_{ub1} + k_1\right) AK \qquad \textbf{(1)}$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [AK]). We in turn store this value (the solution returned by the *solve* command) in [AK]. This is performed by the following command.**

$$AK := solve(dAK, AK)$$

$$AK := K\,A\,c_1 \qquad \textbf{(2)}$$

**This operation is repeated for the other complexes and substrate forms as well.**

$$ApK := solve(dApK, ApK) :$$

$AppP := solve(dAppP, AppP)$ :
$ApP := solve(dApP, ApP)$ :

$A := solve(dA, A)$ :
$App := solve(dApp, App)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$App = \frac{K\,Ap\,c_2\,k_2}{c_1\,P\,k_1}$$
$$A = \frac{P\,Ap\,c_2\,k_2}{c_1\,K\,k_1}$$

## Proof for invariant in the asymmetric branches

**We know that PCon and KCon are both individually always equal to zero. Thus at a given steady state, PCon - KCon must also be equal to zero.**

**We thus introduce the term T = (KCon - PCon) = 0 and also introduce a new ratio, $\epsilon$ = [K]/[P]. Note: As discussed in the main text, the symmetric steady state is one where [K] = [P] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$T := KCon - PCon = 0$ :
$K := \epsilon \cdot P$ :

**The following Maple command (*simplify*), simplifies the expression algebraically.**

$$simplify(T) = -\frac{P\left(Ap\left(k_1 - k_2\right)c_2 + k_1\right)(\epsilon - 1)}{k_1} = 0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term $\left(Ap\left(k_1 - k_2\right)c_2 + k_1\right)$ in the expression needs to be zero. We note that this term is a expression in the partial substrate form [Ap] and kinetic constants. Thus solving this to isolate the partial substrate form we get the following. Here we use the *solve* command from Maple to solve T for [Ap] as shown below.**

$Ap = solve(T, Ap)$

$$Ap = -\frac{k_1}{\left(k_1 - k_2\right)c_2} \tag{1.1}$$

**Thus we can see that in an asymmetric steady state, the value of the partially modified substrate is fixed and is given by the above expression involving only a few key kinetic**

**parameters.**

## Necessary conditions for symmetry breaking

**Since substrate concentrations are always positive, the expression in equation 1.1 should be positive. As the involved terms (kinetic constants) are by definition positive, we get the necessary condition for asymmetric states to exist as follows.**

1. $k_2 > k_1$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient for an asymmetric steady state to exist for some positive $A_{Total}$ value. i.e. We show that upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that a feasible steady state in this context is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that the asymmetric states defined by the invariant concentration of [Ap] described above is indeed a feasible solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an asymmetric steady state, as seen above the concentration of [Ap] is fixed by a few kinetic constants,**

$$Ap := -\frac{k_1}{(k_1 - k_2)\, c_2} = -\frac{k_1}{(k_1 - k_2)\, c_2}$$

**The other variables in this asymmetric state are thus given by (we obtain this by using the correlation obtained earlier between the different concentrations)**

$$A = -\frac{k_2}{(k_1 - k_2)\, c_1\, \epsilon}$$

$$App = -\frac{\epsilon\, k_2}{(k_1 - k_2)\, c_1}$$

$$AK = -\frac{P\, k_2}{k_1 - k_2}$$

$$ApK = -\frac{\epsilon\, P\, k_1}{k_1 - k_2}$$

$$AppP = -\frac{P\, \epsilon\, k_2}{k_1 - k_2}$$

$$ApP = -\frac{P\,k_1}{k_1 - k_2}$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify(dA) = 0$$
$$simplify(dAp) = 0$$
$$simplify(dApp) = 0$$
$$simplify(dAK) = 0$$
$$simplify(dApK) = 0$$
$$simplify(dAppP) = 0$$
$$simplify(dApP) = 0$$
$$simplify(dK) = 0$$
$$simplify(dP) = 0$$

**Hence all that remains to be shown is that the variables (as described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition $(k_2 > k_1)$ is satisfied**
2. **P and $\epsilon$ are positive.**

**Now by using the total conservation of phosphatase in the system (PCon = 0), we obtain an algebraic expression for [P] in terms of $P_{Total}$, kinetic constants and $\epsilon$ as shown below. The Maple command *solve* is used for this purpose.**

$$P := solve(PCon, P) = -\frac{P_{Total}\left(k_1 - k_2\right)}{k_2\,(\epsilon + 1)}$$

**Hence if $\epsilon$ is positive P is automatically positive (provided necessary conditions are satisfied).**

**Thus this means that for every value of $\epsilon \neq 1$, all concentrations are positive and the conservation of kinase and phosphatase is also satisfied.**

**Since the concentrations are all positive there exists a unique finite $A_{Total}$ value for the given $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

## Prediction of position of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitchfork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two facts.**

**1. The system is still symmetric, hence [A] = [App] and [K] = [P].**
**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$$K := P : \quad \epsilon := 1 :$$
$$Ap := -\frac{k_1}{(k_1 - k_2)\, c_2} :$$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point)**

$$A_{Total} = simplify\left(solve\left(simplify(ACon), A_{Total}\right)\right)$$

$$A_{Total} = \frac{\left(-c_1\, c_2\, P_{Total} - 2\, c_2\right) k_2^{\,2} - k_1\, c_1\, k_2 + P_{Total}\, k_1^{\,2}\, c_1\, c_2}{(k_1 - k_2)\, c_1\, c_2\, k_2} \tag{4.1}$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figures (Fig 2A).**

# Double Site Ordered Distributive Phosphorylation : Separate Kinase Separate Phosphatase Case 1 Symmetry - Present and Cannot Break

In this Maple file we analytically show the infeasibility of symmetry breaking in ordered distributive DSP network with a unique kinase and unique phosphatase effecting phosphorylation and dephoshorylation on each modification site. We do this by first describing the model as a system of ODEs and imposing the kinetic constraints pertaining to case 1 symmetry. We further describe the enzyme and substrate conservations associated with the model. By solving for the steady state of the system of ODEs we obtain relations between substrate variables (concentrations) in terms of each other and the free enzyme concentrations. In ordered distributive DSP (with separate kinase separate phosphatase), $K1_{Total} = P2_{Total}$ and $K2_{Total} = P1_{Total}$ is required for exact case 1 symmetry to be present. We use this information to get a further simplified expression describing all possible steady states of the system. By subsequently analyzing the steady state of the system using the conservation conditions, we show that a steady state violating case 1 symmetry between [A] and [App] is not possible in this network for any choice of kinetics or total enzyme concentrations, thus ruling out symmetry breaking.

**We initialize the maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer to Appendix 2 figure 10). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA := k_4 \cdot ApP1 + k_{ub1} \cdot AK1 - k_{b1} \cdot A \cdot K1 :$$
$$dAp := k_1 \cdot AK1 + k_3 \cdot AppP2 + k_{ub2} \cdot ApK2 + k_{ub4} \cdot ApP1 - k_{b2} \cdot Ap \cdot K2 - k_{b4} \cdot Ap \cdot P1 :$$
$$dApp := k_2 \cdot ApK2 + k_{ub3} \cdot AppP2 - k_{b3} \cdot App \cdot P2 :$$

$$dAK1 := k_{b1} \cdot A \cdot K1 - (k_{ub1} + k_1) \cdot AK1 :$$
$$dApK2 := k_{b2} \cdot Ap \cdot K2 - (k_{ub2} + k_2) \cdot ApK2 :$$
$$dAppP2 := k_{b3} \cdot App \cdot P2 - (k_{ub3} + k_3) \cdot AppP2 :$$
$$dApP1 := k_{b4} \cdot Ap \cdot P1 - (k_{ub4} + k_4) \cdot ApP1 :$$

$$dK1 := -k_{b1} \cdot A \cdot K1 + (k_{ub1} + k_1) \cdot AK1 :$$
$$dK2 := -k_{b2} \cdot Ap \cdot K2 + (k_{ub2} + k_2) \cdot ApK2 :$$
$$dP2 := -k_{b3} \cdot App \cdot P2 + (k_{ub3} + k_3) \cdot AppP2 :$$
$$dP1 := -k_{b4} \cdot Ap \cdot P1 + (k_{ub4} + k_4) \cdot ApP1 :$$

**The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, P1Con, P2Con, K1Con and K2Con for the substrate**

**and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A - Ap - App - AK1 - ApK2 - AppP2 - ApP1 :$$
$$P1Con := P1_{Total} - P1 - ApP1 :$$
$$K1Con := K1_{Total} - K1 - AK1 :$$
$$P2Con := P2_{Total} - P2 - AppP2 :$$
$$K2Con := K2_{Total} - K2 - ApK2 :$$

**Kinetic constraints for case 1 symmetry (refer main text) are imposed on the original model.**

$$k_3 := k_1 : k_{ub3} := k_{ub1} : k_{b3} := k_{b1} :$$
$$k_4 := k_2 : k_{ub4} := k_{ub2} : k_{b4} := k_{b2} :$$

**In addition to the kinetic constraints the total enzyme concentrations of corresponding kinases and phosphatases also need to be equal for exact case 1 symmetry to be present. This is imposed as shown below.**

$$K1_{Total} := P2_{Total} :$$
$$K2_{Total} := P1_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$ and $c_2$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ([K1], [K2], [P1] and [P2]) and concentration of the partially modified substrate ([Ap]). In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [AK1]) is given below in detail.**
**The differential equation of [AK1] is given by,**

$$\frac{d\,[AK1]}{dt} = dAK1 \qquad \frac{d\,[AK1]}{dt} = c_1 \left(k_{ub1} + k_1\right) A\,K1 - \left(k_{ub1} + k_1\right) AK1$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [AK1]). We in turn store this value (the solution returned by the *solve* command) in [AK1]. This is performed by the following command.**

$$AK1 := solve(dAK1, AK1)$$
$$AK1 := K1\,A\,c_1 \qquad \textbf{(1)}$$

**This operation is repeated for the other complexes and substrate forms as well.**

$ApK2 := solve(dApK2, ApK2)$ :
$AppP2 := solve(dAppP2, AppP2)$ :
$ApP1 := solve(dApP1, ApP1)$ :

$A := solve(dA, A)$ :
$App := solve(dApp, App)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A = \frac{P1\ Ap\ c_2\ k_2}{c_1\ K1\ k_1}$$
$$App = \frac{K2\ Ap\ c_2\ k_2}{P2\ c_1\ k_1}$$

## Proof for infeasibility of symmetry breaking

**We know that P1Con, K1Con, P2Con and K2Con are all individually equal to zero always. Thus at a given steady state, [K2Con - P1Con] and [K1Con - P2Con] must also be equal to zero. We thus introduce terms T = (K2Con - P1Con) = 0 and Q = (K1Con - P2Con) = 0.**

$T := (K2Con - P1Con) = 0$ :
$Q := (K1Con - P2Con) = 0$ :

**Simplifying T = [K2Con - P1Con] using the *simplify* command, we get the following expression.**

$$simplify(T)\quad -(K2 - P1)\left(Ap\ c_2 + 1\right) = 0$$

**Since all kinetic constants and concentrations of variables are positive, in order for T to be equal to zero, [K2] is necessarily equal to [P1]. Using this information and simplifying Q = [K1Con - P2Con] = 0 we get the following expression.**

$K2 := P1$ :
$$simplify(Q)\quad -K1 + P2 = 0$$

**Thus [K1] is equal to [P2] at any given steady state. Using this information, the earlier expression for [A] and [App] reduces to the following.**

$K1 := P2$ :

$$A = \frac{P1\ Ap\ c_2\ k_2}{c_1\ P2\ k_1}$$

$$App = \frac{P1\ Ap\ c_2\ k_2}{c_1\ P2\ k_1}$$

**Thus we can see that irrespective of enzyme concentrations and kinetic parameters, [A] is always equal to [App]. i.e. There is no possibility of an asymmetric branch in this model implying infeasibility of symmetry breaking.**

# Triple Site Ordered Distributive Phosphorylation : Common Kinase Common Phosphatase Case 1 symmetry - Present and Breaks

In this Maple file we analytically show the presence of symmetry breaking in ordered distributive TSP network with common kinase and common phosphatase effecting phosphorylation and dephoshorylation respectively. We do this by first describing the model as a system of ODEs and imposing the kinetic constraints pertaining to case 1 symmetry. We further describe the substrate and enzyme conservations associated with the model. By solving for the steady state of the system of ODEs we obtain relations between substrate variables (concentrations) in terms of each other and the free enzyme concentrations. In ordered distributive TSP (with common kinase common phosphatase), $K_{Total} = P_{Total}$ is required for exact case 1 symmetry to be present. We use this information to get a further simplified expression describing all possible steady states of the system. With case 1 symmetry in the ordered distributive MSP, the free enzyme kinase and phosphatase share a strict symmetry in the symmetric steady state ([K] = [P]). Thus by isolating steady states not of this type from the earlier correlation, we ascertain the features of the asymmetric state emerging from symmetry breaking. This procedure is carried out in detail below using built in Maple commands.

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The ordered distributive TSP system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer to Appendix 2 figure 10). Here dA represents d[A]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA := k_{ub1}\cdot AK + k_6\cdot ApP - k_{b1}\cdot A\cdot K :$$
$$dAp := k_1\cdot AK + k_5\cdot AppP + k_{ub2}\cdot ApK + k_{ub6}\cdot ApP - k_{b6}\cdot Ap\cdot P - k_{b2}\cdot Ap\cdot K :$$
$$dApp := k_2\cdot ApK + k_4\cdot ApppP + k_{ub3}\cdot AppK + k_{ub5}\cdot AppP - k_{b3}\cdot App\cdot K - k_{b5}\cdot App\cdot P :$$
$$dAppp := k_3\cdot AppK + k_{ub4}\cdot ApppP - k_{b4}\cdot Appp\cdot P :$$

$$dAK := k_{b1}\cdot A\cdot K - \left(k_{ub1} + k_1\right)\cdot AK :$$
$$dApK := k_{b2}\cdot Ap\cdot K - \left(k_{ub2} + k_2\right)\cdot ApK :$$
$$dAppK := k_{b3}\cdot App\cdot K - \left(k_{ub3} + k_3\right)\cdot AppK :$$
$$dApppP := k_{b4}\cdot Appp\cdot P - \left(k_{ub4} + k_4\right)\cdot ApppP :$$
$$dAppP := k_{b5}\cdot App\cdot P - \left(k_{ub5} + k_5\right)\cdot AppP :$$
$$dApP := k_{b6}\cdot Ap\cdot P - \left(k_{ub6} + k_6\right)\cdot ApP :$$

$$dK := -k_{b1}\cdot A\cdot K - k_{b2}\cdot Ap\cdot K - k_{b3}\cdot App\cdot K + \left(k_{ub1} + k_1\right)\cdot AK + \left(k_{ub2} + k_2\right)\cdot ApK + \left(k_{ub3} + k_3\right)\cdot AppK :$$
$$dP := -k_{b4}\cdot Appp\cdot P - k_{b5}\cdot App\cdot P - k_{b6}\cdot Ap\cdot P + \left(k_{ub4} + k_4\right)\cdot ApppP + \left(k_{ub5} + k_5\right)\cdot AppP + \left(k_{ub6}\right.$$

$+ k_6 \cdot ApP :$

**The above equations are also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$ACon := A_{Total} - A - Ap - App - Appp - AK - ApP - AppP - ApppP - ApK - AppK :$
$KCon := K_{Total} - K - AK - ApK - AppK :$
$PCon := P_{Total} - P - ApP - AppP - ApppP :$

**Kinetic constraints for case 1 symmetry (refer main text) are imposed on the original model.**

$k_4 := k_1 : k_{ub4} := k_{ub1} : k_{b4} := k_{b1} :$
$k_5 := k_2 : k_{ub5} := k_{ub2} : k_{b5} := k_{b2} :$
$k_6 := k_3 : k_{ub6} := k_{ub3} : k_{b6} := k_{b3} :$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase need to be equal for exact case 1 symmetry to be present. This is imposed as shown below.**

$K_{Total} := P_{Total} :$

**At this stage we introduce auxiliary constants $c_1$, $c_2$ and $c_3$ in place of the binding constants so as to make further analytical expressions more accessible.**

$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$
$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$
$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) :$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ([K] & [P]) and the concentration of the partially modified substrate ([Ap]). In order to do this, we use the Maple command *solve*, which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [AK]) is given below in detail.**

**The differential equation of [AK] is given by,**

$$\frac{d\,[AK]}{dt} = dAK \qquad \frac{d\,[AK]}{dt} = c_1 \left(k_{ub1} + k_1\right) A\,K - \left(k_{ub1} + k_1\right) AK$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [AK]). We in turn store this value (the solution returned by the *solve* command) in [AK]. This is performed by the following command.**

$AK := solve(dAK, AK)$

$$AK := K\,A\,c_1 \tag{1}$$

**This operation is repeated for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions of the substrate concentrations are extracted from this vector using the *eval* command.**

$ApK := solve(dApK, ApK)$ :
$AppK := solve(dAppK, AppK)$ :
$ApP := solve(dApP, ApP)$ :
$AppP := solve(dAppP, AppP)$ :
$ApppP := solve(dApppP, ApppP)$ :

$Sol := solve(\{dA, dAppp, dApp\}, \{A, Appp, App\})$ :

$A := eval(A, Sol)$ :
$Ap := eval(Ap, Sol)$ :
$App := eval(App, Sol)$ :
$Appp := eval(Appp, Sol)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A = \frac{P\,Ap\,c_3\,k_3}{K\,c_1\,k_1}$$
$$Ap = Ap$$
$$App = \frac{Ap\,K}{P}$$
$$Appp = \frac{Ap\,K^2\,c_3\,k_3}{P^2\,c_1\,k_1}$$

## Proof for invariant in the asymmetric branches

**We know that PCon and KCon are both individually equal to zero always. Thus at a given steady state, PCon - KCon must also be equal to zero.**

**Here we introduce the term T = (KCon - PCon) = 0 and also introduce a new ratio, $\epsilon$ = [K]/[P]. Note: As discussed in the main text, the symmetric steady state is one where [K] = [P] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$T := KCon - PCon = 0$ :
$K := \epsilon \cdot P$ :

**The following command (*simplify*), simplifies the expression T algebraically as shown below**

$$simplify(T) = -\frac{\left(Ap\ (\epsilon+1)\ \left(k_1-k_3\right)c_3+k_1\right)P\ (\epsilon-1)}{k_1} = 0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the expression $Ap\ (\epsilon+1)\ \left(k_1-k_3\right)c_3+k_1$ needs to be zero. This term is an expression involving of the concentrations of the partial substrate form, $\epsilon$ and kinetic constants. Thus solving this to isolate the partial substrate form in terms of the kinetic parameters and $\epsilon$, we get the following. Here we use the *solve* command from Maple to solve T for [Ap].**

$$Ap := solve(T, Ap) = -\frac{k_1}{\left(\epsilon k_1-\epsilon k_3+k_1-k_3\right)c_3}$$

**Using this information, we evaluate [App] from the correlations obtained earlier**

$$App = -\frac{k_1\,\epsilon}{\left(\epsilon k_1-\epsilon k_3+k_1-k_3\right)c_3}$$

**We can see a pattern here. Adding [Ap] and [App] we get,**

$$simplify(Ap + App)$$

$$-\frac{k_1}{c_3\left(k_1-k_3\right)} \qquad \textbf{(1.1)}$$

**Thus we can see that in an asymmetric steady state, the sum of the partially modified substrates is fixed and is given by the above expression involving only a few key kinetic parameters.**

## Necessary conditions

**Since substrate concentrations are always positive, the expression in equation 1.1 should be positive. This gives us the necessary condition for an asymmetric state to exist as follows.**

1. $k_3 > k_1$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient for an asymmetric steady state to exist for some positive $A_{Total}$ value. i.e. We show that upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that a feasible steady state in this context is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that the asymmetric states defined by the invariant concentration of [Ap] and [App] described above is indeed a feasible solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an asymmetric steady state, as seen above the concentration of [Ap] and [App] are fixed by a few kinetic constants,**

$$Ap := simplify\left(-\frac{k_1}{(\epsilon k_1 - \epsilon k_3 + k_1 - k_3)\, c_3}\right) = -\frac{k_1}{(\epsilon+1)\,(k_1 - k_3)\, c_3}$$

$$App := simplify\left(-\frac{k_1 \epsilon}{(\epsilon k_1 - \epsilon k_3 + k_1 - k_3)\, c_3}\right) = -\frac{k_1 \epsilon}{(\epsilon+1)\,(k_1 - k_3)\, c_3}$$

**The other variables in this asymmetric state are thus given by (we obtain this by using the correlation obtained earlier between the different concentrations)**

$$A = -\frac{k_3}{(\epsilon+1)\,(k_1 - k_3)\,\epsilon\, c_1}$$

$$Appp = -\frac{\epsilon^2 k_3}{(\epsilon+1)\,(k_1 - k_3)\, c_1}$$

$$AK = -\frac{P k_3}{(\epsilon+1)\,(k_1 - k_3)}$$

$$ApK = -\frac{\epsilon P k_1 c_2}{(\epsilon+1)\,(k_1 - k_3)\, c_3}$$

$$AppK = -\frac{\epsilon^2 P k_1}{(\epsilon+1)\,(k_1 - k_3)}$$

$$ApppP = -\frac{P \epsilon^2 k_3}{(\epsilon+1)\,(k_1 - k_3)}$$

$$AppP = -\frac{\epsilon P k_1 c_2}{(\epsilon+1)\,(k_1 - k_3)\, c_3}$$

$$ApP = -\frac{P k_1}{(\epsilon+1)\,(k_1 - k_3)}$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify(dA) = 0$$
$$simplify(dAp) = 0$$
$$simplify(dApp) = 0$$
$$simplify(dAppp) = 0$$
$$simplify(dAK) = 0$$
$$simplify(dApK) = 0$$

$simplify(dAppK) = 0$
$simplify(dApppP) = 0$
$simplify(dAppP) = 0$
$simplify(dApP) = 0$
$simplify(dK) = 0$
$simplify(dP) = 0$

**Hence all that remains to be shown is that the variables (as described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition $(k_3 > k_1)$ is satisfied**
2. **P and $\epsilon$ are positive.**

**Now by using the total conservation of phosphatase in the system (PCon = 0), we obtain an algebraic expression for [P] in terms of $P_{Total}$, kinetic constants and $\epsilon$ as shown below. The Maple command *solve* is used for this purpose as shown below.**

$P := simplify(solve(PCon, P))$

$$P := -\frac{P_{Total}\,(\epsilon+1)\,(k_1-k_3)\,c_3}{\epsilon^2 k_3 c_3 + ((-k_1+k_3)c_3 + c_2 k_1)\,\epsilon + c_3 k_3} \tag{3.1}$$

**Hence if $\epsilon$ is positive P is automatically positive (provided necessary conditions are satisfied).**

**Thus this means that for every positive value of $\epsilon \neq 1$, all concentrations are positive and the conservation of kinase and phosphatase is also satisfied.**

**Since the concentrations are all positive, there exists a unique $A_{Total}$ value for every $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

## Prediction of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitch fork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two insights.**

**1. The system is still symmetric, hence [A] = [Appp], [Ap] = [App] and [K] = [P].**
**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$$K := P : \epsilon := 1 :$$

$$Ap := -\frac{k_1}{(\epsilon+1)\left(k_1 - k_3\right) c_3} :$$

$$App := -\frac{k_1 \epsilon}{(\epsilon+1)\left(k_1 - k_3\right) c_3} :$$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point)**

$$A_{Total} = \left(collect\left(solve\left(ACon, A_{Total}\right), P_{Total}\right)\right)$$

$$A_{Total} = \frac{\left(2 c_1 c_2 c_3 k_1^{\,2} - 2 c_1 c_2 c_3 k_1 k_3 + 2 c_1 c_3^{\,2} k_1^{\,2} - 2 c_1 c_3^{\,2} k_3^{\,2}\right) P_{Total}}{\left(k_1 - k_3\right) c_1 c_3 \left(c_2 k_1 - k_1 c_3 + 3 c_3 k_3\right)} + \frac{-c_1 c_2 k_1^{\,2} + c_1 c_3 k_1^{\,2} - 3 c_1 c_3 k_1 k_3 - c_2 c_3 k_1 k_3 + c_3^{\,2} k_1 k_3 - 3 c_3^{\,2} k_3^{\,2}}{\left(k_1 - k_3\right) c_1 c_3 \left(c_2 k_1 - k_1 c_3 + 3 c_3 k_3\right)} \tag{4.1}$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figure (Appendix 2 figure 8).**

## Random DSP System 1 : Common Kinase Common Phosphatase

| Case 1 | Case 2 | Case 3 |
|---|---|---|
| **Present and Breaks Invariant identified** | **Present and Can't Break** | **Present and Breaks Invariant Identified** |

In this Maple file we analytically show the presence of case 1 and case 3 symmetry breaking in Random DSP network with common kinase and common phosphatase (System 1) effecting phosphorylation and dephoshorylation respectively. We also show the infeasibility of case 2 symmetry breaking in this network.

In each case, we do this by first describing the model as a system of ODEs along with the associated substrate and enzyme conservations. We then impose the kinetic constraints (and constraints on total enzyme concentrations) pertaining to the specific case of symmetry. By algebraically solving for the steady state of the resulting system of ODEs we obtain relations between concentrations of the substrate variables in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. In Random System 1, case 1 symmetry requires symmetry between [K] & [P] and [A00] & [A11]; case 2 symmetry breaking requires symmetry between [A01] & [A10]; case 3 symmetry breaking requires symmetry between [K] & [P], [A01] & [A10] and [A00] & [A11]. In each case by leveraging this insight and isolating steady states not of this type, we ascertain the features of the asymmetric state emerging through symmetry breaking. In a similar way we show the infeasibility of case 2 symmetry breaking by revealing that [A01] and [A10] are always equal for any given feasible steady state. These procedures are carried out in detail below using built in Maple commands.

Note: A subscript is used to distinguish between the two different complexes formed between [K] and [A00] ([$A00K_1$ and $A00K_2$]). Similarly a subscript is used to distinguish between the two distinct complexes formed between [P] and [A11] ([$A11P_1$ and $A11P_2$]).

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer to Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := k_4 \cdot A01P + a_4 \cdot A10P + k_{ub1} \cdot A00K_1 + a_{ub1} \cdot A00K_2 - k_{b1} \cdot A00 \cdot K - a_{b1} \cdot A00 \cdot K :$$
$$dA01 := k_1 \cdot A00K_1 + k_3 \cdot A11P_1 + k_{ub2} \cdot A01K + k_{ub4} \cdot A01P - k_{b2} \cdot A01 \cdot K - k_{b4} \cdot A01 \cdot P :$$
$$dA10 := a_1 \cdot A00K_2 + a_3 \cdot A11P_2 + a_{ub2} \cdot A10K + a_{ub4} \cdot A10P - a_{b2} \cdot A10 \cdot K - a_{b4} \cdot A10 \cdot P :$$
$$dA11 := k_2 \cdot A01K + a_2 \cdot A10K + k_{ub3} \cdot A11P_1 + a_{ub3} \cdot A11P_2 - k_{b3} \cdot A11 \cdot P - a_{b3} \cdot A11 \cdot P :$$

$$dA00K_1 := k_{b1} \cdot A00 \cdot K - \left(k_1 + k_{ub1}\right) \cdot A00K_1 :$$

$$dA01K := k_{b2}\cdot A01\cdot K - \left(k_2 + k_{ub2}\right)\cdot A01K :$$
$$dA00K_2 := a_{b1}\cdot A00\cdot K - \left(a_{ub1} + a_1\right)\cdot A00K_2 :$$
$$dA10K := a_{b2}\cdot A10\cdot K - \left(a_{ub2} + a_2\right)\cdot A10K :$$
$$dA11P_1 := k_{b3}\cdot A11\cdot P - \left(k_{ub3} + k_3\right)\cdot A11P_1 :$$
$$dA01P := k_{b4}\cdot A01\cdot P - \left(k_{ub4} + k_4\right)\cdot A01P :$$
$$dA11P_2 := a_{b3}\cdot A11\cdot P - \left(a_{ub3} + a_3\right)\cdot A11P_2 :$$
$$dA10P := a_{b4}\cdot A10\cdot P - \left(a_{ub4} + a_4\right)\cdot A10P :$$

$$dK := -k_{b1}\cdot A00\cdot K + \left(k_1 + k_{ub1}\right)\cdot A00K_1 - a_{b1}\cdot A00\cdot K + \left(a_{ub1} + a_1\right)\cdot A00K_2 - k_{b2}\cdot A01\cdot K + \left(k_2 + k_{ub2}\right)\cdot A01K - a_{b2}\cdot A10\cdot K + \left(a_{ub2} + a_2\right)\cdot A10K :$$
$$dP := -k_{b3}\cdot A11\cdot P + \left(k_{ub3} + k_3\right)\cdot A11P_1 - a_{b3}\cdot A11\cdot P + \left(a_{ub3} + a_3\right)\cdot A11P_2 - k_{b4}\cdot A01\cdot P + \left(k_{ub4} + k_4\right)\cdot A01P - a_{b4}\cdot A10\cdot P + \left(a_{ub4} + a_4\right)\cdot A10P :$$

**The above equations are also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K_1 - A01K - A00K_2 - A10K - A11P_1 - A10P - A11P_2 - A01P :$$
$$KCon := K_{Total} - K - A00K_1 - A10K - A00K_2 - A01K :$$
$$PCon := P_{Total} - P - A11P_1 - A10P - A11P_2 - A01P :$$

**Until now we have modelled the System 1 Random DSP network with common kinase effecting phosphorylation and a common phosphatase for dephosphorylation - without any impositions on kinetics or total concentrations of enzymes. In the following segments, we specifically do this for each class of symmetry. The codes for each symmetry are modular and in order to run a particular symmetry, please run the code until this point and then run only the code for the specific class of symmetry.**
**Note: Do not run the whole script at the same time, as this will impose all symmetries at the same time and give incorrect results.**

## Case 1 Symmetry : Present and Breaks

**Kinetic constraints for case 1 symmetry (refer main text) are imposed on the original model.**

$$k_3 := k_1 : k_{b3} := k_{b1} : k_{ub3} := k_{ub1} :$$
$$k_4 := k_2 : k_{b4} := k_{b2} : k_{ub4} := k_{ub2} :$$

$$a_3 := a_1 : a_{b3} := a_{b1} : a_{ub3} := a_{ub1} :$$
$$a_4 := a_2 : a_{b4} := a_{b2} : a_{ub4} := a_{ub2} :$$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase need to be equal for exact case 1 symmetry to be present. This is imposed as shown below.**

$$K_{Total} := P_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $d_1$ and $d_2$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$

$$a_{b1} := d_1 \cdot \left(a_1 + a_{ub1}\right) :$$
$$a_{b2} := d_2 \cdot \left(a_2 + a_{ub2}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K] & [P]) and concentration of the partially modified substrate ([A10]). In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K$_1$]) is given below in detail.**

**The differential equation of [A00K$_1$] is given by,**

$$\frac{d\left[A00K_1\right]}{dt} = dA00K_1 \qquad \frac{d\left[A00K_1\right]}{dt} = c_1 \left(k_1 + k_{ub1}\right) A00\,K - \left(k_1 + k_{ub1}\right) A00K_1$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K$_1$]). We in turn store this value (the solution returned by the *solve* command) in [A00K$_1$]. This is performed by the following command.**

$$A00K_1 := solve\left(dA00K_1, A00K_1\right)$$

$$A00K_1 := K\,A00\,c_1 \tag{1.1}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A10K := solve(dA10K, A10K) :$$
$$A00K_2 := solve\left(dA00K_2, A00K_2\right) :$$
$$A01K := solve(dA01K, A01K) :$$
$$A11P_1 := solve\left(dA11P_1, A11P_1\right) :$$

$A11P_2 := solve(dA11P_2, A11P_2)$ :
$A01P := solve(dA01P, A01P)$ :
$A10P := solve(dA10P, A10P)$ :

$Sol := solve(\{dA00, dA01, dA11\}, \{A00, A01, A11\})$ :

$A00 := eval(A00, Sol)$ :
$A01 := eval(A01, Sol)$ :
$A11 := eval(A11, Sol)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = \frac{A10\, d_2\, a_2\, P}{a_1\, d_1\, K}$$
$$A01 = \frac{A10\, d_2\, a_2\, c_1\, k_1}{a_1\, d_1\, c_2\, k_2}$$
$$A10 = A10$$
$$A11 = \frac{K\, A10\, d_2\, a_2}{P\, a_1\, d_1}$$

## Proof for invariant in the asymmetric branches

**We know that PCon and KCon are both individually equal to zero always. Thus at a given steady state, PCon - KCon must also be equal to zero.**

**We thus introduce the term T = (KCon - PCon) = 0 and also introduce a new ratio, $\epsilon$ = [K]/[P]. Note: As discussed in the main text, the symmetric steady state is one where [K] = [P] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$T := KCon - PCon = 0$ :
$K := \epsilon \cdot P$ :

**The following command (*simplify*), simplifies the expression**

$$simplify(T) = -\frac{(\epsilon - 1)\left(\left(\left(\left(a_1 - a_2\right) d_1 - a_2\, c_1\right) A10\, d_2 + a_1\, d_1\right) k_2 + A10\, a_2\, c_1\, d_2\, k_1\right) P}{a_1\, d_1\, k_2} = 0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term $\left(\left(A10\left(\left(k_1 - k_2\right) c_1 - d_1\, k_2\right) c_2 + c_1\, k_1\right) p_2 + A10\, c_2\, d_1\, k_2\, p_1\right)$ in the expression needs to be zero. We note that this term is a expression in the partial substrate form and kinetic**

**constants. Thus solving this to isolate the partial substrate form we get the following. Here we use the *solve* command from Maple to solve T for [A10] as shown below.**

$$A10 := simplify(solve(T, A10))$$

$$A10 := -\frac{a_1 d_1 k_2}{d_2 \left( \left( \left( -c_1 - d_1 \right) a_2 + a_1 d_1 \right) k_2 + a_2 c_1 k_1 \right)} \quad \textbf{(1.1.1)}$$

**We can thus see that in an asymmetric state [A10] is always a constant given by only a few key kinetic parameters. Using this expression back in the relation between [A01] we ascertain that [A01] is also fixed at a constant concentration at an asymmetric state.**

$$A01 := simplify(A01)$$

$$A01 := -\frac{a_2 c_1 k_1}{\left( \left( \left( k_1 - k_2 \right) c_1 - d_1 k_2 \right) a_2 + a_1 d_1 k_2 \right) c_2} \quad \textbf{(1.1.2)}$$

**Thus we can see that in an asymmetric steady state, the value of the partially modified substrates ([A01] and [A10]) is fixed and is given by the above expressions involving only a few key kinetic parameters.**

## Necessary Conditions

**Since substrate concentration are always necessarily positive, the expression in equation 1.1.1 and 1.1.2 should be positive. The numerator is only a function of kinetic parameters which are always positive, thus the denominator must necessarily be negative to ensure that the resulting concentration is positive. This gives us the necessary condition for the asymmetric state to exist as follows.**

1. $c_1 a_2 \left( k_2 - k_1 \right) + d_1 k_2 \left( a_2 - a_1 \right) > 0$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient for an asymmetric steady state to exist for some positive $A_{Total}$ value. or upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that in the context of this system, a feasible steady state is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that an asymmetric state defined by the invariant concentrations described above indeed is a solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an asymmetric steady state, as seen above [A01] and [A10] are fixed by a few kinetic constants,**

$$A10 := expand\left(-\frac{a_1 d_1 k_2}{d_2\left(\left(\left(-c_1 - d_1\right) a_2 + a_1 d_1\right) k_2 + a_2 c_1 k_1\right)}\right) = -\frac{a_1 d_1 k_2}{\left(a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2\right) d_2}$$

$$A01 := expand\left(-\frac{a_2 c_1 k_1}{c_2\left(\left(\left(k_1 - k_2\right) c_1 - d_1 k_2\right) a_2 + a_1 d_1 k_2\right)}\right) = -\frac{a_2 c_1 k_1}{\left(a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2\right) c_2}$$

**The other variables in this asymmetric state are thus given by**

$$expand(A00) = -\frac{k_2 a_2}{\left(a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2\right)\epsilon}$$

$$expand(A11) = -\frac{\epsilon k_2 a_2}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand\left(A00K_1\right) = -\frac{P k_2 a_2 c_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand\left(A00K_2\right) = -\frac{P k_2 a_2 d_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand(A01K) = -\frac{\epsilon P a_2 c_1 k_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand(A10K) = -\frac{\epsilon P a_1 d_1 k_2}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand\left(A11P_1\right) = -\frac{P \epsilon k_2 a_2 c_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand\left(A11P_2\right) = -\frac{P \epsilon k_2 a_2 d_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand(A01P) = -\frac{P a_2 c_1 k_1}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

$$expand(A10P) = -\frac{P a_1 d_1 k_2}{a_1 d_1 k_2 + a_2 c_1 k_1 - a_2 c_1 k_2 - a_2 d_1 k_2}$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify(dA00) = 0$$
$$simplify(dA11) = 0$$
$$simplify(dA01) = 0$$

$simplify(dA10) = 0$
$simplify(dA00K_1) = 0$
$simplify(dA00K_2) = 0$
$simplify(dA01P) = 0$
$simplify(dA10P) = 0$
$simplify(dA01K) = 0$
$simplify(dA10K) = 0$
$simplify(dA11P_1) = 0$
$simplify(dA11P_2) = 0$
$simplify(dP) = 0$
$simplify(dK) = 0$

**Hence all that remains to be shown is that the variables (as described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition is satisfied**
2. **P and $\epsilon$ are positive.**

**Now by using the total conservation of phosphatase in the system (PCon = 0), we obtain an algebraic expression for [P] in terms of $P_{Total}$, kinetic constants and $\epsilon$ as shown below. The Maple command *solve* is used for this purpose as shown below.**

$P := expand(solve(PCon, P))$

$$P := -\frac{P_{Total}\, a_1\, d_1}{a_2\left(c_1\,\epsilon + d_1\,\epsilon + c_1 + d_1\right)} - \frac{P_{Total}\, c_1\, k_1}{k_2\left(c_1\,\epsilon + d_1\,\epsilon + c_1 + d_1\right)} + \frac{P_{Total}\, c_1}{c_1\,\epsilon + d_1\,\epsilon + c_1 + d_1} + \frac{P_{Total}\, d_1}{c_1\,\epsilon + d_1\,\epsilon + c_1 + d_1} \tag{1.3.1}$$

$simplify(P)$

$$-\frac{\left(\left(\left(k_1 - k_2\right) c_1 - d_1\, k_2\right) a_2 + a_1\, d_1\, k_2\right) P_{Total}}{a_2\, k_2\,(\epsilon + 1)\left(c_1 + d_1\right)} \tag{1.3.2}$$

**Hence if $\epsilon$ is positive P is automatically positive (provided necessary conditions are satisfied) . Hence we have shown that the individual conservation equations for the enzymes (KCon and PCon) and the system of ODEs is solved by the expressions above, and represent feasible solutions provided $\epsilon$ is positive.**

**Since the concentrations are all positive there exists a unique $A_{Total}$ value for every $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

### Prediction of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitch fork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two insights.**

**1. The system is still symmetric, hence [A00] = [A11] and [K] = [P].**
**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$$K := P : \epsilon := 1 :$$

$$A10 := -\frac{a_1 d_1 k_2}{d_2 \left( \left( \left( -c_1 - d_1 \right) a_2 + a_1 d_1 \right) k_2 + a_2 c_1 k_1 \right)} :$$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point).**

$$A_{Total} = simplify\left( solve\left( ACon, A_{Total} \right) \right)$$

$$A_{Total} = \left( d_2 \left( -c_2 \left( c_1 + d_1 \right) \left( P_{Total}\, c_1 + P_{Total}\, d_1 + 2 \right) k_2^{\,2} - c_1 k_1 \left( c_1 + d_1 \right) k_2 + c_1^{\,2} c_2 k_1^{\,2} P_{Total} \right) a_2^{\,2} + 2 d_1 \left( \left( -\frac{c_1}{2} - \frac{d_1}{2} \right) k_2 + c_1 d_2 k_1 P_{Total} \right) c_2 k_2 a_1 a_2 + a_1^{\,2} c_2 d_1^{\,2} d_2 k_2^{\,2} P_{Total} \right) \Big/ \left( \left( c_1 + d_1 \right) a_2 c_2 \left( \left( \left( -c_1 - d_1 \right) k_2 + c_1 k_1 \right) a_2 + a_1 d_1 k_2 \right) d_2 k_2 \right) \qquad \textbf{(1.4.1)}$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figures (Fig 2B).**

## Case 2 Symmetry : Present and Can't Break

**Kinetic constraints for case 2 symmetry (refer main text) are imposed on the original model.**

$$a_1 := k_1 : a_{b1} := k_{b1} : a_{ub1} := k_{ub1} :$$
$$a_2 := k_2 : a_{b2} := k_{b2} : a_{ub2} := k_{ub2} :$$
$$a_3 := k_3 : a_{b3} := k_{b3} : a_{ub3} := k_{ub3} :$$
$$a_4 := k_4 : a_{b4} := k_{b4} : a_{ub4} := k_{ub4} :$$

**There are no constraints on the total enzyme concentrations for case 2 symmetry to be present. At this stage we introduce auxiliary constants $c_1$, $c_2$, $c_3$ and $c_4$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) :$$
$$k_{b2} := c_2 \cdot (k_2 + k_{ub2}) :$$
$$k_{b3} := c_3 \cdot (k_3 + k_{ub3}) :$$
$$k_{b4} := c_4 \cdot (k_4 + k_{ub4}) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K] & [P]) and concentration of the completely unmodified substrate ([A00]). In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K$_1$]) is given below in detail.**

**The differential equation of [A00K$_1$] is given by,**

$$\frac{d\left[A00K_1\right]}{dt} = dA00K_1$$

$$\frac{d\left[A00K_1\right]}{dt} = c_1 \left(k_1 + k_{ub1}\right) A00\,K - \left(k_1 + k_{ub1}\right) A00K_1 \qquad \textbf{(2.1)}$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K$_1$]). We in turn store this value (the solution returned by the *solve* command) in [A00K$_1$]. This is performed by the following command.**

$$A00K_1 := solve\left(dA00K_1, A00K_1\right)$$

$$A00K_1 := K\,A00\,c_1 \qquad \textbf{(2.2)}$$

**This operation is performed for the other complexes and substrate forms as well. Here we**

**simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A10K := solve(dA10K, A10K) :$$
$$A00K_2 := solve\left(dA00K_2, A00K_2\right) :$$
$$A01K := solve(dA01K, A01K) :$$
$$A11P_1 := solve\left(dA11P_1, A11P_1\right) :$$
$$A11P_2 := solve\left(dA11P_2, A11P_2\right) :$$
$$A01P := solve(dA01P, A01P) :$$
$$A10P := solve(dA10P, A10P) :$$

$$Sol := solve(\{dA10, dA01, dA11\}, \{A10, A01, A11\}) :$$

$$A10 := eval(A10, Sol) :$$
$$A01 := eval(A01, Sol) :$$
$$A11 := eval(A11, Sol) :$$

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = A00$$
$$A01 = \frac{K\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$
$$A10 = \frac{K\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$
$$A11 = \frac{A00\,K^2\,c_1\,c_2\,k_1\,k_2}{c_4\,k_4\,P^2\,c_3\,k_3}$$

## Proof for impossibility of symmetry breaking

**Thus from this we can clearly see that irrespective of kinetic parameters, the concentration of [A01] is always going to be equal to the concentration of [A10]. Thus there is no scope for any asymmetric steady state or case 2 symmetry breaking.**

$$A01 = \frac{K\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$

$$A10 = \frac{K\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$

## Case 3 Symmetry : Present and Breaks

**Kinetic constraints for case 3 symmetry to be seen (refer main text) are imposed on the original model.**

$$a_3 := k_1 : a_{b3} := k_{b1} : a_{ub3} := k_{ub1} :$$
$$a_4 := k_2 : a_{b4} := k_{b2} : a_{ub4} := k_{ub2} :$$
$$a_1 := k_3 : a_{b1} := k_{b3} : a_{ub1} := k_{ub3} :$$
$$a_2 := k_4 : a_{b2} := k_{b4} : a_{ub2} := k_{ub4} :$$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase also need to be equal for exact case 3 symmetry to be present. This is imposed as shown below.**

$$K_{Total} := P_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $c_3$ and $c_4$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) :$$
$$k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K] & [P]) and concentration of the fully modified substrate ([A11]). In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K$_1$]) is given below in detail.**

**The differential equation of [A00K$_1$] is given by,**

$$\frac{d\left[A00K_1\right]}{dt} = dA00K_1$$

$$\frac{d\left[A00K_1\right]}{dt} = c_1\left(k_1 + k_{ub1}\right) A00\,K - \left(k_1 + k_{ub1}\right) A00K_1 \qquad \textbf{(3.1)}$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K$_1$]). We in turn store this value (the solution returned by the *solve* command) in [A00K$_1$]. This is performed by the following command.**

$A00K_1 := solve\left(dA00K_1, A00K_1\right)$

$$A00K_1 := K\,A00\,c_1 \tag{3.2}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$A10K := solve(dA10K, A10K)$ :
$A00K_2 := solve\left(dA00K_2, A00K_2\right)$ :
$A01K := solve(dA01K, A01K)$ :
$A11P_1 := solve\left(dA11P_1, A11P_1\right)$ :
$A11P_2 := solve\left(dA11P_2, A11P_2\right)$ :
$A01P := solve(dA01P, A01P)$ :
$A10P := solve(dA10P, A10P)$ :

$Sol := solve(\{dA00, dA01, dA10, dA11\}, \{A00, A01, A10, A11\})$ :

$A00 := eval(A00, Sol)$ :
$A01 := eval(A01, Sol)$ :
$A10 := eval(A10, Sol)$ :
$A11 := eval(A11, Sol)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$A00 =$

$$\frac{P^2\,A11\left(K\,c_1\,c_2^{\,2}\,k_1\,k_2^{\,2}+K\,c_3\,c_4^{\,2}\,k_3\,k_4^{\,2}+P\,c_1\,c_2\,c_4\,k_1\,k_2\,k_4+P\,c_2\,c_3\,c_4\,k_2\,k_3\,k_4\right)}{K^2\left(K\,c_1\,c_2\,c_4\,k_1\,k_2\,k_4+K\,c_2\,c_3\,c_4\,k_2\,k_3\,k_4+P\,c_1\,c_2^{\,2}\,k_1\,k_2^{\,2}+P\,c_3\,c_4^{\,2}\,k_3\,k_4^{\,2}\right)} \tag{3.3}$$

$A01 =$

$$\frac{P\,A11\left(K\,c_1\,c_3\,c_4\,k_1\,k_3\,k_4+K\,c_3^{\,2}\,c_4\,k_3^{\,2}\,k_4+P\,c_1^{\,2}\,c_2\,k_1^{\,2}\,k_2+P\,c_1\,c_2\,c_3\,k_1\,k_2\,k_3\right)}{K\left(K\,c_1\,c_2\,c_4\,k_1\,k_2\,k_4+K\,c_2\,c_3\,c_4\,k_2\,k_3\,k_4+P\,c_1\,c_2^{\,2}\,k_1\,k_2^{\,2}+P\,c_3\,c_4^{\,2}\,k_3\,k_4^{\,2}\right)} \tag{3.4}$$

$A10 =$

$$\frac{P\left(K\,c_1^{\,2}\,c_2\,k_1^{\,2}\,k_2+K\,c_1\,c_2\,c_3\,k_1\,k_2\,k_3+P\,c_1\,c_3\,c_4\,k_1\,k_3\,k_4+P\,c_3^{\,2}\,c_4\,k_3^{\,2}\,k_4\right)A11}{K\left(K\,c_1\,c_2\,c_4\,k_1\,k_2\,k_4+K\,c_2\,c_3\,c_4\,k_2\,k_3\,k_4+P\,c_1\,c_2^{\,2}\,k_1\,k_2^{\,2}+P\,c_3\,c_4^{\,2}\,k_3\,k_4^{\,2}\right)} \tag{3.5}$$

$A11 =$

$$A11 \tag{3.6}$$

***Proof for invariant in the asymmetric branches***

**We know that PCon and KCon are both individually equal to zero always. Thus at a given steady state, PCon - KCon must also be equal to zero.**

**We thus introduce the term T = (KCon - PCon) = 0 and also introduce a new ratio, $\epsilon$ = [K]/[P]. Note: As discussed in the main text, the symmetric steady state is one where [K] = [P] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$$T := KCon - PCon = 0 :$$
$$K := \epsilon \cdot P :$$

**The following command (*simplify*), simplifies the expression**

$$simplify(T)$$

$$-\left(P\,(\epsilon-1)\left(c_1\,{c_2}^2\,\epsilon\,k_1\,{k_2}^2+c_4\left(c_1\,k_1+c_3\,k_3\right)\left(\left(\left(\epsilon^2-A11\left(c_1+c_3\right)\epsilon-A11\left(c_1+c_3\right)\right)k_4+A11\,c_1\,k_1\,(\epsilon+1)\right)k_2+k_4\,A11\,c_3\,k_3\,(\epsilon+1)\right)c_2+c_3\,{c_4}^2\,\epsilon\,k_3\,{k_4}^2\right)\right)\Big/\left(\epsilon\left(c_1\,{c_2}^2\,k_1\,{k_2}^2+k_4\,\epsilon\,c_4\,k_2\left(c_1\,k_1+c_3\,k_3\right)c_2+c_3\,{c_4}^2\,k_3\,{k_4}^2\right)\right)=0 \qquad \textbf{(3.1.1)}$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term (**

$$c_1\,{c_2}^2\,\epsilon\,k_1\,{k_2}^2+\left(c_1\,k_1+c_3\,k_3\right)c_4\left(\left(\left(\epsilon^2-A11\left(c_1+c_3\right)\epsilon-A11\left(c_1+c_3\right)\right)k_4+A11\,c_1\,k_1\,(\epsilon+1)\right)k_2+k_4\,A11\,c_3\,k_3\,(\epsilon+1)\right)c_2+c_3\,{c_4}^2\,\epsilon\,k_3\,{k_4}^2$$

**) in the expression needs to be equal to zero. We note that this term is a expression in the fully modified [A11], $\epsilon$ and kinetic constants. Thus solving this to isolate the fully modified substrate form we get the following. Here we use the *solve* command from Maple to solve T for [A11] as shown below.**

$$A11 := simplify(solve(T, A11)) =$$
$$-\frac{\epsilon\left(c_1\,{c_2}^2\,k_1\,{k_2}^2+k_4\,\epsilon\,c_4\,k_2\left(c_1\,k_1+c_3\,k_3\right)c_2+c_3\,{c_4}^2\,k_3\,{k_4}^2\right)}{c_2\left(c_1\,k_1+c_3\,k_3\right)c_4\left(\left(\left(-c_1-c_3\right)k_4+c_1\,k_1\right)k_2+c_3\,k_3\,k_4\right)(\epsilon+1)}$$

**Substituting it back into the expressions for [A01] and [A10], we get the following correlations**

$$simplify(A01) =$$
$$\frac{-c_3\,c_4\,\epsilon\,k_3\,k_4-c_1\,c_2\,k_1\,k_2}{c_2\,c_4\left(\left(\left(-c_1-c_3\right)k_4+c_1\,k_1\right)k_2+c_3\,k_3\,k_4\right)(\epsilon+1)} \qquad \textbf{(3.1.2)}$$

$$simplify(A10) =$$
$$\frac{-c_1\,c_2\,\epsilon\,k_1\,k_2-c_3\,c_4\,k_3\,k_4}{c_2\,c_4\left(\left(\left(-c_1-c_3\right)k_4+c_1\,k_1\right)k_2+c_3\,k_3\,k_4\right)(\epsilon+1)} \qquad \textbf{(3.1.3)}$$

**From this we can see that the sum of concentrations of the partially modified substrates ( [A01] and [A10]) are independent of ϵ and are dependent only on a few key kinetic parameters. This thus shows that at an asymmetric steady state, the sum of the concentrations of the partially modified substrates is fixed at a given value.**

$$simplify(A01 + A10)$$

$$\frac{-c_1 c_2 k_1 k_2 - c_3 c_4 k_3 k_4}{c_2 c_4 \left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right)} \tag{3.1.4}$$

## Necessary Conditions

**Since substrate concentrations are always necessarily positive, the expression in equation 3.1.1 should be positive. The numerator is only a function of kinetic parameters all of which are positive always, thus the denominator must necessarily be positive to ensure that the resulting concentration is positive. This gives us the necessary condition for symmetric state to exist as follows.**

1. $c_3 k_4 \left(k_2 - k_3\right) - c_1 k_2 \left(k_1 - k_4\right) > 0$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient for an asymmetric steady state to exist for some positive $A_{Total}$ value. or upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that in the context of this system, a feasible steady state is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that an asymmetric state defined by the invariant concentrations described above indeed is a solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an symmetric steady state, as seen above [Ap] is fixed by a few kinetic constants,**

$$A01 := simplify\left(\frac{-c_3 c_4 \epsilon k_3 k_4 - c_1 c_2 k_1 k_2}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) c_2 c_4 (\epsilon + 1)}\right) =$$

$$\frac{-c_3 c_4 \epsilon k_3 k_4 - c_1 c_2 k_1 k_2}{c_2 c_4 \left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1)}$$

$$A10 := simplify\left(\frac{-c_1 c_2 \epsilon k_1 k_2 - c_3 c_4 k_3 k_4}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) c_2 c_4 (\epsilon + 1)}\right) =$$

$$\frac{-c_1 c_2 \epsilon k_1 k_2 - c_3 c_4 k_3 k_4}{c_2 c_4 \left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1)} =$$

**The other variables in this asymmetric state are thus given by**

$$simplify(A11) = -\frac{\epsilon \left(c_1 {c_2}^2 k_1 {k_2}^2 + k_4 \epsilon c_4 k_2 \left(c_1 k_1 + c_3 k_3\right) c_2 + c_3 {c_4}^2 k_3 {k_4}^2\right)}{c_2 \left(c_1 k_1 + c_3 k_3\right) c_4 \left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1)}$$

$$simplify(A00) = \frac{-c_1 {c_2}^2 \epsilon k_1 {k_2}^2 - c_2 c_4 k_4 \left(c_1 k_1 + c_3 k_3\right) k_2 - c_3 {c_4}^2 \epsilon k_3 {k_4}^2}{c_2 \left(c_1 k_1 + c_3 k_3\right) c_4 \left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1) \epsilon}$$

$$simplify\left(A00K_1\right) = -\frac{c_1 P \left(c_2 k_1 k_2 \left(c_2 \epsilon k_2 + c_4 k_4\right) c_1 + c_3 c_4 k_3 k_4 \left(c_4 \epsilon k_4 + c_2 k_2\right)\right)}{\left(c_1 k_1 + c_3 k_3\right) \left(k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right)\right) c_2 (\epsilon + 1) c_4}$$

$$simplify\left(A00K_2\right) = -\frac{c_3 P \left(c_2 k_1 k_2 \left(c_2 \epsilon k_2 + c_4 k_4\right) c_1 + c_3 c_4 k_3 k_4 \left(c_4 \epsilon k_4 + c_2 k_2\right)\right)}{\left(c_1 k_1 + c_3 k_3\right) \left(k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right)\right) c_2 (\epsilon + 1) c_4}$$

$$simplify(A01K) = -\frac{\epsilon P \left(c_3 c_4 \epsilon k_3 k_4 + c_2 k_1 k_2 c_1\right)}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1) c_4}$$

$$simplify(A10K) = -\frac{\epsilon P \left(c_1 c_2 \epsilon k_1 k_2 + c_3 c_4 k_3 k_4\right)}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) c_2 (\epsilon + 1)}$$

$$simplify\left(A11P_1\right) = -\frac{c_3 P \epsilon \left(c_4 k_3 k_4 \left(c_2 \epsilon k_2 + c_4 k_4\right) c_3 + c_2 k_1 k_2 c_1 \left(c_4 \epsilon k_4 + c_2 k_2\right)\right)}{\left(c_1 k_1 + c_3 k_3\right) \left(k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right)\right) c_2 (\epsilon + 1) c_4}$$

$$simplify\left(A11P_2\right) = -\frac{c_1 P \epsilon \left(c_4 k_3 k_4 \left(c_2 \epsilon k_2 + c_4 k_4\right) c_3 + c_2 k_1 k_2 c_1 \left(c_4 \epsilon k_4 + c_2 k_2\right)\right)}{\left(c_1 k_1 + c_3 k_3\right) \left(k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right)\right) c_2 (\epsilon + 1) c_4}$$

$$simplify(A01P) = -\frac{P \left(c_3 c_4 \epsilon k_3 k_4 + c_2 k_1 k_2 c_1\right)}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) c_2 (\epsilon + 1)}$$

$$simplify(A10P) = -\frac{P \left(c_1 c_2 \epsilon k_1 k_2 + c_3 c_4 k_3 k_4\right)}{\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right) (\epsilon + 1) c_4}$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify(dA00) = 0$$
$$simplify(dA11) = 0$$
$$simplify(dA01) = 0$$
$$simplify(dA10) = 0$$
$$simplify\left(dA00K_1\right) = 0$$
$$simplify\left(dA00K_2\right) = 0$$
$$simplify(dA01P) = 0$$
$$simplify(dA10P) = 0$$
$$simplify(dA01K) = 0$$
$$simplify(dA10K) = 0$$
$$simplify\left(dA11P_1\right) = 0$$

$$simplify\left(dA11P_2\right) = 0$$
$$simplify(dP) = 0$$
$$simplify(dK) = 0$$

**Hence all that remains to be shown is that the variables (As described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition is satisfied**
2. **P and $\epsilon$ are positive.**

**Now by using the total conservation of phosphatase in the system (PCon = 0), we obtain an algebraic expression for [P] in terms of $P_{Total}$, kinetic constants and $\epsilon$ as shown below. The Maple command *solve* is used for this purpose as shown below.**

$$P := simplify(solve(PCon, P)) =$$
$$-\left(c_2\left(c_1 k_1 + c_3 k_3\right) c_4\left(k_2\left(k_1 - k_4\right) c_1 - c_3 k_4\left(k_2 - k_3\right)\right) P_{Total}\left(\epsilon + 1\right)\right) \Big/ \left(c_2 c_4 k_2 k_4\left(c_1 + c_3\right)\left(c_1 k_1 + c_3 k_3\right)\epsilon^2 + \left(c_2 k_2\left(\left(k_1 + k_2\right) c_2 - c_4\left(k_1 - k_4\right)\right) k_1 {c_1}^2 + \left(k_1 k_2\left(k_2 + k_3\right) {c_2}^2 - c_4\left(\left(\left(-k_1 - k_3\right) k_4 + k_1 k_3\right) k_2 + k_1 k_3 k_4\right) c_2 + {c_4}^2 k_3 k_4\left(k_1 + k_4\right)\right) c_3 c_1 + c_4\left(\left(k_2 - k_3\right) c_2 + c_4\left(k_3 + k_4\right)\right) k_4 {c_3}^2 k_3\right)\epsilon + c_2 c_4 k_2 k_4\left(c_1 + c_3\right)\left(c_1 k_1 + c_3 k_3\right)\right)$$

**Hence if $\epsilon$ is positive P is automatically positive (provided necessary conditions are satisfied) . Hence we have shown that the individual conservation equations for the enzymes (KCon and PCon) and the system of ODEs is solved by the expressions above, and represent feasible solutions provided $\epsilon$ is positive.**

**Since the concentrations are all positive there exists a unique $A_{Total}$ value for every $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

## Prediction of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitch fork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two insights.**

**1. The system is still symmetric.**

**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$$\epsilon := 1 :$$

$$A11 := simplify\left(eval\left(-\frac{\left(c_1 c_2^{\,2} k_1 k_2^{\,2} + k_4 \epsilon c_4 k_2 \left(c_1 k_1 + c_3 k_3\right) c_2 + c_3 c_4^{\,2} k_3 k_4^{\,2}\right)\epsilon}{\left(c_1 k_1 + c_3 k_3\right)\left(\left(\left(-c_1 - c_3\right) k_4 + c_1 k_1\right) k_2 + c_3 k_3 k_4\right)(\epsilon + 1)\, c_4 c_2}, \epsilon = 1\right)\right)$$

$$A11 := -\frac{\left(c_2 k_2 + k_4 c_4\right)\left(c_1 c_2 k_1 k_2 + c_3 c_4 k_3 k_4\right)}{2 c_2 \left(c_1 k_1 + c_3 k_3\right) c_4 \left(k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right)\right)} \qquad \textbf{(3.4.1)}$$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point).**

$$A_{Total} = simplify\left(solve\left(ACon, A_{Total}\right)\right)$$

$$A_{Total} = \left(2\left(c_1 c_2 k_1 k_2 + c_3 c_4 k_3 k_4\right)\left(-\frac{k_2^{\,2}\left(\left(k_1 + k_2\right) c_1 + c_3\left(k_2 + k_3\right)\right) c_1 k_1 c_2^{\,3}}{2}\right.\right. \qquad \textbf{(3.4.2)}$$

$$+ \left(\left(k_1 k_2 P_{Total}\left(k_1 - k_4\right)\left(k_1 + k_2\right) c_1^{\,3} + \left(\left(\left(\left(-2 k_1 - k_3\right) k_4 + k_1\left(k_1\right.\right.\right.\right.\right.$$

$$\left.\left.\left.\left.\left. + k_3\right)\right) k_2^{\,2} + \left(-k_1\left(k_1 + k_3\right) k_4 + 2 k_1^{\,2} k_3\right) k_2 + k_1^{\,2} k_3 k_4\right) P_{Total} c_3\right.\right.$$

$$\left.\left. + \frac{k_2 k_1\left(\left(k_1 - 4 k_4\right) k_2 - k_4 k_1\right)}{2}\right) c_1^{\,2} + \left(\left(\left(\left(-k_1 - 2 k_3\right) k_4 + k_1 k_3\right) k_2^{\,2}\right.\right.\right.$$

$$\left.\left.\left. + k_1 k_3\left(k_3 - k_4\right) k_2 + 2 k_1 k_3^{\,2} k_4\right) P_{Total} c_3 + \frac{k_2^{\,2}\left(\left(-4 k_1 - 3 k_3\right) k_4 + k_1 k_3\right)}{2}\right)\right.$$

$$\left. c_3 c_1 - \left(\left(k_2^{\,2} P_{Total} - k_3^{\,2} P_{Total}\right) c_3 + \frac{3\left(k_2 - \frac{k_3}{3}\right) k_2}{2}\right) k_4 c_3^{\,2} k_3\right) c_4$$

$$
\left. - \frac{\left(c_1 k_1 + c_3 k_3\right) k_2 \left(\left(k_1 + k_2\right) c_1 + c_3 \left(k_2 + k_3\right)\right) c_1 k_1}{2} \right) {c_2}^2
$$

$$
+ c_4 \left( \left( \left( k_1 k_2 P_{Total} \left(k_1 - k_4\right) \left(k_1 + k_4\right) {c_1}^3 + \left( 2 \left( \left( \left( -k_1 - \frac{k_3}{2} \right) {k_4}^2 - \frac{k_1 k_3 k_4}{2} \right.\right.\right.\right.\right.
$$

$$
\left.\left.\left.\left.\left. + {k_1}^2 k_3 \right) k_2 + \frac{k_1 k_3 k_4 \left(k_1 + k_4\right)}{2} \right) P_{Total} c_3 + \frac{k_1 k_2 k_4 \left(k_1 - 3 k_4\right)}{2} \right) {c_1}^2 \right.\right.
$$

$$
+ \left( \left( \left( \left( -k_1 - 2 k_3 \right) {k_4}^2 - k_3 \left(k_1 + k_3\right) k_4 + k_1 {k_3}^2 \right) k_2 + k_3 \left(k_1 + k_3\right) {k_4}^2 \right.\right.
$$

$$
\left.\left. + 2 k_1 {k_3}^2 k_4 \right) P_{Total} c_3 - \frac{3 {k_4}^2 \left( \left( k_1 + \frac{4 k_3}{3} \right) k_2 - \frac{k_1 k_3}{3} \right)}{2} \right) c_3 c_1
$$

$$
- \left( P_{Total} \left(k_3 + k_4\right) \left(k_2 - k_3\right) c_3 + \left( \frac{k_3}{2} + 2 k_4 \right) k_2 - \frac{k_4 k_3}{2} \right) k_4 {c_3}^2 k_3 \Bigg) c_4
$$

$$
+ \left. \frac{\left(c_1 k_1 + c_3 k_3\right)^2 \left( k_2 \left(k_1 - 3 k_4\right) c_1 - 3 \left( k_2 - \frac{k_3}{3} \right) k_4 c_3 \right)}{2} \right) c_2
$$

$$
\left. - \frac{{c_4}^2 \left(c_1 k_1 + c_3 k_3 + k_4 c_4\right) k_4 \left(\left(k_1 + k_4\right) c_1 + c_3 \left(k_3 + k_4\right)\right) c_3 k_3}{2} \right) \Bigg) \Bigg/
$$

$$
\left( c_2 \left(c_1 k_1 + c_3 k_3\right) c_4 \left( k_2 \left(\left(k_1 + k_2\right) c_1 + c_3 \left(k_2 + k_3\right)\right) c_1 k_1 {c_2}^2 - \left(c_1 k_1 \right.\right.\right.
$$

$$
\left.\left.\left. + c_3 k_3 \right) c_4 \left( k_2 \left(k_1 - 3 k_4\right) c_1 - 3 \left( k_2 - \frac{k_3}{3} \right) k_4 c_3 \right) c_2 + {c_4}^2 k_4 \left(\left(k_1 + k_4\right) c_1 \right.\right.\right.
$$

$$
\left.\left.\left. + c_3 \left(k_3 + k_4\right)\right) c_3 k_3 \right) \left( k_2 \left(k_1 - k_4\right) c_1 - c_3 k_4 \left(k_2 - k_3\right) \right) \right)
$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figure (Fig 3A).**

# Random distributive DSP System 2 : Separate Kinase Common Phosphatase

| Case 1 | Case 2 | Case 3 |
|---|---|---|
| Symmetry Doesn't Exist | Present and Can't Break | Symmetry Doesn't Exist |

In this Maple file we analytically show the infeasibility of case 2 symmetry breaking in Random distributive DSP network with separate kinase and common phosphatase (System 2) effecting phosphorylation and dephoshorylation respectively.We do this by first describing the model as a system of ODEs along with the associated enzyme and substrate conservations. We then impose the kinetic constraints (and constraints on total enzyme concentrations) pertaining to case 2 symmetry. By algebraically solving for the steady state of the resulting system of ODEs we obtain relations between concentrations of the substrate variables in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. case 2 symmetry breaking requires symmetry between [A01] & [A10]; By leveraging this insight, we show the infeasibility of case 2 symmetry breaking by revealing that [A01] and [A10] are always equal for any given feasible steady state. These procedures are carried out in detail below using built in Maple commands.

Note: A subscript is used to distinguish between the two different complexes formed between [P] and [A11] ([$A11P_1$ and $A11P_2$]).

**We initialize the Maple file with the restart command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := k_4 \cdot A01P + a_4 \cdot A10P + k_{ub1} \cdot A00K1 + a_{ub1} \cdot A00K2 - k_{b1} \cdot A00 \cdot K1 - a_{b1} \cdot A00 \cdot K2 :$$
$$dA01 := k_1 \cdot A00K1 + k_3 \cdot A11P_1 + k_{ub2} \cdot A01K2 + k_{ub4} \cdot A01P - k_{b2} \cdot A01 \cdot K2 - k_{b4} \cdot A01 \cdot P :$$
$$dA10 := a_1 \cdot A00K2 + a_3 \cdot A11P_2 + a_{ub2} \cdot A10K1 + a_{ub4} \cdot A10P - a_{b2} \cdot A10 \cdot K1 - a_{b4} \cdot A10 \cdot P :$$
$$dA11 := k_2 \cdot A01K2 + a_2 \cdot A10K1 + k_{ub3} \cdot A11P_1 + a_{ub3} \cdot A11P_2 - k_{b3} \cdot A11 \cdot P - a_{b3} \cdot A11 \cdot P :$$
$$dA00K1 := k_{b1} \cdot A00 \cdot K1 - (k_1 + k_{ub1}) \cdot A00K1 :$$
$$dA10K1 := a_{b2} \cdot A10 \cdot K1 - (a_2 + a_{ub2}) \cdot A10K1 :$$
$$dA00K2 := a_{b1} \cdot A00 \cdot K2 - (a_{ub1} + a_1) \cdot A00K2 :$$
$$dA01K2 := k_{b2} \cdot A01 \cdot K2 - (k_{ub2} + k_2) \cdot A01K2 :$$
$$dA11P_1 := k_{b3} \cdot A11 \cdot P - k_{ub3} \cdot (A11P_1) - k_3 \cdot (A11P_1) :$$
$$dA10P := a_{b4} \cdot A10 \cdot P - (a_{ub4} + a_4) \cdot A10P :$$
$$dA11P_2 := a_{b3} \cdot A11 \cdot P - (a_{ub3} + a_3) \cdot A11P_2 :$$
$$dA01P := k_{b4} \cdot A01 \cdot P - (k_{ub4} + k_4) \cdot A01P :$$

$$dK1 := -k_{b1}\cdot A00\cdot K1 + \left(k_1 + k_{ub1}\right)\cdot A00K1 - a_{b2}\cdot A10\cdot K1 + \left(a_2 + a_{ub2}\right)\cdot A10K1 :$$
$$dK2 := -a_{b1}\cdot A00\cdot K2 + \left(a_{ub1} + a_1\right)\cdot A00K2 - k_{b2}\cdot A01\cdot K2 + \left(k_{ub2} + k_2\right)\cdot A01K2 :$$
$$dP := -k_{b3}\cdot A11\cdot P + \left(k_{ub3} + k_3\right)\cdot A11P_1 - a_{b4}\cdot A10\cdot P + \left(a_{ub4} + a_4\right)\cdot A10P - a_{b3}\cdot A11\cdot P + \left(a_{ub3} + a_3\right)\cdot A11P_2 - k_{b4}\cdot A01\cdot P + \left(k_{ub4} + k_4\right)\cdot A01P :$$

**The above equations are also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon, K1Con and K2Con for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K1 - A01K2 - A00K2 - A10K1 - A11P_1 - A10P - A11P_2 - A01P :$$
$$K1Con := K1_{Total} - K1 - A00K1 - A10K1 :$$
$$K2Con := K2_{Total} - K2 - A00K2 - A01K2 :$$
$$PCon := P_{Total} - P - A11P_1 - A10P - A11P_2 - A01P :$$

## Case 2 Symmetry : Present and Can't Break

**Kinetic constraints for case 2 symmetry (refer main text) are imposed on the original model.**

$$a_1 := k_1 : a_{b1} := k_{b1} : a_{ub1} := k_{ub1} :$$
$$a_2 := k_2 : a_{b2} := k_{b2} : a_{ub2} := k_{ub2} :$$
$$a_3 := k_3 : a_{b3} := k_{b3} : a_{ub3} := k_{ub3} :$$
$$a_4 := k_4 : a_{b4} := k_{b4} : a_{ub4} := k_{ub4} :$$

**In addition to the kinetic constraints the total enzyme concentrations of the two kinases need to be equal for exact case 2 symmetry to be present. This is imposed as shown below.**

$$K1_{Total} := K_{Total} :$$
$$K2_{Total} := K_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $c_3$ and $c_4$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1\cdot\left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2\cdot\left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3\cdot\left(k_3 + k_{ub3}\right) :$$
$$k_{b4} := c_4\cdot\left(k_4 + k_{ub4}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K1], [K2] & [P]) and concentration of the completely unmodified substrate ([A00]). In order to**

**do this, we use the Maple command *solve,* which algebraically solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1]) is given below in detail.**

**The differential equation of [A00K1] is given by,**

$$\frac{d\,[A00K1]}{dt} = dA00K1$$

$$\frac{d\,[A00K1]}{dt} = c_1\left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1 \qquad \textbf{(1.1)}$$

**The solve command by Maple, uses this solves this equation for the given variable (in this case [A00K1]). This is performed by the following command. Here we assign [A00K1], the solution returned by the *solve* command.**

$$A00K1 := solve(dA00K1, A00K1)$$

$$A00K1 := K1\,A00\,c_1 \qquad \textbf{(1.2)}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A01K2 := solve(dA01K2, A01K2):$$
$$A00K2 := solve(dA00K2, A00K2):$$
$$A10K1 := solve(dA10K1, A10K1):$$
$$A11P_1 := solve\left(dA11P_1, A11P_1\right):$$
$$A10P := solve(dA10P, A10P):$$
$$A11P_2 := solve\left(dA11P_2, A11P_2\right):$$
$$A01P := solve(dA01P, A01P):$$

$$Sol := solve(\{dA10, dA01, dA11\}, \{A10, A01, A11\}):$$

$$A10 := eval(A10, Sol):$$
$$A01 := eval(A01, Sol):$$
$$A11 := eval(A11, Sol):$$

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = A00$$
$$A01 = \frac{K1\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$
$$A10 = \frac{K2\,A00\,c_1\,k_1}{c_4\,k_4\,P}$$

$$A11 = \frac{A00\ K1\ K2\ c_1\ c_2\ k_1\ k_2}{c_4\ k_4\ P^2\ c_3\ k_3}$$

### Proof for impossibility of symmetry breaking

**We know that K1Con and K2Con are both individually equal to zero always. Thus at a given steady state, K1Con - K2Con must also be equal to zero.**

**We thus introduce the term T = (K1Con - K2Con) = 0 and also introduce a new ratio, $\epsilon$ = [K2]/[K1]. Note: As discussed in the main text, the symmetric steady state is one where [K1] = [K2] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$T := K1Con - K2Con = 0$ :
$K2 := \epsilon \cdot K1$ :

**The following command (*simplify*), simplifies the expression algebraically**

$$simplify(T) = K1\ (\epsilon - 1)\ (A00\ c_1 + 1) = 0$$

**Thus from this we can clearly see that irrespective of kinetic parameters, [K1] is always going to be equal to [K2]. This means, from the expressions given above for [A01] and [A10], that [A01] is always equal to [A10] irrespective of kinetic parameters or total concentrations of enzymes or substrate. Thus there is no scope for any asymmetric steady state or case 2 symmetry breaking.**

## Random DSP System 3 : Separate Kinase Separate Phosphatase

| Case 1 | Case 2 | Case 3 |
|---|---|---|
| **Present and Breaks Invariant identified** | **Present and Breaks Invariant identified** | **Present and breaks** |

In this maple file we analytically show presence of case 1 and case 2 symmetry breaking in Random DSP network with separate kinase and separate phosphatase (System 3) effecting phosphorylation and dephoshorylation respectively. In each case, we do this by first describing the model as a system of ODEs along with the associated enzyme and substrate conservations. By algebraically solving for the steady state of the system of ODEs we obtain relations between substrate variables (concentrations) in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. In Random System 3, case 1 symmetry symmetry requires symmetry between [A00] & [A11]; case 2 symmetry breaking requires symmetry between [A01] & [A10].

We then proceed along two lines. One hand by leveraging this insight, we isolate correlations pertaining to asymmetric steady states involving substrate forms and free enzymes. Through an alternate approach considering the conservation of individual species and an overall flux balance around substrate forms, we ascertain a secondary correlation that is pertinent to all feasible steady states of the system. By bringing the two conditions together, we then finally ascertain the necessary conditions and features of symmetry breaking in case 1 and case 2 symmetry. These procedures are carried out in detail below using built in maple commands.

**We initialize the maple file with the *restart* command and load the relevant libraries of inbuilt maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := k_4 \cdot A01P1 + a_4 \cdot A10P2 + k_{ub1} \cdot A00K1 + a_{ub1} \cdot A00K2 - k_{b1} \cdot A00 \cdot K1 - a_{b1} \cdot A00 \cdot K2 :$$
$$dA01 := k_1 \cdot A00K1 + k_3 \cdot A11P2 + k_{ub2} \cdot A01K2 + k_{ub4} \cdot A01P1 - k_{b4} \cdot A01 \cdot P1 - k_{b2} \cdot A01 \cdot K2 :$$
$$dA10 := a_1 \cdot A00K2 + a_3 \cdot A11P1 + a_{ub2} \cdot A10K1 + a_{ub4} \cdot A10P2 - a_{b2} \cdot A10 \cdot K1 - a_{b4} \cdot A10 \cdot P2 :$$
$$dA11 := k_2 \cdot A01K2 + a_2 \cdot A10K1 + k_{ub3} \cdot A11P2 + a_{ub3} \cdot A11P1 - k_{b3} \cdot A11 \cdot P2 - a_{b3} \cdot A11 \cdot P1 :$$
$$dA00K1 := k_{b1} \cdot A00 \cdot K1 - \left(k_1 + k_{ub1}\right) \cdot A00K1 :$$
$$dA01K2 := k_{b2} \cdot A01 \cdot K2 - \left(k_2 + k_{ub2}\right) \cdot A01K2 :$$
$$dA11P2 := k_{b3} \cdot A11 \cdot P2 - \left(k_3 + k_{ub3}\right) \cdot A11P2 :$$
$$dA01P1 := k_{b4} \cdot A01 \cdot P1 - \left(k_4 + k_{ub4}\right) \cdot A01P1 :$$
$$dA00K2 := a_{b1} \cdot A00 \cdot K2 - \left(a_1 + a_{ub1}\right) \cdot A00K2 :$$
$$dA10K1 := a_{b2} \cdot A10 \cdot K1 - \left(a_2 + a_{ub2}\right) \cdot A10K1 :$$
$$dA11P1 := a_{b3} \cdot A11 \cdot P1 - \left(a_3 + a_{ub3}\right) \cdot A11P1 :$$

$$dA10P2 := a_{b4}\cdot A10\cdot P2 - \left(a_4 + a_{ub4}\right)\cdot A10P2 :$$
$$dP2 := -k_{b3}\cdot A11\cdot P2 + \left(k_3 + k_{ub3}\right)\cdot A11P2 - a_{b4}\cdot A10\cdot P2 + \left(a_4 + a_{ub4}\right)\cdot A10P2 :$$
$$dP1 := -a_{b3}\cdot A11\cdot P1 + \left(a_3 + a_{ub3}\right)\cdot A11P1 - k_{b4}\cdot A01\cdot P1 + \left(k_4 + k_{ub4}\right)\cdot A01P1 :$$
$$dK1 := -k_{b1}\cdot A00\cdot K1 + \left(k_1 + k_{ub1}\right)\cdot A00K1 - a_{b2}\cdot A10\cdot K1 + \left(a_2 + a_{ub2}\right)\cdot A10K1 :$$
$$dK2 := -a_{b1}\cdot A00\cdot K2 + \left(a_1 + a_{ub1}\right)\cdot A00K2 - k_{b2}\cdot A01\cdot K2 + \left(k_2 + k_{ub2}\right)\cdot A01K2 :$$

**The above equations are also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, K1Con, K2Con, P1Con and P2Con for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K1 - A01K2 - A11P2 - A01P1 - A00K2 - A10K1 - A11P1 - A10P2 :$$
$$K1Con := K1_{Total} - A00K1 - A10K1 - K1 :$$
$$K2Con := K2_{Total} - A01K2 - A00K2 - K2 :$$
$$P1Con := P1_{Total} - A11P1 - A01P1 - P1 :$$
$$P2Con := P2_{Total} - A10P2 - A11P2 - P2 :$$

**Until now we have modelled the System 3 Random DSP network with separate kinase effecting phosphorylation and a separate phosphatase for dephosphorylation - without any impositions on kinetics or concentrations for symmetry. In the following segments, we specifically do this for each class of symmetry. The codes for each symmetry are modular and in order to run a particular symmetry, please run the code until this point and then run only the code for the specific class of symmetry.**

**Note: Do not run the whole script at the same time, as this will impose all symmetries at the same time and give incorrect results.**

## Case 1 Symmetry : Present and Breaks

**Kinetic constraints for case 1 symmetry (refer main text) are imposed on the original model.**

$$k_3 := k_1 : k_{b3} := k_{b1} : k_{ub3} := k_{ub1} :$$
$$k_4 := k_2 : k_{b4} := k_{b2} : k_{ub4} := k_{ub2} :$$

$$a_3 := a_1 : a_{b3} := a_{b1} : a_{ub3} := a_{ub1} :$$
$$a_4 := a_2 : a_{b4} := a_{b2} : a_{ub4} := a_{ub2} :$$

**In addition to the kinetic constraints the total enzyme concentrations of corresponding kinases and phosphatases need to be equal for exact case 1 symmetry in Random DSP System 3. This is imposed as shown below.**

$$K1_{Total} := P2_{Total} :$$

$K2_{Total} := P1_{Total}$ :

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $d_1$ and $d_2$ in place of the binding constants so as to make further analytical expressions more accessible.**

$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right)$ :
$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right)$ :

$a_{b1} := d_1 \cdot \left(a_1 + a_{ub1}\right)$ :
$a_{b2} := d_2 \cdot \left(a_2 + a_{ub2}\right)$ :

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K1], [K2], [P1] & [P2]) and concentration of the partially modified substrate [A01]. In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1]) is given below in detail.**

**The differential equation of [A00K1] is given by,**

$$\frac{d\,[A00K1]}{dt} = dA00K1 \qquad \frac{d\,[A00K1]}{dt} = c_1\left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K1]). We in turn store this value (the solution returned by the *solve* command) in [A00K1]. This is performed by the following command.**

$A00K1 := solve(dA00K1, A00K1)$

$$A00K1 := K1\,A00\,c_1 \tag{1.1}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$A01K2 := solve(dA01K2, A01K2)$ :
$A11P2 := solve(dA11P2, A11P2)$ :
$A01P1 := solve(dA01P1, A01P1)$ :
$A00K2 := solve(dA00K2, A00K2)$ :
$A10K1 := solve(dA10K1, A10K1)$ :
$A11P1 := solve(dA11P1, A11P1)$ :
$A10P2 := solve(dA10P2, A10P2)$ :

$Sol := solve(\{dA00, dA11, dA10\}, \{A00, A11, A10\})$ :

$A00 := eval(A00, Sol)$ :

$A01 := eval(A01, Sol)$ :
$A10 := eval(A10, Sol)$ :
$A11 := eval(A11, Sol)$ :

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = \frac{A01\ P1\ c_2\ k_2}{c_1\ k_1\ K1}$$
$$A01 = A01$$
$$A10 = \frac{A01\ K2\ P1\ a_1\ c_2\ d_1\ k_2}{P2\ c_1\ k_1\ K1\ a_2\ d_2}$$
$$A11 = \frac{K2\ A01\ c_2\ k_2}{P2\ c_1\ k_1}$$

### Proof for invariant in asymmetric branches

**We know that** $\left(P1_{Total} \cdot P2_{Total} - K1_{Total} \cdot K2_{Total} = 0\right)$ **in this system under case 1 symmetry (Since** $K1_{Total} = P2_{Total}$ **and** $K2_{Total} = P1_{Total}$**). Thus at a given steady state,** $\left(P1_{Total} \cdot P2_{Total} - K1_{Total} \cdot K2_{Total}\right)$ **must be equal to 0. We introduce a term T =** $\left(P1_{Total} \cdot P2_{Total} - K1_{Total} \cdot K2_{Total}\right)$ **as shown below. Here we write** $P1_{Total}$ **as** $P1_{Total}$ **- P1Con. Note: This comes from the expression used for P1Con earlier (P1Con =** $P1_{Total}$ **- P1 - A11P1 - A01P1) and the fact that PCon = 0. This way we have represented the independent enzyme concentrations as a sum of the free enzyme and the complexes in which it is sequestered.**

**Similarly for the other total enzyme concentrations as well.**

$$T := \left(P1_{Total} - P1Con\right) \cdot \left(P2_{Total} - P2Con\right) - \left(K2_{Total} - K2Con\right) \cdot \left(K1_{Total} - K1Con\right) = 0 :$$

**At this stage we evaluate the ratio of [A00] and [A11]. Note that with case 1 symmetry breaking, we are looking for asymmetric states such that [A00] not equal to [A11].**

$$simplify\left(\frac{A00}{A11}\right) = \frac{P1\ P2}{K1\ K2}$$

**We denote this ratio as** $\epsilon$**. i.e.** $\epsilon$ **= (P2*P1)/(K2*K1).**

$$P2 := \frac{\epsilon \cdot K1 \cdot K2}{P1} :$$
$$simplify\left(\frac{A00}{A11}\right) = \epsilon$$

**After introducing this ratio and algebraically simplifying T using the Maple command *simplify*, we get the following equation.**

$$simplify(T) = \frac{1}{KI\,c_1\,\epsilon\,k_1^{\,2}\,a_2}\left((\epsilon-1)\left(KI^2\,\epsilon\,a_2\,c_1\left(A01\,c_2+1\right)k_1^{\,2}+k_2\,PI^2\,A01\,a_1\,c_2\,d_1\left(A01\,c_2+1\right)k_1 - A01^2\,PI^2\,a_2\,c_2^{\,2}\,d_1\,k_2^{\,2}\right)K2\right)=0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term**
$\left(KI^2\,a_2\,\epsilon\,c_1\left(A01\,c_2+1\right)k_1^{\,2}+a_1\,PI^2\,A01\,c_2\,d_1\,k_2\left(A01\,c_2+1\right)k_1 - A01^2\,PI^2\,c_2^{\,2}\,d_1\,k_2^{\,2}\,a_2\right)$ **in the numerator of the expression needs to be zero. This term is an expression in the partial substrate form [A01], $\epsilon$, the free enzymes and kinetic constants.**

**We show here that this term is nothing but the numerator of the following expression where [A01] written as a function of [A10].**

$$r := \left(\left(\frac{A01}{k_1}\cdot\frac{A10}{a_1} - \left(\frac{A01}{k_2}+\frac{1}{c_2\cdot k_2}\right)\cdot\left(\frac{A10}{a_2}+\frac{1}{d_2\cdot a_2}\right)\right)\right):$$

**This is verified below. The *numer* command in Maple isolates only the numerator of an expression. In this context we isolate the numerator of 'r' and add the term denoting asymmetry from earlier. We can see that all expressions in the term cancel out revealing that they are identical.**

$$simplify\left(numer(r)+\left(KI^2\,a_2\,\epsilon\,c_1\left(A01\,c_2+1\right)k_1^{\,2}+a_1\,PI^2\,A01\,c_2\,d_1\,k_2\left(A01\,c_2+1\right)k_1 - A01^2\,PI^2\,c_2^{\,2}\,d_1\,k_2^{\,2}\,a_2\right)\right) = 0$$

**Thus, here we have a clear representation of the asymmetric steady state in terms of a correlation between only the concentrations of the partially modified enzymes and key kinetic constants (i.e. r = 0). r in turn can be rewritten for simplicity as follows.**

$$restart:$$
$$r := \lambda_{A01}\cdot\lambda_{A10} - 1:$$

**Where the $\lambda_{A01}$ and $\lambda_{A10}$ expressions are given as follows.**

$$\lambda_{A01} := simplify\left(\frac{\frac{A01}{k_1}}{\left(\frac{A01}{k_2}+\frac{1}{c_2\cdot k_2}\right)}\right) = \frac{A01\,c_2\,k_2}{k_1\left(A01\,c_2+1\right)}$$

$$\lambda_{A10} := simplify\left(\frac{\frac{A10}{a_1}}{\left(\frac{A10}{a_2}+\frac{1}{d_2\cdot a_2}\right)}\right) = \frac{A10\,d_2\,a_2}{a_1\left(A10\,d_2+1\right)}$$

**Now we use an auxiliary analysis to ascertain other correlations between [A01] and [A10] representing all steady states of the system. We refresh Maple here to rewrite some new expressions.**

*restart* :

**We begin by evaluating the following ratio of enzymes** $\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)}$. **We know from the individual conservation equations that they correspond to the sum of the free enzyme and the complexes in which they are sequestered. Thus,**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)} = \frac{(A10P2+A11P2+P2)+(A00K1+A10K1+K1)}{(A11P1+A01P1+P1)+(A01K2+A00K2+K2)} :$$

**This can be further simplified as follows, (by rewriting the concentration of the complexes at steady state and simplifying the resulting expression)**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)} = \frac{(A11P2+A00K1)+\left(1+d_2\cdot A10\right)\cdot(K1+P2)}{(A00K2+A11P1)+\left(1+c_2\cdot A01\right)\cdot(K2+P1)}$$

**Now from the dynamical representation of the network as a system of ODEs earlier (particularly equations dA01 and dA10) we know that at steady state,**

$$A11P2+A00K1 = \frac{k_2}{k_1}\cdot(A01K2+A01P1) :$$

$$A11P1+A00K2 = \frac{a_2}{a_1}\cdot(A10P2+A10K1) :$$

**Substituiting this in the earlier expression we get the following,**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)}=\frac{\frac{k_2}{k_1}\cdot(A01K2+A01P1)+\left(1+d_2\cdot A10\right)\cdot(K1+P2)}{\frac{a_2}{a_1}\cdot(A10P2+A10K1)+\left(1+c_2\cdot A01\right)\cdot(K2+P1)}:$$

**Which simplifies to,**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)}=\frac{\frac{k_2}{k_1}\cdot c_2\cdot A01\cdot(K2+P1)+\left(1+d_2\cdot A10\right)\cdot(K1+P2)}{\frac{a_2}{a_1}\cdot d_2\cdot A10\cdot(P2+K1)+\left(1+c_2\cdot A01\right)\cdot(K2+P1)}:$$

**Now we introduce a new ratio,** $\alpha=\frac{\left(1+c_2\cdot A01\right)\cdot(K2+P1)}{\left(1+d_2\cdot A10\right)\cdot(K1+P2)}$ **Substituting this ratio in the above expression we get the following simplified form,**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)}=\frac{\frac{k_2}{k_1}\cdot\frac{c_2\cdot A01\cdot\text{alpha}}{\left(1+c_2\cdot A01\right)}+1}{\frac{a_2}{a_1}\cdot\frac{d_2\cdot A10}{\left(1+d_2\cdot A10\right)}+\alpha}:$$

**Reintroducting the nomenclature of** $\lambda_{A01}$ **and** $\lambda_{A10}$ **the above equation simplifies as,**

$$\frac{\left(P2_{Total}+K1_{Total}\right)}{\left(P1_{Total}+K2_{Total}\right)}=\frac{\lambda_{A01}\cdot\alpha+1}{\lambda_{A10}+\alpha}:$$

**Under case 1 symmetry,** $K2_{Total}=P1_{Total}$**, and** $K1_{Total}=P2_{Total}$**. Here we bring in the asymmetry correlation ascertained earlier, i.e. r = 1 or** $\lambda_{A01}=\frac{1}{\lambda_{A10}}$ **. Using these two insights, the above equation simplifies to**

$$\frac{P2_{Total}}{P1_{Total}}=\lambda_{A01}\cdot\frac{\left(\lambda_{A01}\cdot\alpha+1\right)}{\left(1+\alpha\cdot\lambda_{A01}\right)}$$

**Since the expression on the right is always positive (** $\alpha$ **and** $\lambda_{A01}$ **are all positive sums of kinetic constants and species concentrations), we can cancel it in the numerator and the denominator, without making any assumptions about the steady state.**

$$\lambda_{A01} = \frac{P2_{Total}}{P1_{Total}}$$

**Thus,**

$$\lambda_{A10} = \frac{P1_{Total}}{P2_{Total}}$$

**Solving these equations for their respective substrate concentrations we find that in an asymmetric state, the concentrations of [A01] and [A10], the partially modified substrates are fixed and is given by few key kinetic constants.**

$$A01 := -\frac{P2_{Total}\, k_1}{c_2\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right)} = -\frac{P2_{Total}\, k_1}{c_2\left(k_1\, P2_{Total} - k_2\, P1_{Total}\right)}$$

$$A10 := -\frac{P1_{Total}\, a_1}{d_2\left(P1_{Total}\, a_1 - a_2\, P2_{Total}\right)} = -\frac{P1_{Total}\, a_1}{d_2\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)}$$

## Necessary Conditions

**Since substrate concentration are always necessarily positive, the expression for the concentrations of [A00] and [A11] should be positive. The numerator is only a function of kinetic parameters and total enzyme concentrations which are always positive, thus the denominator must necessarily be positive to ensure that the resulting concentration is positive. This gives us the necessary condition for the asymmetric state to exist as follows.**

1. $k_2 \cdot P1_{Total} > k_1 \cdot P2_{Total}$
2. $a_2 \cdot P2_{Total} > a_1 \cdot P1_{Total}$

## Sufficiency of necessary conditions

**In this section we show that the necessary conditions generated above are sufficient to ensure symmetry breaking at some finite $A_{Total}$.**

**As shown above, should an asymmetric branch exist the associated invariants (of concentrations) need to be true irrespective of $A_{Total}$ values. To show that the asymmetric state is a feasible steady state for the system, we need to show that the concentrations of all variables involved (substrates, complexes and free enzymes) are all positive. Simultaneoulsy we need ensure that the system satisfies the conservation conditions associated with the concentrations of the substrate and the respective enzymes.**

**We show this in the following manner.**

**1. We evaluate the steady state of the system to obtain expressions of concentration of all substrates and complexes in terms of a few key concentrations ([A01], [K1], [P1], [K2], [P2] as done earlier).**
**Note: This is done by running the code until before the section 'Proof for invariant in asymmetric branches'**

**2. We then substitute one of the invariants [A01] concentration into the system.**

$$A01 := -\frac{P2_{Total}\, k_1}{c_2\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right)} :$$

**3. Further we solve the conservation expressions K1Con and P1Con, for expressions of the free enzymes [K1] and [P1] in terms of the concentrations of other free enzymes ([K2] and [P2]) and the invariant substituted earlier for [A01] concentration.**

$$P1 := solve(P1Con, P1) :$$
$$K1 := solve(K1Con, K1) :$$

**We can thus ascertain concentrations of the following substrates as a function of free enzymes [K2], [P2] and constants. Here we introduce a new ratio** $\alpha = \frac{K2}{P2}$

$$K2 := \alpha \cdot P2 :$$
$$simplify(A10) = -\frac{P1_{Total}\, a_1}{\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right) d_2}$$
$$simplify(K1) = -\frac{P2_{Total}\, \alpha\, d_1\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)}{\left(\alpha\, d_1\, P2_{Total} + c_1\, P1_{Total}\right) a_2}$$
$$simplify(P1) = -\frac{P1_{Total}\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right) c_1}{k_2\left(\alpha\, d_1\, P2_{Total} + c_1\, P1_{Total}\right)}$$
$$simplify(A00) = -\frac{P1_{Total}\, a_2}{\alpha\, d_1\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)}$$
$$simplify(A11) = -\frac{\alpha\, P2_{Total}\, k_2}{\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right) c_1}$$

**We can see that automatically we get the second invariant for the concentration of [A10] to be true. From the expressions of the variables we note that when [K2] and [P2] are both positive the variable concentrations are positive.**

**4. Now we solve K2Con (the conservation condition associated with the enzyme [K2]) for the concentration of free enzyme [P2] and get the following correlation.**

$$P2 := simplify(solve(P2Con, P2)) =$$

$$P2 := -\frac{\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right)}{k_2\left(a_1\,\alpha - a_2\right) P1_{Total} + a_2\, P2_{Total}\left(-\alpha\, k_2 + k_1\right)} \qquad \textbf{(1.3.1)}$$

**Note that the denominator can be simplified as**
$\alpha\, k_2\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right) + a_2\left(k_1\, P2_{Total} - k_2\, P1_{Total}\right)$

**So again, if $\alpha$ is positive, [P2] is positive implying [K2] is positive.**

**In an asymmetric condition we have no restriction on the value that this $\alpha$ can take. Note that in a symmetric state, by mandating [A01] = [A10] we fix the value of $\alpha$ in terms of kinetic constants and total enzyme concentrations.**

$simplify(K1Con) = 0$
$simplify(K2Con) = 0$
$simplify(P1Con) = 0$
$simplify(P2Con) = 0$

$simplify(dA00) = 0$
$simplify(dA11) = 0$
$simplify(dA01) = 0$
$simplify(dA10) = 0$
$simplify(dA00K1) = 0$
$simplify(dA00K2) = 0$
$simplify(dA01P1) = 0$
$simplify(dA10P2) = 0$
$simplify(dA01K2) = 0$
$simplify(dA10K1) = 0$
$simplify(dA11P1) = 0$
$simplify(dA11P2) = 0$
$simplify(dK2) = 0$
$simplify(dK1) = 0$
$simplify(dP1) = 0$
$simplify(dP2) = 0$

**We can thus ensure that for any positive value of $\alpha$, all concentrations are positive. This thus implies that there exists a finite positive ATotal value where this asymmetric state is realized.**

**Hence we have shown that the necessary conditions are sufficient for symmetry breaking at some finite $A_{Total}$ value.**

### Position of symmetry breaking (Pitchfork bifurcation)

**At the position of symmetry breaking, we know two insights.**
**1. The system is still symmetric, hence [A00] = [A11], [K1] = [P2], [P1] = [K2]**

**2. The invariants descibing the asymmetric steady state are also true.**

**Using these two information, we can simplify the original system considerably as follows. Note:**

$K1 := P2$ :
$P1 := K2$ :

$$A00 = \frac{K2\, A01\, c_2\, k_2}{P2\, c_1\, k_1}$$
$$A01 = A01$$
$$A10 = \frac{A01\, K2^2\, a_1\, c_2\, d_1\, k_2}{P2^2\, c_1\, k_1\, a_2\, d_2}$$
$$A11 = \frac{K2\, A01\, c_2\, k_2}{P2\, c_1\, k_1}$$

**But we also know that [A01] and [A10] are invariants given by the following expressions.**

$$A01 := -\frac{P2_{Total}\, k_1}{c_2\left(P2_{Total}\, k_1 - k_2\, P1_{Total}\right)} = -\frac{P2_{Total}\, k_1}{c_2\left(k_1\, P2_{Total} - k_2\, P1_{Total}\right)}$$
$$A10 := -\frac{P1_{Total}\, a_1}{d_2\left(P1_{Total}\, a_1 - a_2\, P2_{Total}\right)} - \frac{P1_{Total}\, a_1}{d_2\left(a_1\, P1_{Total} - a_2\, P2_{Total}\right)}$$

**Using this information we can find the value of [K2] and [P2]. Substituiting this into the expressions for conservation of the individual enzymes, we get the concentrations of the respective free enzymes.**

$$A_{Total} = simplify\left(solve\left(ACon, A_{Total}\right)\right)$$
$$A_{Total} = \left(-2\, a_2\left(\left(P2\left(k_2 + k_1\right) c_1 + k_2\left(K2\, d_1 + 1\right)\right) K2\, c_2 + \frac{P2\, c_1\, k_1}{2}\right) d_2\, P2_{Total}{}^2 \right. \quad \textbf{(1.4.1)}$$
$$+ 2\, P1_{Total}\left(\left(\left(P2\left(k_1\, P2 + K2\left(k_2 + k_1\right)\right)\right) c_1 + K2\, k_2\left(K2\, d_1 + 1\right)\right) d_2\right.$$
$$\left. + \frac{P2\, c_1\, k_1}{2}\right) c_2 + \frac{P2\, c_1\, d_2\, k_1}{2}\Bigg) a_1\, P2_{Total} - 2\, P1_{Total}{}^2\, P2\, c_1\left(P2\, d_2\right.$$
$$\left. + \frac{1}{2}\right) a_1\, k_2\, c_2\Bigg) \Big/ \left(\left(-P2_{Total}\, k_1 + k_2\, P1_{Total}\right) c_1\, P2\, d_2\left(P1_{Total}\, a_1\right.\right.$$
$$\left.\left. - a_2\, P2_{Total}\right) c_2\right)$$

**An example of this is done in the read me file using the parameters used in generating the**

**figures. A cross verification of the analytical work and bifurcation analysis is done in this manner.**

## Case 2 Symmetry : Present and Breaks

**Kinetic constraints for case 2 symmetry to be present (refer main text) are imposed on the original model.**

$$a_1 := k_1 : a_{b1} := k_{b1} : a_{ub1} := k_{ub1} :$$
$$a_2 := k_2 : a_{b2} := k_{b2} : a_{ub2} := k_{ub2} :$$
$$a_3 := k_3 : a_{b3} := k_{b3} : a_{ub3} := k_{ub3} :$$
$$a_4 := k_4 : a_{b4} := k_{b4} : a_{ub4} := k_{ub4} :$$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase need to be equal for exact case 2 symmetry to be present. This is imposed as shown below.**

$$K1_{Total} := K_{Total} :$$
$$K2_{Total} := K_{Total} :$$
$$P1_{Total} := P_{Total} :$$
$$P2_{Total} := P_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $c_3$ and $c_4$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) :$$
$$k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K1], [K2], [P1] & [P2]) and concentration of the completely unmodified substrate [A00]. In order to do this, we use the Maple command *solve,* which algebraically solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1]) is given below in detail.**

**The differential equation of [A00K1] is given by,**

$$\frac{d\,[A00K1]}{dt} = dA00K1 \qquad \frac{d\,[A00K1]}{dt} = c_1 \left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1$$

**The solve command by Maple, uses this solves this equation for the given variable (in this case [A00K1]). This is performed by the following command. Here we assign [A00K1], the solution**

**returned by the solve command.**

$$A00K1 := solve(dA00K1, A00K1)$$

$$A00K1 := K1\ A00\ c_1 \tag{2.1}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously *solve* expressions for the substrate forms using the solve command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A01K2 := solve(dA01K2, A01K2):$$
$$A11P2 := solve(dA11P2, A11P2):$$
$$A01P1 := solve(dA01P1, A01P1):$$
$$A00K2 := solve(dA00K2, A00K2):$$
$$A10K1 := solve(dA10K1, A10K1):$$
$$A11P1 := solve(dA11P1, A11P1):$$
$$A10P2 := solve(dA10P2, A10P2):$$

$$Sol := solve(\{dA01, dA10, dA11\}, \{A01, A10, A11\}):$$

$$A00 := eval(A00, Sol):$$
$$A01 := eval(A01, Sol):$$
$$A10 := eval(A10, Sol):$$
$$A11 := eval(A11, Sol):$$

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = A00$$
$$A01 = \frac{A00\ K1\ c_1\ k_1}{c_4\ k_4\ P1}$$
$$A10 = \frac{K2\ A00\ c_1\ k_1}{P2\ c_4\ k_4}$$
$$A11 = \frac{A00\ K1\ K2\ c_1\ c_2\ k_1\ k_2}{P2\ c_4\ k_4\ P1\ c_3\ k_3}$$

### Proof for invariant in asymmetric branches

**We know that $P1_{Total}$*$K2_{Total}$ = $K1_{Total}$*$P2_{Total}$ in this system under case 1 symmetry (Since $K1_{Total}$ = $K2_{Total}$ and $P2_{Total}$ = $P1_{Total}$). Thus at a given steady state, ($P1_{Total}$* $K2_{Total}$ = $K1_{Total}$*$P2_{Total}$) must be equal to 0. We introduce a term T = ($P1_{Total}$*$K2_{Total}$ = $K1_{Total}$*$P2_{Totall}$) as shown below. Here we write $P1_{Total}$ as P1Con + $P1_{Total}$. Note: This comes from the expression used for P1Con earlier and the fact that PCon = 0. This way we have represented the independent enzyme concentrations as a sum of the free enzyme and the complexes in which it is sequestered.**

**Similarly for the other total enzyme concentrations as well.**

$$T := \left(P1Con - P1_{Total}\right)\cdot\left(K2Con - K2_{Total}\right) - \left(P2Con - P2_{Total}\right)\cdot\left(K1Con - K1_{Total}\right) = 0 :$$

**This way we have represented the independent enzyme concentrations as a sum of the free enzyme and the complexes in which it is sequestered.**

**At this stage we evaluate the ratio of [A01] and [A10]. Note that with case 2 symmetry breaking, we are looking for asymmetric states such that [A01] not equal to [A10].**

$$simplify\left(\frac{A01}{A10}\right) = \frac{K1\ P2}{P1\ K2}$$

**We denote this ratio as $\epsilon$. i.e. $\epsilon$ = (P2*K1)/(K2*P1).**

$$P2 := \frac{\epsilon\cdot P1\cdot K2}{K1} :$$
$$simplify\left(\frac{A01}{A10}\right) = \epsilon =$$

**After introducing this ratio and algebraically simplifying T using the Maple command *simplify*, we get the following equation.**

$$simplify(T)$$
$$\frac{1}{P1\ c_4\ k_4^{\,2}\ k_3\ \epsilon}\left((\epsilon - 1)\left(-k_3\ \epsilon\ P1^2\ c_4\left(A00\ c_1 + 1\right)k_4^{\,2} - K1^2\ A00\ c_1\ c_2\ k_1\ k_2\left(A00\ c_1 + 1\right)k_4 + A00^2\ K1^2\ c_1^{\,2}\ c_2\ k_1^{\,2}\ k_3\right)K2\right) = 0 \quad \textbf{(2.1.1)}$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term (**
$-k_3\ \epsilon\ P1^2\ c_4\left(A00\ c_1 + 1\right)k_4^{\,2} - K1^2\ A00\ c_1\ c_2\ k_1\ k_2\left(A00\ c_1 + 1\right)k_4 + A00^2\ K1^2\ c_1^{\,2}\ c_2\ k_1^{\,2}\ k_3$**) in the numerator of the expression needs to be zero. This term is an expression in the partial substrate form [A00], $\epsilon$, the free enzymes and kinetic constants.**

**We show here that this term is nothing but the numerator of the following expression multiplied by the ratio ($\epsilon\ P1^2\ c_4\ k_4$) where [A01] written as a function of [A10].**

$$r := \left(\left(\frac{k_3\cdot c_3\cdot A11}{k_2\cdot\left(1 + c_3\cdot A11\right)}\right)\cdot\left(\frac{k_1\cdot c_1\cdot A00}{k_4\cdot\left(1 + c_1\cdot A00\right)}\right) - 1\right) :$$

**This is verified below. The *numer* command in Maple isolates only the numerator of an expression. In this context we isolate the numerator of 'r', divide it by the factor ($\epsilon\ P1^2\ c_4\ k_4$) and add the term denoting asymmetry from earlier. We can see that all expressions in the**

**term cancel out revealing that they are identical.**

$$factor\left(-k_3 \in PI^2 c_4 \left(A00\, c_1 + 1\right) k_4^{\,2} - KI^2\, A00\, c_1 c_2 k_1 k_2 \left(A00\, c_1 + 1\right) k_4 + A00^2\, KI^2 c_1^{\,2} c_2 k_1^{\,2} k_3 - \frac{(numer(r))}{\in PI^2 c_4 k_4}\right)$$

**Thus, here we have a clear representation of the asymmetric steady state in terms of a correlation between only the concentrations of the partially modified enzymes and key kinetic constants (i.e. r = 0). r in turn can be rewritten for simplicity as follows.**

```
restart :
r := λ_A00·λ_A11 - 1 :
```

**Where the $\lambda_{A00}$ and $\lambda_{A11}$ expressions are given as follows.**

$$\lambda_{A11} := \frac{k_3 \cdot c_3 \cdot A11}{k_2 \cdot \left(1 + c_3 \cdot A11\right)} :$$

$$\lambda_{A00} := \frac{k_1 \cdot c_1 \cdot A00}{k_4 \cdot \left(1 + c_1 \cdot A00\right)} :$$

**Now we use an auxiliary analysis to ascertain other correlations between [A00] and [A11] that is valid for all steady states of the system (not just asymmetric ones). We refresh Maple here to rewrite some new expressions.**

```
restart :
```

**We begin by evaluating the following ratio of enzymes $\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)}$. We know from the individual conservation equations that they correspond to the sum of the free enzyme and the complexes in which they are sequestered. Thus,**

$$\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)} = \frac{(A11P1 + A01P1 + P1) + (A10P2 + A11P2 + P2)}{(A00K1 + A10K1 + K1) + (A01K2 + A00K2 + K2)} :$$

**Now from the dynamical representation of the network as a system of ODEs earlier (particularly equations dA01 and dA10) we know that at steady state,**

$$A01P1 + A10P2 = \frac{k_1}{k_4} \cdot (A00K1 + A00K2) :$$

$$A01K2 + A10K1 = \frac{k_3}{k_2} \cdot (A11P1 + A11P2) :$$

Thus,

$$\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)} = \frac{\frac{k_1}{k_4} \cdot (A00K1 + A00K2) + P1 + P2 + A11P1 + A11P2}{\frac{k_3}{k_2} \cdot (A11P1 + A11P2) + A00K1 + K1 + K2 + A00K2} :$$

**This can be further simplified as follows, (by rewriting the concentration of the complexes at steady state and simplifying the resulting expression)**

$$\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)} = \frac{\frac{k_1}{k_4} \cdot c_1 \cdot A00 \cdot (K1 + K2) + \left(1 + c_3 \cdot A11\right) \cdot (P1 + P2)}{\frac{k_3}{k_2} \cdot c_3 \cdot A11 \cdot (P1 + P2) + \left(1 + c_1 \cdot A00\right) \cdot (K1 + K2)} :$$

**Now we introduce a new ratio,** $\alpha = \frac{\left(1 + c_1 \cdot A00\right) \cdot (K1 + K2)}{\left(1 + c_3 \cdot A11\right) \cdot (P1 + P2)}$ **Substituting this ratio in the above expression we get the following simplified form,**

$$\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)} = \frac{\frac{k_1}{k_4} \cdot \frac{c_1 \cdot A00 \cdot \alpha}{\left(1 + c_1 \cdot A00\right)} + 1}{\frac{k_3}{k_2} \cdot \frac{c_3 \cdot A11}{\left(1 + c_3 \cdot A11\right)} + \alpha} :$$

**Reintroducting the nomenclature of** $\lambda_{A00}$ **and** $\lambda_{A11}$ **the above equation simplifies as,**

$$\frac{\left(P1_{Total} + P2_{Total}\right)}{\left(K1_{Total} + K2_{Total}\right)} = \frac{\lambda_{A00} \cdot \alpha + 1}{\lambda_{A11} + \alpha} :$$

**Under case 2 symmetry,** $K2_{Total} = K1_{Total}$**, and** $P1_{Total} = P2_{Total}$**. Here we bring in the asymmetry correlation ascertained earlier, i.e. r = 1 or** $\lambda_{A00} = \frac{1}{\lambda_{A11}}$ **. Using these two insights, the above equation simplifies to**

$$\frac{\left(P_{Total}\right)}{\left(K_{Total}\right)} = \frac{\left(\lambda_{A00} \cdot \alpha \cdot D + 1\right)}{\left(1 + \alpha \cdot D \cdot \lambda_{A00}\right)} \cdot \lambda_{A00} :$$

**Since the expression on the right is always positive ( $\alpha$ and $\lambda_{A00}$ are all positive sums of kinetic constants and species concentrations), we can cancel it in the numerator and the denominator, without making any assumptions about the steady state.**

$$\lambda_{A00} := \frac{P_{Total}}{K_{Total}} :$$

**Thus**

$$\lambda_{A11} := \frac{K_{Total}}{P_{Total}} :$$

**Solving these equations for their respective substrate concentrations we find that in an asymmetric state, the concentrations of [A00] and [A11], the partially modified substrates are fixed and is given by few key kinetic constants.**

$$A00 = solve\left(\frac{k_1 \cdot c_1 \cdot A00}{k_4 \cdot \left(1 + c_1 \cdot A00\right)} = \frac{P_{Total}}{K_{Total}}, A00\right)$$

$$A00 = \frac{P_{Total}\, k_4}{c_1 \left(k_1\, K_{Total} - P_{Total}\, k_4\right)} \tag{2.1.2}$$

$$A11 = solve\left(\frac{k_3 \cdot c_3 \cdot A11}{k_2 \cdot \left(1 + c_3 \cdot A11\right)} = \frac{K_{Total}}{P_{Total}}, A11\right)$$

$$A11 = -\frac{K_{Total}\, k_2}{c_3 \left(K_{Total}\, k_2 - k_3\, P_{Total}\right)} \tag{2.1.3}$$

## Necessary conditions

**Since substrate concentration are always necessarily positive, the expression for the concentrations of [A00] and [A11] should be positive. The numerator is only a function of kinetic parameters and total enzyme concentrations which are always positive, thus the denominator must necessarily be positive to ensure that the resulting concentration is positive. This gives us the necessary condition for the asymmetric state to exist as follows.**

1. $k_1 \cdot K_{Total} > k_4 \cdot P_{Total}$
2. $k_3 \cdot P_{Total} > k_2 \cdot K_{Total}$

## Sufficiency of necessary conditions

**In this section we show that the necessary conditions generated above are sufficient to ensure symmetry breaking at some finite $A_{Total}$.**

**As shown above, should an asymmetric branch exist the associated invariants (of concentrations) need to be true irrespective of $A_{Total}$ values. To show that the asymmetric state is a feasible steady state for the system, we need to show that the concentrations of all variables involved (substrates, complexes and free enzymes) are all positive. Simultaneoulsy we need ensure that the system satisfies the conservation conditions associated with the concentrations of the substrate and the respective enzymes.**

**We show this in the following manner.**

**1. We evaluate the steady state of the system to obtain expressions of concentration of all substrates and complexes in terms of a few key concentrations ([A00], [K1], [P1], [K2], [P2] as done earlier).**
**Note: This is done by running the code until before the section 'Proof for invariant in asymmetric branches'**

**2. We then substitute one of the invariants [A00] concentration into the system.**

$$A00 := \frac{P_{Total}\, k_4}{c_1 \left(k_1 K_{Total} - P_{Total}\, k_4\right)} :$$

**3. Further we solve the conservation expressions K1Con and P1Con, for expressions of the free enzymes [K1] and [P1] in terms of the concentrations of other free enzymes ([K2] and [P2]) and the invariant substituted earlier for [A00] concentration.**

$$K1 := simplify(solve(K1Con, K1)) :$$
$$P1 := simplify(solve(P1Con, P1)) :$$

**We can thus ascertain concentrations of the following substrates as a function of free enzymes [K2], [P2] and constants. Here wee introduce a new ratio $\alpha = \frac{K2}{P2}$**

$$K2 := \alpha \cdot P2 :$$

$$simplify(A11) = -\frac{K_{Total}\, k_2}{\left(k_2 K_{Total} - k_3 P_{Total}\right) c_3}$$

$$simplify(K1) = \frac{K_{Total} \left(k_1 K_{Total} - P_{Total}\, k_4\right) c_4}{k_1 \left(c_2 \epsilon P_{Total} + c_4 K_{Total}\right)}$$

$$simplify(P1) = -\frac{P_{Total} \epsilon c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)}{\left(c_2 \epsilon P_{Total} + c_4 K_{Total}\right) k_3}$$

$$simplify(A10) = \frac{\epsilon P_{Total}\, k_1}{\left(k_1 K_{Total} - P_{Total}\, k_4\right) c_4}$$

$$simplify(A01) = -\frac{K_{Total}\, k_3}{\epsilon c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)}$$

**We can see that automatically we get the second invariant for the concentration of [A11] to**

**be true. From the expressions of the variables we note that when [K2] and [P2] are both positive the variable concentrations are positive.**

**4. Now we solve K2Con (the conservation condition associated with the enzyme [K2]) for the concentration of free enzyme [P2] and get the following correlation.**

$$P2 := simplify(solve(P2Con, P2))$$

$$P2 := \frac{\left(k_2 K_{Total} - k_3 P_{Total}\right)\left(k_1 K_{Total} - P_{Total} k_4\right)}{k_1\left(\epsilon k_2 - k_3\right) K_{Total} - k_3 P_{Total}\left(\epsilon k_1 - k_4\right)} \quad \textbf{(2.3.1)}$$

**Note that the denominator can be simplified as**
$\epsilon k_1 \left(k_2 K_{Total} - k_3 P_{Total}\right) - k_3 \left(k_1 K_{Total} - P_{Total} k_4\right)$

**So again, if $\alpha$ is positive, [P2] is positive implying [K2] is positive.**

**Note that in a symmetric state, by mandating [A01] = [A10] we fix the value of $\alpha$ in terms of kinetic constants and total enzyme concentrations.**

$simplify(K1Con) = 0$
$simplify(K2Con) = 0$
$simplify(P1Con) = 0$
$simplify(P2Con) = 0$

$simplify(dA00) = 0$
$simplify(dA11) = 0$
$simplify(dA01) = 0$
$simplify(dA10) = 0$
$simplify(dA00K1) = 0$
$simplify(dA00K2) = 0$
$simplify(dA01P1) = 0$
$simplify(dA10P2) = 0$
$simplify(dA01K2) = 0$
$simplify(dA10K1) = 0$
$simplify(dA11P1) = 0$
$simplify(dA11P2) = 0$
$simplify(dK2) = 0$
$simplify(dK1) = 0$
$simplify(dP1) = 0$
$simplify(dP2) = 0$

**We can thus ensure that for any positive value of $\alpha$, all concentrations are positive. This thus implies that there exists a finite positive $A_{Total}$ value where this asymmetric state is realized.**

**Hence we have shown that the necessary conditions are sufficient for symmetry breaking at**

**some finite $A_{Total}$ value.**

### Position of symmetry breaking (Pitchfork bifurcation)

**At the position of symmetry breaking, we know two insights.**
**1. The system is still symmetric, hence [A01] = [A10], [K1] = [K2], [P1] = [P2]**
**2. The invariants descibing the asymmetric steady state are also true.**

**Using these two information, we can simplify the original system considerably as follows.**
**Note:**

$$K1 := K2 :$$
$$P1 := P2 :$$

$$A00 = A00$$
$$A01 = \frac{K2\ A00\ c_1\ k_1}{P2\ c_4\ k_4}$$
$$A10 = \frac{K2\ A00\ c_1\ k_1}{P2\ c_4\ k_4}$$
$$A11 = \frac{A00\ K2^2\ c_1\ c_2\ k_1\ k_2}{P2^2\ c_4\ k_4\ c_3\ k_3}$$

**But we also know that [A00] and [A11] are invariants given by the following expressions.**

$$A00 := \frac{P_{Total}\ k_4}{c_1\ \left(k_1\ K_{Total} - P_{Total}\ k_4\right)} :$$
$$A11 := -\frac{K_{Total}\ k_2}{\left(k_2\ K_{Total} - k_3\ P_{Total}\right)\ c_3} :$$

**Using this information we can find the value of [K2] and [P2]. Substituiting this into the expressions for conservation of the individual enzymes, we get the concentrations of the respective free enzymes.**

$$A_{Total} = simplify\left(solve\left(ACon, A_{Total}\right)\right)$$
$$A_{Total} = \frac{1}{P2^2\ c_4\ k_4\ c_3\ k_3}\left(2\left(P2\left(\left(K2\left(k_4 + k_1\right)c_1 + \frac{k_4}{2}\right)c_4\ k_3\ P2 + \left(\left(K2\ c_2 + 1\right)k_3 + K2\ c_2\ k_2\right)c_1\ k_1\ K2\right)c_3 + \frac{K2^2\ c_1\ c_2\ k_1\ k_2}{2}\right)A00\right) \quad \textbf{(2.4.1)}$$

**An example of this is done in the read me file using the parameters used in generating the figures. A cross verification of the analytical work and bifurcation analysis is done in this manner.**

## Case 3 Symmetry : Present and Breaks (Arguments for necessary conditions)

**Unlike case 3 symmetry breaking in other models of Random DSP with common enzymes or case 1-2 symmetry breaking in System 3 Random DSP with separate kinases and phosphatases, there exists no simple linear invariant of concentrations at asymmetric steady states post symmetry breaking (upon bifurcation along $A_{Total}$)**

**However the asymmetric steady states in this case 3 System 3 Random DSP exhibit their own signature of symmetry breaking in the form of asymptotic values for concentrations of specific substrates to increasing $A_{Total}$. These asymptotic values manifest as approximate concentration robustness for concentrations of the specific substrates.**

**In this analysis below we show the presence of such symmetries and how the underlying kinetics dictates both the specific pair that exhibits asymptotic response to increasing $A_{Total}$ and the individual asymptote concentrations.**

**Through this process we also obtain necessary conditions for case 3 symmetry to break in System 3 Random DSP (including the necessary conditions for approximate robustness in relevant substrate concentrations through asymptotic behavior)**

**Kinetic constraints for case 3 symmetry (refer main text) are imposed on the original model.**

$$a_3 := k_1 : a_{b3} := k_{b1} : a_{ub3} := k_{ub1} :$$
$$a_4 := k_2 : a_{b4} := k_{b2} : a_{ub4} := k_{ub2} :$$
$$a_1 := k_3 : a_{b1} := k_{b3} : a_{ub1} := k_{ub3} :$$
$$a_2 := k_4 : a_{b2} := k_{b4} : a_{ub2} := k_{ub4} :$$

**In addition to the kinetic constraints the total enzyme concentrations of kinase and phosphatase also need to be equal for exact case 3 symmetry to be present. This is imposed as shown below.**

$$K1_{Total} := P1_{Total} :$$
$$K2_{Total} := P2_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, $c_3$ and $c_4$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) :$$

$k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K1], [K2], [P1] & [P2]) and concentration of the substrates. In order to do this, we use the Maple command *solve,* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1]) is given below in detail.**

**The differential equation of [A00K$_1$] is given by,**

$$\frac{d\left[A00K1\right]}{dt} = dA00K1$$

$$\frac{d\left[A00K1\right]}{dt} = c_1 \left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1 \qquad \textbf{(3.1)}$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K1]). We in turn store this value (the solution returned by the *solve* command) in [A00K$_1$]. This is performed by the following command.**

$A00K1 := solve(dA00K1, A00K1)$

$$A00K1 := K1\,A00\,c_1 \qquad \textbf{(3.2)}$$

**This operation is performed for the other complexes as well as shown below.**

$A01K2 := solve(dA01K2, A01K2) :$
$A11P2 := solve(dA11P2, A11P2) :$
$A01P1 := solve(dA01P1, A01P1) :$
$A00K2 := solve(dA00K2, A00K2) :$
$A10K1 := solve(dA10K1, A10K1) :$
$A11P1 := solve(dA11P1, A11P1) :$
$A10P2 := solve(dA10P2, A10P2) :$

**At this stage we solve for the individual free enzyme concentrations of [P1] and [P2] as shown below. Note that since K1Con, K2Con, P1Con and P2Con are all equal to zero at any given steady state, K1Con-P1Con and K2Con-P2Con is also equal to zero. We define these terms as T = K1Con-P1Con and Q = K2Con-P2Con and use these expressions to solve for [P1] and [P2] as shown below**

$T := K1Con - P1Con :$
$Q := K2Con - P2Con :$

$P1 := solve(T, P1)$

$$P1 := \frac{K1\left(A00\,c_1 + A10\,c_4 + 1\right)}{A01\,c_4 + A11\,c_1 + 1} \qquad \textbf{(3.3)}$$

$P2 := solve(Q, P2)$

$$P2 := \frac{K2\left(A00\,c_3 + A01\,c_2 + 1\right)}{A10\,c_2 + A11\,c_3 + 1} \tag{3.4}$$

**Now at this stage we simplify the remaining unsolved equations from the system of ODE description of the model (dA00, dA01, dA10, dA11). (Note: dP1, dP2, dK1, dK2 are all equal to zero since we solved the respective complex equations in the sytem of ODE description)**

**We definte two terms, M = dA00 + dA01 and N = dA00 + dA10. Note that since dA00, dA01 and dA10 are each equally to zero at any given steady state, M and N are also equal to zero.**

$M := simplify(dA00 + dA01)$

$$M := \frac{1}{A10\,c_2 + A11\,c_3 + 1}\left(K2\left(\left(\left(k_2 - k_3\right)\left(A00\,A10 - A01\,A11\right)c_3 - k_2\left(A01 - A10\right)\right)c_2 - c_3\,k_3\left(A00 - A11\right)\right)\right) \tag{3.5}$$

$N := simplify(dA00 + dA10)$

$$N := -\frac{1}{A01\,c_4 + A11\,c_1 + 1}\left(K1\left(\left(\left(k_1 - k_4\right)\left(A00\,A01 - A10\,A11\right)c_4 + k_1\left(A00 - A11\right)\right)c_1 - c_4\,k_4\left(A01 - A10\right)\right)\right) \tag{3.6}$$

**Thus from M and N we can ascertain that for any given feasible steady stae the following expressions styled m and n need to be equal to zero.**

$m := \left(\left(k_2 - k_3\right)\left(A00\,A10 - A01\,A11\right)c_3 - k_2\left(A01 - A10\right)\right)c_2 - c_3\,k_3\left(A00 - A11\right):$
$n := \left(\left(k_1 - k_4\right)\left(A00\,A01 - A10\,A11\right)c_4 + k_1\left(A00 - A11\right)\right)c_1 - c_4\,k_4\left(A01 - A10\right):$

**At this stage, we introduce the ratio $\epsilon = \frac{A00}{A11}$ and thus m and n simplifies as shown below**

$A00 := \text{epsilon} \cdot A11:$

$simplify(m)$

$$\left(-A11\left(k_2 - k_3\right)\left(-A10\,\epsilon + A01\right)c_3 - k_2\left(A01 - A10\right)\right)c_2 - A11\,c_3\,k_3\left(\epsilon - 1\right) \tag{3.7}$$

$simplify(n)$

$$A11\left(\left(k_1 - k_4\right)\left(A01\,\epsilon - A10\right)c_4 + k_1\left(\epsilon - 1\right)\right)c_1 - c_4\,k_4\left(A01 - A10\right) \tag{3.8}$$

**We first solve for the concentration of [A11] from m as shown below.**

$A11 := simplify(solve(M, A11))$

$$A11 := -\frac{c_2\,k_2\left(A01 - A10\right)}{\left(\left(k_2 - k_3\right)\left(-A10\,\epsilon + A01\right)c_2 + k_3\left(\epsilon - 1\right)\right)c_3} \tag{3.9}$$

**Substituiting this back in n, we get the following**

$$simplify(n)$$

$$\left(\left(\left(\left(\left(\left(\left(A01\,c_1 + A10\,c_3\right)k_4 - A01\,c_1\,k_1\right)\epsilon + \left(-A01\,c_3 - A10\,c_1\right)k_4 + A10\,c_1\,k_1\right)k_2 + k_4\,c_3\,k_3\,\left(-A10\,\epsilon + A01\right)\right)c_4 - c_1\,k_1\,k_2\,(\epsilon - 1)\right)c_2 - k_4\,c_3\,c_4\,k_3\,(\epsilon - 1)\right)(A01 - A10)\right) \Big/ \left(c_3\left(\left(k_2 - k_3\right)\left(-A10\,\epsilon + A01\right)c_2 + k_3\,(\epsilon - 1)\right)\right) \tag{3.10}$$

**We now solve this expression for [A01]. Note that since we are isolating solutions of asymmetry (i.e. [A01] ≠ [A10]), we ignore the symmetric solution while solving the system as shown below.**

$$A01 := simplify(solve(N, A01)[2])$$

$$A01 := \left(c_2\left(A10\,c_4\left(c_3\,\epsilon - c_1\right)k_4 + c_1\,k_1\left(A10\,c_4 - \epsilon + 1\right)\right)k_2 - c_3\,c_4\,k_3\,k_4\left(A10\,c_2\,\epsilon + \epsilon - 1\right)\right) \Big/ \left(c_2\left(\left(\left(-c_1\,\epsilon + c_3\right)k_4 + \epsilon\,c_1\,k_1\right)k_2 - k_4\,c_3\,k_3\right)c_4\right) \tag{3.11}$$

**Finally we use the expressions obtained for [A11] and [A01] to solve for the free enzymes [K1] and [K2] from their respective conservation expressions as shown below.**

$$simplify(A11)$$

$$\frac{\left(-A10\left(c_1\,k_1\,k_2 - k_4\left(\left(c_1 + c_3\right)k_2 - c_3\,k_3\right)\right)c_4 - c_1\,k_1\,k_2\right)c_2 - c_3\,c_4\,k_3\,k_4}{c_3\,c_1\left(\left(k_2 - k_3\right)\left(A10\left(k_1 - k_4\right)(\epsilon + 1)\,c_4 + k_1\right)c_2 - \epsilon\,c_4\,k_3\left(k_1 - k_4\right)\right)} \tag{3.12}$$

$$K1 := simplify(solve(simplify(K1Con), K1))$$

$$K1 := \left(P1_{Total}\left(\left(k_2 - k_3\right)\left(A10\left(k_1 - k_4\right)(\epsilon + 1)\,c_4 + k_1\right)c_2 - \epsilon\,c_4\,k_3\left(k_1 - k_4\right)\right)c_3\right) \Big/ \left(\left(\left(A10\,c_3\left(k_2 - k_3\right)(\epsilon + 1)\,c_4 + \left(k_2 - k_3\right)c_3 - \epsilon\,c_1\,k_2\right)c_2 - c_3\,\epsilon\,k_3\,c_4\right)\left(A10\left(k_1 - k_4\right)c_4 + k_1\right)\right) \tag{3.13}$$

$$K2 := simplify(solve(K2Con, K2))$$

$$K2 := \left(c_4\,P2_{Total}\left(\left(k_1 - k_4\right)\left(A10\left(k_2 - k_3\right)(\epsilon + 1)\,c_2 - \epsilon\,k_3\right)c_4 + c_2\,k_1\left(k_2 - k_3\right)\right)c_1\left(\epsilon\,k_2\left(k_1 - k_4\right)c_1 + c_3\,k_4\left(k_2 - k_3\right)\right)\right) \Big/ \left(\left(\left(\left(k_2 - k_3\right)A10\left(k_2\left(k_1 - k_4\right)c_1 + \epsilon\,c_3\,k_4\left(k_2 - k_3\right)\right)c_2 - k_3\left(k_2\left(k_1 - k_4\right)c_1 + c_3\,k_4\left(k_2 - k_3\right)\right)\epsilon\right)c_4 - c_1\,c_2\,k_1\,k_2\left(k_2 - k_3\right)(\epsilon - 1)\right)\left(\left(A10\,c_1\left(k_1 - k_4\right)(\epsilon + 1)\,c_2 + \epsilon\left(k_1 - k_4\right)c_1 + c_3\,k_4\right)c_4 + c_1\,c_2\,k_1\right)\right) \tag{3.14}$$

**At this stage, by solving a majority of the system of equations describing the sytem (at an asymmetric steady state), we have also obtained correlations between concentrations of variables, primarily, concentrations of substrates and enzymes are obtained as functions of [A10] and ϵ.**

$$simplify(K1Con) = 0$$
$$simplify(K2Con) = 0$$
$$simplify(P1Con) = 0$$
$$simplify(P2Con) = 0$$

$simplify(dA00K1) = 0$
$simplify(dA00K2) = 0$
$simplify(dA01P1) = 0$
$simplify(dA10P2) = 0$
$simplify(dA01K2) = 0$
$simplify(dA10K1) = 0$
$simplify(dA11P1) = 0$
$simplify(dA11P2) = 0$
$simplify(dK2) = 0$
$simplify(dK1) = 0$
$simplify(dP1) = 0$
$simplify(dP2) = 0$

**At an asymmetric steady state, $\epsilon \neq 1$. Thus in order to show the presence of symmetry breaking, we need to show the presence of positive solutions to the above correlations admitting $\epsilon \neq 1$. However due to the complexity of the system this is not possible to do symbolically however can be easily verified for various kinetic values.**

## Necessary conditions for and location of symmetry breaking (pitchfork bifrucation)

**In this section we show the necessary condition for a pitchfork bifurcation/symmetry breaking to exist. The pitchfork bifurcation is by definition the intersection of an asymmetric branch and the symmetric branch of the steady states. Thus the correlation can be easily discerned from the expressions obtained earlier for asymmetric solutions where now $\epsilon = 1$.**

**Thus evaluating the expressions obtained for [A11] or [A00] at epsilon =1 (At symmetric steady states [A00] = [A11])**

$$simplify(eval(eval(simplify(A11)), \text{epsilon} = 1)) \tag{3.1.1}$$

$$\frac{\left(-A10\left(c_1 k_1 k_2 - k_4\left(\left(c_1 + c_3\right) k_2 - c_3 k_3\right)\right) c_4 - c_1 k_1 k_2\right) c_2 - c_3 c_4 k_3 k_4}{2\left(\left(k_2 - k_3\right)\left(A10\left(k_1 - k_4\right) c_4 + \frac{k_1}{2}\right) c_2 - \frac{c_4 k_3 \left(k_1 - k_4\right)}{2}\right) c_1 c_3}$$

$$simplify(eval(eval(simplify(A00)), \text{epsilon} = 1)) \tag{3.1.2}$$

$$\frac{\left(-A10\left(c_1 k_1 k_2 - k_4\left(\left(c_1 + c_3\right) k_2 - c_3 k_3\right)\right) c_4 - c_1 k_1 k_2\right) c_2 - c_3 c_4 k_3 k_4}{2\left(\left(k_2 - k_3\right)\left(A10\left(k_1 - k_4\right) c_4 + \frac{k_1}{2}\right) c_2 - \frac{c_4 k_3 \left(k_1 - k_4\right)}{2}\right) c_1 c_3}$$

$$simplify(eval(eval(simplify(A01)), \text{epsilon} = 1)) \tag{3.1.3}$$

$$A10$$

$$simplify(eval(eval(simplify(A10)), \text{epsilon} = 1)) \tag{3.1.4}$$

$$A10$$

**The above expression for A11 can be simplified further as the term t = 0 as shown below**

*restart* :

$$t := \frac{A11 - \dfrac{k_2}{c_3 \cdot (k_3 - k_2)}}{A10 + \dfrac{k_3}{(k_3 - k_2) \cdot c_2}} = \frac{-A11 + \dfrac{k_4}{(k_1 - k_4) \cdot c_1}}{A10 + \dfrac{k_1}{(k_1 - k_4) \cdot c_4}} :$$

**Simplifying the above equation we get the following,**

$$collect\left(simplify(t), \{k_2, k_3, k_4, k_1\}\right)$$

$$\frac{\left((c_3\, A11 + 1)\, k_2 - c_3\, k_3\, A11\right) c_2}{c_3 \left(c_2\, k_2\, A10 + \left(-c_2\, A10 - 1\right) k_3\right)} = -\frac{c_4 \left(c_1\, k_1\, A11 + \left(-c_1\, A11 - 1\right) k_4\right)}{\left((c_4\, A10 + 1)\, k_1 - c_4\, k_4\, A10\right) c_1} \qquad \textbf{(3.1.5)}$$

**We now use the individual conservation expressions for the enzymes [P1] and [P2]**

$$P1_{Total} = \left(1 + c_4 \cdot A10 + c_1 \cdot A11\right) P1 :$$
$$P2_{Total} = \left(1 + c_2 \cdot A10 + c_3 \cdot A11\right) P2 :$$

**This can be rearranged as shown below**

$$1 + c_4 \cdot A10 = \frac{P1_{Total}}{P1} - c_1 \cdot A11 :$$
$$1 + c_1 \cdot A11 = \frac{P1_{Total}}{P1} - c_4 \cdot A10 :$$
$$1 + c_2 \cdot A10 = \frac{P2_{Total}}{P2} - c_3 \cdot A11 :$$
$$1 + c_3 \cdot A11 = \frac{P2_{Total}}{P2} - c_2 \cdot A10 :$$

**Substituting these expressions in t, we get the following simplification for t**

$$l := simplify\left(\frac{\left(\left(\dfrac{P2_{Total}}{P2} - c_2 \cdot A10\right) k_2 - c_3\, k_3\, A11\right) c_2}{c_3 \left(c_2\, k_2\, A10 + \left(-\dfrac{P2_{Total}}{P2} + c_3 \cdot A11\right) k_3\right)} = -\frac{c_4 \left(c_1\, k_1\, A11 + \left(-\dfrac{P1_{Total}}{P1} + c_4 \cdot A10\right) k_4\right)}{\left(\left(\dfrac{P1_{Total}}{P1} - c_1 \cdot A11\right) k_1 - c_4\, k_4\, A10\right) c_1}\right)$$

$$l := -\frac{\left(A10\,P2\,c_2\,k_2 + A11\,P2\,c_3\,k_3 - k_2\,P2_{Total}\right)c_2}{c_3\left(A10\,P2\,c_2\,k_2 + A11\,P2\,c_3\,k_3 - k_3\,P2_{Total}\right)} \qquad \textbf{(3.1.6)}$$

$$= \frac{c_4\left(c_4\,k_4\,A10\,P1 + A11\,P1\,c_1\,k_1 - k_4\,P1_{Total}\right)}{\left(c_4\,k_4\,A10\,P1 + A11\,P1\,c_1\,k_1 - k_1\,P1_{Total}\right)c_1}$$

**Analyzing the left hand side of the equation 3.1.6**

**Assume k3 is greater than k2. We know that the denominator is positve from equation 3.1.5. Thus the numerator is also positive (since k2<k3), thus the entire left hand side term is strictly negative.**

**Thus the right hand side expression has to be negative as well for the equation to be true.**

**Now suppose k4 > k1. Then the numerator as it appears in 3.1.6 is positive. And since k4 > k1, the denominator also has to be positive. Thus the right hand side is entirely positive.**

**This is a contradiction and so the equation cannot be true under this condition.**

**Hence, should the correlation above hold good, k1 > k4 when k3 > k2. Similarly one can reason that when k2 > k3, k4 > k1 is necessary.**

**Thus these become necessary kinetic constraints for symmetry breaking to be feasible. Note that these are not shown to be suffiient however extensive computational simulations show indications that these conditions are indeed sufficient for symmetry breaking to exist at some finite total substrate concentration.**

**The exact value at the pitchfork bifurcation point can be known by substituting equation 3.1.2 in the following correlation between the concentrations at the symmetric point involving total enzyme concentrations of [P1] and [P2].**

$$E := \frac{P1_{Total}}{P2_{Total}} \cdot \frac{\left(1 + c_2 \cdot A10 + c_3 \cdot A11\right)}{\left(1 + c_4 \cdot A10 + c_1 \cdot A11\right)} - \frac{\left(c_3 \cdot k_3 \cdot A11 - c_2 \cdot k_2 \cdot A10\right)}{\left(-c_1 \cdot k_1 \cdot A11 + c_4 \cdot k_4 \cdot A10\right)} :$$

**System 3 Random DSP breaks symmetry with a unique asymmetric signature in the form of asymptotes for the individual substrates. However the value of the asymptote and the substrate associated with it is given entirely by the kinetics.**

**Upon bifurcation along $A_{Total}$ two scenarios are possible, as seen computationally**

1. [A00] becomes 0(Infinity) as [A11] becomes Infinity(0) (On the asymmetric branches)

In this scenario, [A01] and [A10] each inidivually saturate towards an unique asymptotic value.

2. [A01] becomes 0(Infinity) as [A10] becomes Infinity(0) (On the asymmetric branches)

In this scenario, [A01] and [A10] each inidivually saturate towards an unique asymptotic value.

### Scenario 1 : Asymptotes (If [A01] & [A10] is robust, [A00] & [A11] becomes 0/infinite)

**Since [A00] or [A11] becomes infinite, we are interested in solutions of the kind, epsilon = 0/infinity**

**1. [A11] = infinity, [A00] = 0, epsilon = 0**

$$simplify(eval(eval(simplify(A11)), \text{epsilon} = 0))$$

$$\frac{\left(-A10\left(\left(\left(-c_1 - c_3\right)k_4 + c_1 k_1\right)k_2 + k_4 c_3 k_3\right)c_2 - k_4 c_3 k_3\right)c_4 - c_1 c_2 k_1 k_2}{\left(A10\left(k_1 - k_4\right)c_4 + k_1\right)c_2 c_1 c_3\left(k_2 - k_3\right)} \quad (3.2.1)$$

**This value is infinite - so the denominator must be equal to zero. This gives us the asymptotic value of [A10] as shown below**

$$A10 = simplify(solve(denom(simplify(eval(eval(simplify(A11)), \text{epsilon} = 0))), A10))$$

$$A10 = -\frac{k_1}{c_4\left(k_1 - k_4\right)} \quad (3.2.2)$$

**At this value, we can evaluate the other concentrations, and so we have the second asymptote for [A01].**

$$simplify\left(eval\left(eval(A01, \text{epsilon} = 0), A10 = -\frac{k_1}{c_4\left(k_1 - k_4\right)}\right)\right)$$

$$\frac{k_3}{\left(k_2 - k_3\right)c_2} \quad (3.2.3)$$

**1. [A11] = 0, [A00] = Infinity, epsilon = infinity**

$$simplify(A00)$$

$$-\frac{\left(\left(A10\left(c_1 k_1 k_2 - k_4\left(\left(c_1 + c_3\right)k_2 - c_3 k_3\right)\right)c_4 + c_1 k_1 k_2\right)c_2 + c_3 c_4 k_3 k_4\right)\epsilon}{\left(\left(k_2 - k_3\right)\left(A10\left(k_1 - k_4\right)(\epsilon + 1)c_4 + k_1\right)c_2 - \epsilon c_4 k_3\left(k_1 - k_4\right)\right)c_3 c_1} \quad (3.2.4)$$

**This value is infinite - so the denominator must be equal to zero. This gives us the asymptotic value of [A10] as shown below**

$$simplify(limit(solve(denom(A00), A10), \text{epsilon} = \text{infinity}))$$

$$\frac{k_3}{c_2\left(k_2 - k_3\right)} \quad (3.2.5)$$

**At this value, we can evaluate the other concentrations, and so we have the second asymptote for [A01].**

$$simplify\left(eval\left(limit((A01), \text{epsilon} = \text{infinity}), A10 = \frac{k_3}{c_2\,(k_2 - k_3)}\right)\right)$$

$$-\frac{k_1}{c_4\,(k_1 - k_4)} \qquad \textbf{(3.2.6)}$$

**Thus for symmetry to break and lead to robustness in A10 and A10, the necessary condition is $k_2 > k_3$ and $k_4 > k_1$ and these are the respective asymptotes**

## Scenario 2 : Asymptotes (If [A00] & [A11] is robust, [A01] & [A10] becomes 0/infinite)

**Since [A01] or [A10] becomes infinite, we are interested in solutions of the kind, = 0 and epsilon = infinity**

**1. [A01] = infinity, [A10] = 0**

$$simplify(A01)$$

$$\left(\left(A10\,c_4\,(c_3\,\epsilon - c_1)\,k_4 + c_1\,k_1\,(A10\,c_4 - \epsilon + 1)\right)c_2\,k_2 - c_3\,c_4\,k_3\,k_4\,(A10\,c_2\,\epsilon + \epsilon - 1)\right) \Big/ \left(\left(\left(\left(-c_1\,\epsilon + c_3\right)k_4 + \epsilon\,c_1\,k_1\right)k_2 - k_4\,c_3\,k_3\right)c_2\,c_4\right) \qquad \textbf{(3.3.1)}$$

**This value is infinite - so the denominator must be equal to zero. This gives us the asymptotic value of epsilon as shown below**

$$\text{epsilon} = solve(denom(A01), \text{epsilon})$$

$$\epsilon = -\frac{c_3\,k_4\,(k_2 - k_3)}{k_2\,(k_1 - k_4)\,c_1} \qquad \textbf{(3.3.2)}$$

**At this value, we can evaluate the other concentrations and thus we have our two invariants**

$$simplify\left(eval\left(simplify(A11), \epsilon = -\frac{c_3\,k_4\,(k_2 - k_3)}{k_2\,(k_1 - k_4)\,c_1}\right)\right)$$

$$-\frac{k_2}{c_3\,(k_2 - k_3)} \qquad \textbf{(3.3.3)}$$

$$simplify\left(eval\left(simplify(A00), \epsilon = -\frac{c_3\,k_4\,(k_2 - k_3)}{k_2\,(k_1 - k_4)\,c_1}\right)\right)$$

$$\frac{k_4}{c_1\,(k_1 - k_4)} \qquad \textbf{(3.3.4)}$$

**2. [A01] = 0, [A10] = Infinity**

$$simplify(A01)$$

$$\left(c_2\left(A10\,c_4\left(\epsilon\,c_3-c_1\right)k_4+c_1\,k_1\left(A10\,c_4-\epsilon+1\right)\right)k_2-c_3\,c_4\,k_3\,k_4\left(A10\,c_2\,\epsilon+\epsilon-1\right)\right)\Big/\left(c_4\,c_2\left(\left(\left(-\epsilon\,c_1+c_3\right)k_4+\epsilon\,c_1\,k_1\right)k_2-k_4\,c_3\,k_3\right)\right) \tag{3.3.5}$$

**This value is infinite - so the denominator must be equal to zero. This gives us the asymptotic value of epsilon as shown below, where [A10] = infinity**

$$\text{epsilon} = simplify(limit(solve(A01, \text{epsilon}), A10 = \text{infinity}))$$

$$\epsilon = -\frac{c_1\left(k_1-k_4\right)k_2}{c_3\,k_4\left(k_2-k_3\right)} \tag{3.3.6}$$

**At this value, we can evaluate the other concentrations and thus we have our two invariants**

$$limit\left(simplify\left(eval\left(simplify(A11), \epsilon = -\frac{c_1\left(k_1-k_4\right)k_2}{c_3\,k_4\left(k_2-k_3\right)}\right)\right), A10 = \text{infinity}\right)$$

$$\frac{k_4}{\left(k_1-k_4\right)c_1} \tag{3.3.7}$$

$$limit\left(simplify\left(eval\left(simplify(A00), \epsilon = -\frac{c_1\left(k_1-k_4\right)k_2}{c_3\,k_4\left(k_2-k_3\right)}\right)\right), A10 = \text{infinity}\right)$$

$$-\frac{k_2}{\left(k_2-k_3\right)c_3} \tag{3.3.8}$$

**Thus for symmetry to break and lead to robustness in A00 and A11, the necessary condition is $k_3 > k_2$ and $k_1 > k_4$ and these are the respective asymptotes**

# Mixed Random 1 DSP : Common Kinase Common Phosphatase Case 2 - Present and Doesn't Break

In this Maple file we analytically show the infeasibility of case 2 symmetry breaking in the Mixed Random 1 DSP network with common kinase and common phosphatase effecting distributive random phosphorylation and random processive dephoshorylation respectively. We do this by first describing the model as a system of ODEs along with the associated enzyme and substrate conservations. We then impose the kinetic constraints pertinent to case 2 symmetry. By solving for the steady state of the system of ODEs we obtain relations between concentrations of the substrate variables in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. case 2 symmetry breaking requires symmetry between [A01] & [A10]; By leveraging this insight, we show the infeasibility of case 2 symmetry breaking by revealing that [A01] and [A10] are always equal for any given feasible steady state. These procedures are carried out in detail below using built in Maple commands.

Note: A subscript is used to distinguish between the two different complexes formed between [K] and [A00]. Similarly a subscript is used to distinguish between the two distinct complexes formed between [P] and [A11].

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(VectorCalculus) : with(LinearAlgebra) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := -k_{b1}\,A00\cdot K - a_{b1}\,A00\cdot K + k_{ub1}\cdot A00K_1 + a_{ub1}\cdot A00K_2 + a_4\cdot A10P + k_4\cdot A01P :$$
$$dA01 := -k_{b2}\,A01\cdot K + k_1\,A00K_1 + k_{ub2}\,A01K :$$
$$dA10 := -a_{b2}\,A10\cdot K + a_1\,A00K_2 + a_{ub2}\,A10K :$$
$$dA11 := k_2\,A01K + a_2\,A10K - k_{b3}\cdot A11\cdot P - a_{b3}\cdot A11\cdot P + k_{ub3}\cdot A11P_1 + a_{ub3}\cdot A11P_2 :$$

$$dA00K_1 := k_{b1}\cdot A00\cdot K - \left(k_1 + k_{ub1}\right)\cdot A00K_1 :$$
$$dA10K := a_{b2}\cdot A10\cdot K - \left(a_2 + a_{ub2}\right)\cdot A10K :$$
$$dA00K_2 := a_{b1}\cdot A00\cdot K - \left(a_{ub1} + a_1\right)\cdot A00K_2 :$$
$$dA01K := k_{b2}\cdot A01\cdot K - \left(k_{ub2} + k_2\right)\cdot A01K :$$
$$dA11P_1 := k_{b3}\cdot A11\cdot P - \left(k_{ub3} + k_3\right)\cdot A11P_1 :$$
$$dA10P := a_3\cdot A11P_2 - a_4\cdot A10P :$$
$$dA11P_2 := a_{b3}\cdot A11\cdot P - \left(a_{ub3} + a_3\right)\cdot A11P_2 :$$
$$dA01P := k_3\cdot A11P_1 - k_4\cdot A01P :$$

$$dK := -k_{b1}\cdot A00\cdot K + \left(k_1 + k_{ub1}\right)\cdot A00K_1 - a_{b2}\cdot A10\cdot K + \left(a_2 + a_{ub2}\right)\cdot A10K - a_{b1}\cdot A00\cdot K + \left(a_{ub1}\right.$$

$$+ a_1)\cdot A00K_2 - k_{b2}\cdot A01\cdot K + (k_{ub2} + k_2)\cdot A01K :$$
$$dP := -k_{b3}\cdot A11\cdot P + k_{ub3}\cdot A11P_1 + a_4\cdot A10P - a_{b3}\cdot A11\cdot P + a_{ub3}\cdot A11P_2 + k_4\cdot A01P :$$

**The above equations are also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K_1 - A01K - A00K_2 - A10K - A11P_1 - A10P - A11P_2 - A01P :$$
$$KCon := K_{Total} - K - A00K_1 - A10K - A00K_2 - A01K :$$
$$PCon := P_{Total} - P - A11P_1 - A10P - A11P_2 - A01P :$$

## Proof for impossibility of symmetry breaking

**Kinetic constraints for case 2 symmetry (refer main text) are imposed on the original model.**

$$a_1 := k_1 : a_{b1} := k_{b1} : a_{ub1} := k_{ub1} :$$
$$a_2 := k_2 : a_{b2} := k_{b2} : a_{ub2} := k_{ub2} :$$
$$a_3 := k_3 : a_{b3} := k_{b3} : a_{ub3} := k_{ub3} :$$
$$a_4 := k_4 :$$

**There are no constraints on the total enzyme concentrations for case 2 symmetry to be present in Mixed Random 1 DSP. At this stage we introduce auxiliary constants $c_1$, $c_2$, and $c_3$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1\cdot(k_{ub1} + k_1) :$$
$$k_{b2} := c_2\cdot(k_{ub2} + k_2) :$$
$$k_{b3} := c_3\cdot(k_{ub3} + k_3) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ( [K] & [P]) and the concentrations of the fully unmodified substrate [A00] and the fully modified substrate [A11]. In order to do this, we use the Maple command *solve* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using $[A00K_1]$ is given below in detail).**

**The differential equation of $[A00K_1]$ is given by,**

$$\frac{d[A00K_1]}{dt} = dA00K_1 \qquad \frac{d[A00K_1]}{dt} = c_1(k_1 + k_{ub1})A00K - (k_1 + k_{ub1})A00K_1$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K$_1$]). We in turn store this value (the solution returned by the *solve* command) in [A00K$_1$]. This is performed by the following command.**

$$A00K_1 := solve\left(dA00K_1, A00K_1\right)$$

$$A00K_1 := K\,A00\,c_1 \tag{1.1}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A01K := solve(dA01K, A01K):$$
$$A00K_2 := solve\left(dA00K_2, A00K_2\right):$$
$$A10K := solve(dA10K, A10K):$$
$$A11P_1 := solve\left(dA11P_1, A11P_1\right):$$
$$A10P := solve(dA10P, A10P):$$
$$A11P_2 := solve\left(dA11P_2, A11P_2\right):$$
$$A01P := solve(dA01P, A01P):$$

$$S := solve(\{dA10, dA01\}, \{A01, A10\}):$$

$$A00 := eval(A00, S):$$
$$A01 := eval(A01, S):$$
$$A10 := eval(A10, S):$$
$$A11 := eval(A11, S):$$

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = A00$$
$$A01 = \frac{A00\,c_1\,k_1}{c_2\,k_2}$$
$$A10 = \frac{A00\,c_1\,k_1}{c_2\,k_2}$$
$$A11 = A11$$

## Proof for impossibility of symmetry breaking

**Thus from this we can clearly see that irrespective of kinetic parameters, the concentration of [A01] is always going to be equal to the concentration of [A10]. Thus there is no scope for any asymmetric steady state or case 2 symmetry breaking.**

$$A01 = \frac{A00\, c_1\, k_1}{c_2\, k_2}$$

$$A10 = \frac{A00\, c_1\, k_1}{c_2\, k_2}$$

# Mixed Random 2 DSP : Separate Kinase Common Phosphatase Case 2 - Present and Breaks

In this Maple file we analytically show the presence of case 2 symmetry breaking in the Mixed Random 2 DSP network with separate kinase and common phosphatase effecting distributive random phosphorylation and processive dephoshorylation respectively. We do this by first describing the model as a system of ODEs along with the associated substrate and enzyme conservations. We then impose the kinetic constraints pertaining to case 2 symmetry. By solving for the steady state of the system of ODEs we obtain relations between concentrations of the substrate variables in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. case 2 symmetry breaking requires symmetry between [A01] & [A10] and [K1] & [K2]. By leveraging this insight and isolating steady states not of this type, we ascertain the features of the asymmetric steady state emerging from symmetry breaking. These procedures are carried out in detail below using built in Maple commands.

Note: A subscript is used to distinguish between the two different complexes formed between [P] and [A11].

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := -k_{b1}\,A00\cdot K1 - a_{b1}\,A00\cdot K2 + k_{ub1}\cdot A00K1 + a_{ub1}\cdot A00K2 + a_4\cdot A10P + k_4\cdot A01P :$$
$$dA01 := -k_{b2}\,A01\cdot K2 + k_1\,A00K1 + k_{ub2}\,A01K2 :$$
$$dA10 := -a_{b2}\,A10\cdot K1 + a_1\,A00K2 + a_{ub2}\,A10K1 :$$
$$dA11 := k_2\cdot A01K2 + a_2\cdot A10K1 - k_{b3}\cdot A11\cdot P - a_{b3}\cdot A11\cdot P + k_{ub3}\cdot A11P_1 + a_{ub3}\cdot A11P_2 :$$

$$dA00K1 := k_{b1}\cdot A00\cdot K1 - \left(k_1 + k_{ub1}\right)\cdot A00K1 :$$
$$dA10K1 := a_{b2}\cdot A10\cdot K1 - \left(a_2 + a_{ub2}\right)\cdot A10K1 :$$
$$dA00K2 := a_{b1}\cdot A00\cdot K2 - \left(a_{ub1} + a_1\right)\cdot A00K2 :$$
$$dA01K2 := k_{b2}\cdot A01\cdot K2 - \left(k_{ub2} + k_2\right)\cdot A01K2 :$$
$$dA11P_1 := k_{b3}\cdot A11\cdot P - \left(k_{ub3} + k_3\right)\cdot A11P_1 :$$
$$dA10P := a_3\cdot A11P_2 - a_4\cdot A10P :$$
$$dA11P_2 := a_{b3}\cdot A11\cdot P - \left(a_{ub3} + a_3\right)\cdot A11P_2 :$$
$$dA01P := k_3\cdot A11P_1 - k_4\cdot A01P :$$

$$dK1 := -k_{b1}\cdot A00\cdot K1 + \left(k_1 + k_{ub1}\right)\cdot A00K1 - a_{b2}\cdot A10\cdot K1 + \left(a_2 + a_{ub2}\right)\cdot A10K1 :$$
$$dK2 := -a_{b1}\cdot A00\cdot K2 + \left(a_{ub1} + a_1\right)\cdot A00K2 - k_{b2}\cdot A01\cdot K2 + \left(k_{ub2} + k_2\right)\cdot A01K2 :$$

$$dP := -k_{b3}\cdot A11\cdot P + k_{ub3}\cdot A11P_1 + a_4\cdot A10P - a_{b3}\cdot A11\cdot P + a_{ub3}\cdot A11P_2 + k_4\cdot A01P :$$

**The model is also associated with conservation conditions for the substrate and enzyme concentrations which are described below. Here we store the conservation expressions in ACon, K1Con, K2Con and PCon for the substrate and respective enzymes. The right hand side of each of these expressions are always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K1 - A01K2 - A00K2 - A10K1 - A11P_1 - A10P - A11P_2 - A01P :$$
$$K1Con := K1_{Total} - K1 - A00K1 - A10K1 :$$
$$K2Con := K1_{Total} - K2 - A00K2 - A01K2 :$$
$$PCon := P_{Total} - P - A11P_1 - A10P - A11P_2 - A01P :$$

**Kinetic constraints for case 2 symmetry (refer main text) are imposed on the original model.**

$$a_1 := k_1 : a_{b1} := k_{b1} : a_{ub1} := k_{ub1} :$$
$$a_2 := k_2 : a_{b2} := k_{b2} : a_{ub2} := k_{ub2} :$$
$$a_3 := k_3 : a_{b3} := k_{b3} : a_{ub3} := k_{ub3} :$$
$$a_4 := k_4 :$$

**In addition to the kinetic constraints the total enzyme concentrations of the two kinases need to be equal for exact case 2 symmetry to be present in Mixed-Random 2 DSP. This is imposed as shown below.**

$$K2_{Total} := K1_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, and $c_3$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1\cdot\left(k_{ub1} + k_1\right) :$$
$$k_{b2} := c_2\cdot\left(k_{ub2} + k_2\right) :$$
$$k_{b3} := c_3\cdot\left(k_{ub3} + k_3\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the free enzymes ([K1], [K2] & [P]) and the partially modified substrate [A01]. In order to do this, we use the Maple command *solve* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1] is given below in detail).**

**The differential equation of [A00K1] is given by,**

$$\frac{d\,[A00K1]}{dt} = dA00K1$$

$$\frac{d\,[A00K1]}{dt} = c_1\left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1 \qquad (1)$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K1]). We in turn store this value (the solution returned by the *solve* command) in [A00K1]. This is performed by the following command.**

$$A00K1 := solve(dA00K1, A00K1)$$

$$A00K1 := K1\,A00\,c_1 \qquad (2)$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are extracted from this vector using the *eval* command.**

$$A01K2 := solve(dA01K2, A01K2):$$
$$A00K2 := solve(dA00K2, A00K2):$$
$$A10K1 := solve(dA10K1, A10K1):$$
$$A11P_1 := solve\left(dA11P_1, A11P_1\right):$$
$$A10P := solve(dA10P, A10P):$$
$$A11P_2 := solve\left(dA11P_2, A11P_2\right):$$
$$A01P := solve(dA01P, A01P):$$

$$S := solve(\{dA00, dA11, dA10\}, \{A00, A11, A10\}):$$

$$A00 := eval(A00, S):$$
$$A01 := eval(A01, S):$$
$$A10 := eval(A10, S):$$
$$A11 := eval(A11, S):$$

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = \frac{K2\,A01\,c_2\,k_2}{K1\,c_1\,k_1}$$
$$A01 = A01$$
$$A10 = \frac{K2^2\,A01}{K1^2}$$
$$A11 = \frac{K2\,A01\,c_2\,k_2\,(K1 + K2)}{2\,K1\,P\,c_3\,k_3}$$

### Proof for invariant in the asymmetric branches

**We know that K1Con and K2Con are both individually equal to zero always. Thus at a given**

**steady state, K1Con - K2Con must also be equal to zero.**

**We thus introduce the term T = (K1Con - K2Con) = 0 and also introduce a new ratio, $\epsilon$ = [K1]/[K2]. Note: As discussed in the main text, the symmetric steady state is one where [K1] = [K2] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

$T := K1Con - K2Con = 0$ :
$K1 := \epsilon \cdot K2$ :

**The following command (*simplify*), simplifies the expression algebraically**

$$simplify(T) = \frac{\left(-\epsilon\, k_1 + A01\, c_2\, \left(k_1 - k_2\right)\right) K2\, (\epsilon - 1)}{k_1\, \epsilon} = 0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term $\left(-\epsilon\, k_1 + A01\, c_2\, \left(k_1 - k_2\right)\right)$ in the expression needs to necessarily be zero. This term is an expression in the partial substrate form [A01], $\epsilon$ and kinetic constants. Thus solving this to isolate the partially modified substrate form we get the following. Here we use the *solve* command from Maple to solve T for [A01] and the *simplify* command to algebraically simplify the resulting expression.**

$A01 := simplify(solve(T, A01))$

$$A01 := \frac{\epsilon\, k_1}{\left(k_1 - k_2\right) c_2} \tag{1.1}$$

**Substituting this value for [A01] it back into the expression for the concentration of [A01], we get the following correlation.**

$$A00 = \frac{k_2}{\left(k_1 - k_2\right) c_1}$$

**We note that the concentration of [A00] is fixed in the asymmetric steady state and is given by a few key kinetic parameters. Using this information (the concentration of [A00] at an asymmetric steady state) - we solve for [K2], [P] at this asymmetric steady state using their individual enzyme conservation equations. This is done by finding solutions for [K2] and [P] using equations K2Con and PCon as shown below, using the Maple command *solve*.**

$P := simplify(solve(PCon, P))$ :
$K2 := simplify(solve(K2Con, K2))$ :

$$simplify(P) = \frac{-k_2\, \left(k_3 + k_4\right) K1_{Total} + P_{Total}\, k_3\, k_4}{k_3\, k_4}$$

$simplify(K2) =$

$$\frac{KI_{Total}\left(k_1 - k_2\right)}{k_1\left(\epsilon + 1\right)}$$

**Using these expressions, we find that the concentration of [A11] and [P] is also fixed and is given by only a few kinetic constants and total enzyme concentrations as shown below.**

$$simplify(P) = \frac{-k_2\left(k_3 + k_4\right) KI_{Total} + P_{Total}\, k_3\, k_4}{k_3\, k_4}$$

$$simplify(A11) = -\frac{k_2\, k_4\, KI_{Total}}{2\, c_3\left(k_2\left(k_3 + k_4\right) KI_{Total} - P_{Total}\, k_3\, k_4\right)}$$

## Necessary conditions

**Since substrate concentration are always necessarily positive, the expression for [A00] and [A11] above must be positive. The numerators are only a function of kinetic parameters which are always positive, thus the denominator must necessarily be positive to ensure that the resulting concentration is positive. This gives us the necessary condition for an asymmetric state to exist as follows.**

1. $k_2 < k_1$
2. $k_2\, KI_{total}\left(k_3 + k_4\right) < P_{total}\, k_3\, k_4$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient for an asymmetric steady state to exist for some positive $A_{Total}$ value. i.e. We show that upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that a feasible steady state in this context is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that the asymmetric states defined by the concentration (invariant) of [A00], [A11] & [P] described above is indeed a feasible solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an asymmetric steady state, as seen above the concentration of [A00], [A11] & [P] are fixed by a few kinetic constants and total enzyme concentrations,**

$$A00 := \frac{k_2}{\left(k_1 - k_2\right) c_1} = \frac{k_2}{\left(k_1 - k_2\right) c_1}$$

$$A11 := -\frac{k_2\, k_4\, KI_{Total}}{2\, c_3\left(k_2\left(k_3 + k_4\right) KI_{Total} - P_{Total}\, k_3\, k_4\right)} =$$

$$-\frac{1}{2}\,\frac{k_2\,k_4\,KI_{Total}}{c_3\left(k_2\left(k_3+k_4\right)KI_{Total}-P_{Total}\,k_3\,k_4\right)}$$

$$P := \frac{-k_2\left(k_3+k_4\right)KI_{Total}+P_{Total}\,k_3\,k_4}{k_3\,k_4} = \frac{-k_2\left(k_3+k_4\right)KI_{Total}+P_{Total}\,k_3\,k_4}{k_3\,k_4}$$

**The other variables in this asymmetric state are thus given by**

$$A00 = \frac{k_2}{\left(k_1-k_2\right)c_1}$$

$$A11 = -\frac{k_2\,k_4\,KI_{Total}}{2\,c_3\left(k_2\left(k_3+k_4\right)KI_{Total}-P_{Total}\,k_3\,k_4\right)}$$

$$A00K1 = \frac{\epsilon\,KI_{Total}\,k_2}{k_1\left(\epsilon+1\right)}$$

$$A00K2 = \frac{KI_{Total}\,k_2}{k_1\left(\epsilon+1\right)}$$

$$A01K2 = \frac{KI_{Total}\,\epsilon}{\epsilon+1}$$

$$A10K1 = \frac{KI_{Total}}{\epsilon+1}$$

$$K1 = \frac{\epsilon\,KI_{Total}\left(k_1-k_2\right)}{k_1\left(\epsilon+1\right)}$$

$$K2 = \frac{KI_{Total}\left(k_1-k_2\right)}{k_1\left(\epsilon+1\right)}$$

$$simplify\left(A11P_1\right) = \frac{KI_{Total}\,k_2}{2\,k_3}$$

$$simplify\left(A11P_2\right) = \frac{KI_{Total}\,k_2}{2\,k_3}$$

$$simplify\left(A01P\right) = \frac{KI_{Total}\,k_2}{2\,k_4}$$

$$simplify\left(A10P\right) = \frac{KI_{Total}\,k_2}{2\,k_4}$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify\left(dA00\right) = 0$$
$$simplify\left(dA11\right) = 0$$
$$simplify\left(dA01\right) = 0$$
$$simplify\left(dA10\right) = 0$$

$simplify(dA00K1) = 0$
$simplify(dA00K2) = 0$
$simplify(dA01P) = 0$
$simplify(dA10P) = 0$
$simplify(dA01K2) = 0$
$simplify(dA10K1) = 0$
$simplify(dA11P_1) = 0$
$simplify(dA11P_2) = 0$
$simplify(dP) = 0$
$simplify(dK1) = 0$
$simplify(dK2) = 0$

**Hence all that remains to be shown is that the variables (As described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition $(k_3 > k_1)$ is satisfied**
2. **$\epsilon$ is positive.**

**However if $\epsilon$ is positive all concentrations are automatically positive (provided necessary conditions are satisfied).**

**Thus this means that for every positive value of $\epsilon \neq 1$, all concentrations are positive and the conservation of kinase and phosphatase is also satisfied.**

**Since the concentrations are all positive there exists a unique finite $A_{Total}$ value for every $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

## Prediction of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitchfork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two insights.**

**1. The system is still symmetric, hence [A01] = [A10] and [K1] = [K2].**
**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$K1 := K2 : \epsilon := 1 :$

$$A00 := \frac{k_2}{(k_1 - k_2)\, c_1} :$$

$$A11 := -\frac{k_2\, k_4\, KI_{Total}}{2\, c_3\, (k_2\, (k_3 + k_4)\, KI_{Total} - P_{Total}\, k_3\, k_4)} :$$

$$P := \frac{-k_2\, (k_3 + k_4)\, KI_{Total} + P_{Total}\, k_3\, k_4}{k_3\, k_4} :$$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point)**

$$A_{Total} = simplify\left(solve\left(simplify(ACon), A_{Total}\right)\right)$$

$$\begin{aligned} A_{Total} = \Big( & -2\,(k_3 + k_4)\, c_3\, c_1\, KI_{Total}^{\,2}\, \left((k_1 + k_4)\, k_3 + k_1\, k_4\right) c_2\, k_2^{\,3} + 2\Big(c_3\,\left(c_1\, P_{Total}\, k_4^{\,2}\right. \\ & \left. + k_1\,(c_1\, P_{Total} + 1)\, k_4 + c_1\, k_1^{\,2}\, KI_{Total}\right) k_3^{\,2} + 2\, k_1\, k_4 \left(\left(\frac{1}{2}\, c_1\, c_3\, P_{Total} + \frac{1}{4}\, c_1\right.\right. \\ & \left.\left. + \frac{1}{2}\, c_3\right) k_4 + c_1\, c_3\, k_1\, KI_{Total}\right) k_3 + c_1\, c_3\, k_1^{\,2}\, k_4^{\,2}\, KI_{Total}\Big)\, KI_{Total}\, c_2\, k_2^{\,2} \\ & - 2\, k_3\, k_1\, k_4 \Big(c_3\,\left(c_2\, P_{Total}\, k_4 + c_1\, k_1\, KI_{Total}\,\left(-KI_{Total}\, c_2 + c_2\, P_{Total} - 2\right)\right) k_3 \\ & + c_1\, k_1\, k_4\, KI_{Total}\left(c_2\, c_3\, P_{Total} - c_2\, c_3\, KI_{Total} + \frac{1}{2}\, c_2 - 2\, c_3\right)\Big) k_2 \\ & - 2\, c_1\, c_3\, k_1^{\,2}\, k_3^{\,2}\, k_4^{\,2}\, P_{Total}\,(KI_{Total}\, c_2 + 2)\Big) \Big/ \left(2\, k_3\,(k_1 - k_2)\, k_1\, c_3\,\left(k_2\,(k_3\right.\right. \\ & \left.\left. + k_4)\, KI_{Total} - P_{Total}\, k_3\, k_4\right) c_1\, k_4\, c_2\right) \end{aligned} \qquad \textbf{(4.1)}$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figures (Fig 2E).**

# Mixed Random 2a DSP : Separate Kinase Common Phosphatase Case 2 - Present and Breaks

In this Maple file we analytically show the presence of case 2 symmetry breaking in the Mixed Random 2a DSP network with separate kinase effecting distributive random phosphorylation and unsaturated processive dephoshorylation respectively (modelled as a linear reaction). We do this by first describing the model as a system of ODEs along with the associated substrate and enzyme conservations. We then impose the kinetic constraints pertaining to case 2 symmetry. By solving for the steady state of the system of ODEs we obtain relations between concentrations of the substrate variables in terms of each other and the free enzyme concentrations. After this, we identify key symmetric pairings that represent the symmetric steady state. i.e. case 2 symmetry breaking requires symmetry between [A01] & [A10] and [K1] & [K2]. By leveraging this insight and isolating steady states not of this type, we ascertain the features of the asymmetric steady state emerging from symmetry breaking. These procedures are carried out in detail below using built in Maple commands.

Note: A subscript is used to distinguish between the two different complexes formed between [P] and [A11].

**We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer Appendix 2 figure 10). Here dA00 represents d[A00]/dt and similarly for other expressions. At steady state thus, each of the right hand sides of these expressions will be equal to zero.**

$$dA00 := -k_{b1}\,A00\,K1 - p_{b1}\,A00\,K2 + k_{ub1}\cdot A00K1 + p_{ub1}\cdot A00K2 + k_3\cdot A11 + p_3\cdot A11 :$$
$$dA01 := -k_{b2}\,A01\,K2 + k_1\,A00K1 + k_{ub2}\,A01K2 :$$
$$dA10 := -p_{b2}\,A10\,K1 + p_1\,A00K2 + p_{ub2}\,A10K1 :$$
$$dA11 := k_2\,A01K2 + p_2\,A10K1 - k_3\cdot A11 - p_3\cdot A11 :$$

$$dA00K1 := k_{b1}\cdot A00\cdot K1 - \left(k_1 + k_{ub1}\right)\cdot A00K1 :$$
$$dA10K1 := p_{b2}\cdot A10\cdot K1 - \left(p_2 + p_{ub2}\right)\cdot A10K1 :$$
$$dA00K2 := p_{b1}\cdot A00\cdot K2 - \left(p_{ub1} + p_1\right)\cdot A00K2 :$$
$$dA01K2 := k_{b2}\cdot A01\cdot K2 - \left(k_{ub2} + k_2\right)\cdot A01K2 :$$

$$dK1 := -k_{b1}\cdot A00\cdot K1 + \left(k_1 + k_{ub1}\right)\cdot A00K1 - p_{b2}\cdot A10\cdot K1 + \left(p_2 + p_{ub2}\right)\cdot A10K1 :$$
$$dK2 := -p_{b1}\cdot A00\cdot K2 + \left(p_{ub1} + p_1\right)\cdot A00K2 - k_{b2}\cdot A01\cdot K2 + \left(k_{ub2} + k_2\right)\cdot A01K2 :$$

**The above equations are associated with conservation conditions which are described below. Here we store the conservation expressions in ACon, K1Con and K2Con for the substrate and the respective enzymes. The right hand side of each of these expressions are always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A00 - A10 - A01 - A11 - A00K1 - A01K2 - A00K2 - A10K1 :$$
$$K1Con := K1_{Total} - K1 - A00K1 - A10K1 :$$
$$K2Con := K1_{Total} - K2 - A00K2 - A01K2 :$$

**Kinetic constraints for case 2 symmetry (refer main text) are imposed on the original model.**

$$p_1 := k_1 : p_{b1} := k_{b1} : p_{ub1} := k_{ub1} :$$
$$p_2 := k_2 : p_{b2} := k_{b2} : p_{ub2} := k_{ub2} :$$
$$p_3 := k_3 :$$

**In addition to the kinetic constraints the total enzyme concentrations of the two kinases need to be equal for exact case 2 symmetry to be present in Mixed-Random 3 DSP. This is imposed as shown below.**

$$K2_{Total} := K1_{Total} :$$

**At this stage we introduce auxiliary constants $c_1$, $c_2$, and $c_3$ in place of the binding constants so as to make further analytical expressions more accessible.**

$$k_{b1} := c_1 \cdot \left(k_{ub1} + k_1\right) :$$
$$k_{b2} := c_2 \cdot \left(k_{ub2} + k_2\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the free enzymes ([K1], [K2] & [P]) and the partially modified substrate [A01]. In order to do this, we use the Maple command *solve* which solves the equation supplied for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this (using [A00K1] is given below in detail).**

**The differential equation of [A00K1] is given by,**

$$\frac{d\,[A00K1]}{dt} = dA00K1 \qquad \frac{d\,[A00K1]}{dt} = c_1 \left(k_1 + k_{ub1}\right) A00\,K1 - \left(k_1 + k_{ub1}\right) A00K1$$

**The *solve* command by Maple, solves this equation for the given variable (in this case [A00K1]). We in turn store this value (the solution returned by the *solve* command) in [A00K1]. This is performed by the following command.**

$$A00K1 := solve(dA00K1, A00K1)$$
$$A00K1 := K1\,A00\,c_1 \qquad \textbf{(1)}$$

**This operation is performed for the other complexes and substrate forms as well. Here we simultaneously solve expressions for the substrate forms using the *solve* command as shown below. The solution is stored in a variable labelled *Sol*, and then the respective solutions are**

**extracted from this vector using the *eval* command.**

```
A01K2 := solve(dA01K2, A01K2) :
A00K2 := solve(dA00K2, A00K2) :
A10K1 := solve(dA10K1, A10K1) :

Sol := solve({dA00, dA11, dA10}, {A00, A11, A10}) :

A00 := eval(A00, Sol) :
A01 := eval(A01, Sol) :
A10 := eval(A10, Sol) :
A11 := eval(A11, Sol) :
```

**Doing this results in the following correlations between the concentrations of the various substrate forms at steady state.**

$$A00 = \frac{A01\ K2\ c_2\ k_2}{K1\ c_1\ k_1}$$

$$A01 = A01$$

$$A10 = \frac{A01\ K2^2}{K1^2}$$

$$A11 = \frac{A01\ K2\ c_2\ k_2\ (K1 + K2)}{2\ K1\ k_3}$$

### Proof for invariant in asymmetric branches

**We know that K1Con and K2Con are both individually equal to zero always. Thus at a given steady state, K1Con - K2Con must also be equal to zero.**

**We thus introduce the term T = (K1Con - K2Con) = 0 and also introduce a new ratio, $\epsilon$ = [K1]/[K2]. Note: As discussed in the main text, the symmetric steady state is one where [K1] = [K2] or $\epsilon$ = 1. Since we are isolating solutions of asymmetry, we are primarily interested in solutions that permit, $\epsilon \neq 1$.**

```
T := K1Con - K2Con = 0 :
K1 := ϵ·K2 :
```

**The following command (*simplify*), simplifies the expression**

$$simplify(T) = \frac{K2\ \left(-\epsilon\ k_1 + A01\ c_2\ \left(k_1 - k_2\right)\right)\ (\epsilon - 1)}{k_1\ \epsilon} = 0$$

**From this we can ascertain that, should an asymmetric steady state exist (where $\epsilon \neq 1$) - the term $\left(-\epsilon\ k_1 + A01\ c_2\ \left(k_1 - k_2\right)\right)$ in the expression needs to necessarily be zero. This term is an expression in the partial substrate form [A01], $\epsilon$ and kinetic constants. Thus solving this to**

**isolate the partially modified substrate form in terms of the kinetic parameters we get the following. Here we use the *solve* command from Maple to solve T for [A01]. We use the *simplify* command to algebraically simplify the resulting expression.**

$A01 := simplify(solve(T, A01))$

$$A01 := \frac{\epsilon\, k_1}{(k_1 - k_2)\, c_2} \tag{1.1}$$

**Substituting this value for [A01] it back into the expression for the concentration of [A01], we get the following correlation.**

$$A00 = \frac{k_2}{(k_1 - k_2)\, c_1}$$

**We note that the concentration of [A00] is fixed in the asymmetric steady state and is given by a few key kinetic parameters. Using this information (the concentration of [A00] at an asymmetric steady state) - we solve for the concentration of free enzyme [K2] at this asymmetric steady state using the corresponding enzyme conservation equation (K2Con). This operation is carried out using the Maple command *solve* as shown below.**

$K2 := simplify(solve(K2Con, K2))$ :

$$K2 = \frac{K1_{Total}\,(k_1 - k_2)}{k_1\,(\epsilon + 1)}$$

**Using these expressions, we find that the concentration of [A11] is also fixed and given by only a few key kinetic constants and total enzyme concentrations as shown below.**

$$simplify(A11) = \frac{K1_{Total}\, k_2}{2\, k_3}$$

## Necessary conditions

**Since substrate concentration are always necessarily positive, the expression for [A00] and [A11] above must be positive. The numerators are only a function of kinetic parameters which are always positive, thus the denominator must be positive to ensure that the resulting concentration is positive. This gives us the necessary condition for an asymmetric state to exist as follows.**

1. $k_2 < k_1$

## Sufficiency of necessary conditions

**In this section of the proof we show the necessary conditions shown above are also sufficient**

**for an asymmetric steady state to exist for some positive $A_{Total}$ value. i.e. We show that upon a bifurcation along $A_{Total}$ we are bound to encounter symmetry breaking provided the necessary conditions are satisfied. Note that a feasible steady state in this context is one in which the concentrations of all substrates, complexes and enzymes are positive.**

**We do this by showing that the asymmetric states defined by the invariant concentration of [A00] & [A11] described above is indeed a feasible solution for the system of ODEs at some positive $A_{Total}$ value.**

**In an asymmetric steady state, as seen above the concentration of [A00] and [A11] are fixed by a few kinetic constants and total enzyme concentrations,**

$$A00 := \frac{k_2}{(k_1 - k_2)\, c_1} = \frac{k_2}{(k_1 - k_2)\, c_1}$$
$$A11 := \frac{K1_{Total}\, k_2}{2\, k_3} = \frac{1}{2}\,\frac{K1_{Total}\, k_2}{k_3}$$

**The other variables in this asymmetric state are thus given by**

$$A00 = \frac{k_2}{(k_1 - k_2)\, c_1}$$
$$A11 = \frac{K1_{Total}\, k_2}{2\, k_3}$$
$$A00K1 = \frac{\epsilon\, K1_{Total}\, k_2}{k_1\,(\epsilon + 1)}$$
$$A00K2 = \frac{K1_{Total}\, k_2}{k_1\,(\epsilon + 1)}$$
$$A01K2 = \frac{K1_{Total}\,\epsilon}{\epsilon + 1}$$
$$A10K1 = \frac{K1_{Total}}{\epsilon + 1}$$
$$simplify(A11P_1) = A11P_1$$
$$simplify(A11P_2) = A11P_2$$
$$simplify(A01P) = A01P$$
$$simplify(A10P) = A10P$$

**The system of ODE is also satisfied at this point, as is verified below.**

$$simplify(dA00) = 0$$
$$simplify(dA11) = 0$$
$$simplify(dA01) = 0$$
$$simplify(dA10) = 0$$

$simplify(dA00K1) = 0$
$simplify(dA00K2) = 0$
$simplify(dA01K2) = 0$
$simplify(dA10K1) = 0$
$simplify(dK1) = 0$
$simplify(dK2) = 0$

**Hence all that remains to be shown is that the variables (As described above) are positive for some value of $A_{Total}$.**
**This is true if and only if**

1. **Necessary condition $(k_1 > k_2)$ is satisfied**
2. **$\epsilon$ is positive.**

**However if $\epsilon$ is positive all concentrations are automatically positive (provided necessary conditions are satisfied).**

**Thus this means that for every non singular value of $\epsilon$ all concentrations are positive and the conservation of kinase and phosphatase is also satisfied.**

**Since the concentrations are all positive there exists a unique $A_{Total}$ value for every $\epsilon$ (permitting asymmetric states).**

**Hence we have proved that symmetry breaking is guaranteed for some finite positive $A_{Total}$, provided the necessary conditions above are satisfied - making those conditions sufficient for the behavior.**

## Prediction of pitchfork bifurcation along $A_{Total}$

**Here we predict the value of $A_{Total}$ at which symmetry breaking occurs via a pitch fork bifurcation. This point in the bifurcation is characterized by the intersection of both the symmetric steady state branch and the asymmetric steady state branches.**

**Hence at the position of symmetry breaking, we know two insights.**

**1. The system is still symmetric, hence [A01] = [A10] and [K1] = [K2].**
**2. The invariant describing the asymmetric steady state is also true.**

**Using these two information, we can simplify the original system considerably as follows.**

$K1 := K2 : \epsilon := 1 :$
$A00 := \frac{k_2}{(k_1 - k_2)\, c_1} :$
$A11 := \frac{K1_{Total}\, k_2}{2\, k_3} :$

**Now, by solving the conservation expression for the substrate we can isolate the value of $A_{Total}$ when the asymmetric steady states and the symmetric steady state intersect (indicating the pitchfork bifurcation point)**

$$A_{Total} = simplify\left(solve\left(simplify(ACon), A_{Total}\right)\right)$$

$$A_{Total} = \frac{1}{2\left(k_1 - k_2\right) c_1 c_2 k_3 k_1}\left(c_1\left(KI_{Total}\left(k_2 + 2k_3\right) c_2 + 4k_3\right) k_1^{\,2} - \left(KI_{Total} k_2 c_1 - 2k_3\right) k_2 c_2 k_1 - 2\,KI_{Total} k_2^{\,2} c_1 c_2 k_3\right) \quad \textbf{(4.1)}$$

**A cross verification of this analytical work is carried out in the read me file for the parameters used in generating the figures (Appendix 2 Fig 1).**

# Absolute (exact) concentration robustness
# Ordered DSP - Common kinase common phosphatase

In this maple file, we analytically show three key results pertaining to the presence of (exact) ACR in the ordered DSP system with common kinase common phosphotase effecting phosphorylation and dephosphorylation respectively (without any imposition of symmetry at the kinetic level). Namely,

1. We show that among the substrates, only the partially modified form (Ap) can show ACR in the system irrespective of kinetic parameters; i.e. The compeltely modified or completely unmodified substrate form is incapable of exhibiting ACR. This ACR is again only possible with changing total amounts of substrate concentration ($A_{Total}$).
2. The presence of ACR is (if present) guaranteed on two distinct branches of steady states (for any given $A_{Total}$) value.
3. If an ACR branch is to be present for some $A_{Total}$, then there will always be a non-ACR branch on which the ratio of the free kinase to free phosphatase is exactly constant irrespective of total substrate concentrations.

The ACR considered here is defined here as follows -- "If the substrate form is robust (maintained at an exact concentration) to increasing (or changing) total concentration of either the substrate or the enzymes ($A_{Total}$/$K_{Total}$/$P_{Total}$), that branch of steady state is termed to be an ACR branch, and the substrate form that shows robustness is said to exhibit ACR"

Put together these results provide the insight that granular symmetry is not strictly required in the network for ACR to be present - and other constratints involving kinetic constants and total concentrations of enzymes can guarentee ACR. Further discussion on the relevance of these results is provided in the appendix of the manuscript.

This section of the maple document pertaining to these results is structured with a common text introducing the model and initializing the steady state calculations - followed by different sub-sections each containing the proof to a specific self contained result. When running the code - please take notice to only run one sub-section at a time.

**We begin by initializing the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)**

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

**The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text and supplementary figure (refer to Appendix 2 Figure 10). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.**

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot ApK + k_{ub4} \cdot ApP - k_{b2} \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - \left(k_{ub3} + k_3\right) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dP := -k_{b3} \cdot App \cdot P + \left(k_{ub3} + k_3\right) \cdot AppP - k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

**The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient behavior and at steady state).**

$$ACon := A_{Total} - A - Ap - App - AK - ApK - AppP - ApP :$$
$$PCon := P_{Total} - P - AppP - ApP :$$
$$KCon := K_{Total} - K - AK - ApK :$$

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) :$$
$$k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) :$$
$$k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

**Once this is done, we solve for the steady state of the system in terms of fewer key variables. In this context we want to solve all variables in terms of the concentrations of the free enzymes ([K] & [P]) and the concentration of the partially modified substrate ([Ap]). In order to do this, we use the Maple command *solve,* which solves the supplied equation for a given variable. We first solve for the individual complexes using their corresponding differential equation. An example of this**

**(using [AK]) is given below in detail.**

$$AK := solve(dAK, AK) = K\,A\,c_1$$

**This operation is repeated for the other complexes forms as well.**

$$ApK := solve(dApK, ApK):$$
$$AppP := solve(dAppP, AppP):$$
$$ApP := solve(dApP, ApP):$$

**Here we introduce a new ratio $\epsilon$ which is equal to the concentration of free (unbound) kinase to the free (unbound) phosphotase.**

$$K := \epsilon \cdot P:$$

**We know that PCon is always equal to zero. Here we solve for the steady state concentration of the free phosphatase and in doing so we obtain an expression for it in terms of the ratio epsilion and the substrate forms.**

$$P := solve(PCon, P) = \frac{P_{Total}}{Ap\,c_4 + App\,c_3 + 1}$$

**As stated earlier - different sub-sections pertaining to specific proofs follow from here. Please take notice to run one sub-section at a time.**

## Proof of inability of the ACR in any substrate with changing $K_{Total}$ (or $P_{Total}$)

In this sub-section we show that no substrate is capable of exhibiting ACR for changing enzyme concentrations. In order to do this - we systematically show that the completely unmodified, partially modified and completely modified substrate forms are incapable of exhibiting ACR with increasing or changing $K_{Total}$.

### A

**We begin by solving the remaining equations to obtain expressions for substrates App, Ap at steady state in terms of concentrations of A and the ratio $\epsilon$. We do this by using the inbuilt solve command in *Maple*.**

$$Sol := solve(\{dAp, dApp\}, \{Ap, App\}) = \left\{Ap = \frac{\epsilon\,A\,c_1\,k_1}{c_4\,k_4}, App = \frac{\epsilon^2\,A\,c_1\,k_1\,c_2\,k_2}{c_4\,k_4\,c_3\,k_3}\right\}$$

$$assign(Sol)$$

**Thus at steady state - analyzing the total conservation of the substrate using these expressions leads to (we have used the *numer* and *collect* command to isolate the numerator and reorganize the resulting $A_{Total}$ expression respectively),**

$$\begin{aligned}&collect(numer(-ACon),\epsilon)\\&A^2 c_1^2 c_2^2 k_1^2 k_2^2 \epsilon^4 + \left(A^2 c_1^2 c_2 c_3 k_1^2 k_2 k_3 + A^2 c_1^2 c_2 c_4 k_1^2 k_2 k_3\right)\epsilon^3\\&\quad + \left(A^2 c_1^2 c_3 c_4 k_1^2 k_3^2 + A^2 c_1 c_2 c_3 c_4 k_1 k_2 k_3 k_4 - A c_1 c_2 c_3 c_4 k_1 k_2 k_3 k_4 A_{Total}\right.\\&\quad + A c_1 c_2 c_3 c_4 k_1 k_2 k_3 k_4 P_{Total} + A c_1 c_2 c_3 c_4 k_1 k_3^2 k_4 P_{Total}\\&\quad \left. + A c_1 c_2 c_4 k_1 k_2 k_3 k_4\right)\epsilon^2 + \left(A^2 c_1 c_3 c_4^2 k_1 k_3^2 k_4 - A c_1 c_3 c_4^2 k_1 k_3^2 k_4 A_{Total}\right.\\&\quad + A c_1 c_3 c_4^2 k_1 k_3^2 k_4 P_{Total} + A c_1 c_3 c_4^2 k_3^2 k_4^2 P_{Total} + A c_1 c_3 c_4 k_1 k_3^2 k_4\Big)\epsilon\\&\quad + A c_3 c_4^2 k_3^2 k_4^2 - c_3 c_4^2 k_3^2 k_4^2 A_{Total}\end{aligned} \tag{1.1.1}$$

**Here we proceed by a proof of contradiction - if we assume that A exhibits ACR for a range of $K_{Total}$ values - then this expression simplifies to simply a single variate polynomial in $\epsilon$ (roots of which represents steady state solutions of the system). The key insight here is that since this expression does not contain $K_{Total}$, changing $K_{Total}$ values would not affect the roots of this polynomial, implying $\epsilon$ can only take certain values at steady states with ACR. This is a contradiction as the expression for $K_{Total}$ cannot accomodate a non-changing A and $\epsilon$ with changing $K_{Total}$, as shown below.**

$$\begin{aligned}&collect(numer(KCon),\epsilon)\\&c_4 k_4 \left(A c_1 c_2 k_1 k_2 K_{Total} - A c_1 c_2 k_1 k_3 P_{Total}\right)\epsilon^2 + c_4 k_4 \left(A c_1 c_4 k_1 k_3 K_{Total}\right.\\&\quad \left. - A c_1 c_4 k_3 k_4 P_{Total} - c_4 k_3 k_4 P_{Total}\right)\epsilon + c_4^2 k_4^2 k_3 K_{Total}\end{aligned} \tag{1.1.2}$$

**i.e. By the assumption of ACR in A, it is fixed at a given value (by definition of ACR) for a range of $K_{Total}$ values. Thus in the resulting polynomial from KCon, $\epsilon$ (feasible roots of the resulting polynomial in $\epsilon$) will have to necessarily change as $K_{Total}$ changes. However equation ACon above, which was independent of $K_{Total}$, has fixed roots of $\epsilon$ which are consequently independent of $K_{Total}$. As stated earlier the common feasible root of these two equations would define steady state concentrations of the system. Since the roots in $\epsilon$ change in equation KCon with changing $K_{Total}$ and remain fixed in equation ACon - we have a contradiction. Thus A cannot be fixed exactly for a range of $K_{Total}$.**

**Thus A cannot exhibit ACR.**

## Ap

**We begin by solving the remaining equations to obtain expressions for substrates A, App at steady state in terms of concentrations of Ap and the ratio $\epsilon$. We do this by using the inbuilt**

***solve* command in *Maple*.**

$$Sol := solve(\{dA, dApp\}, \{A, App\}) = \left\{A = \frac{Ap\, c_4\, k_4}{c_1\, \epsilon\, k_1}, App = \frac{\epsilon\, Ap\, c_2\, k_2}{c_3\, k_3}\right\}$$

$$assign(Sol)$$

**Thus at steady state - analyzing the total conservation of the substrate using these expressions leads to (we have used the *numer* and *collect* command to isolate the numerator and reorganize the resulting $A_{Total}$ expression respectively),**

$$collect(numer(-ACon), \epsilon)$$

$$\begin{aligned} &Ap^2\, c_1\, c_2^{\,2}\, \epsilon^3\, k_1\, k_2^{\,2} + \left(Ap^2\, c_1\, c_2\, c_3\, k_1\, k_2\, k_3 + Ap^2\, c_1\, c_2\, c_4\, k_1\, k_2\, k_3\right. \\ &\quad - Ap\, c_1\, c_2\, c_3\, k_1\, k_2\, k_3\, A_{Total} + Ap\, c_1\, c_2\, c_3\, k_1\, k_2\, k_3\, P_{Total} + Ap\, c_1\, c_2\, c_3\, k_1\, k_3^{\,2}\, P_{Total} \\ &\quad \left. + Ap\, c_1\, c_2\, k_1\, k_2\, k_3\right)\epsilon^2 + \left(Ap^2\, c_1\, c_3\, c_4\, k_1\, k_3^{\,2} + Ap^2\, c_2\, c_3\, c_4\, k_2\, k_3\, k_4\right. \\ &\quad - Ap\, c_1\, c_3\, c_4\, k_1\, k_3^{\,2}\, A_{Total} + Ap\, c_1\, c_3\, c_4\, k_1\, k_3^{\,2}\, P_{Total} + Ap\, c_1\, c_3\, c_4\, k_3^{\,2}\, k_4\, P_{Total} \\ &\quad \left. + Ap\, c_1\, c_3\, k_1\, k_3^{\,2} - c_1\, c_3\, k_1\, k_3^{\,2}\, A_{Total}\right)\epsilon + Ap^2\, c_3\, c_4^{\,2}\, k_3^{\,2}\, k_4 + Ap\, c_3\, c_4\, k_3^{\,2}\, k_4 \end{aligned} \tag{1.2.1}$$

**Here we proceed by contradiction - if we assume that Ap exhibits ACR for a range of $K_{Total}$ values - then this expression simplifies to simply a single variate polynomial in $\epsilon$ (roots of which represents steady state solutions of the system). The key insight here is that since this expression does not contain $K_{Total}$, changing $K_{Total}$ values would not affect the roots of this expresssion, implying $\epsilon$ can only take certain values at steady states with ACR. This is a contradiction as the expression for $K_{Total}$ cannot accomodate a non-changing Ap and $\epsilon$ with changing $K_{Total}$, as shown below**

$$collect(numer(KCon), \epsilon)$$

$$\begin{aligned} &\left(Ap\, c_2\, k_1\, k_2\, K_{Total} - Ap\, c_2\, k_1\, k_3\, P_{Total} - k_1\, k_3\, P_{Total}\right)\epsilon + Ap\, c_4\, k_1\, k_3\, K_{Total} \\ &\quad - P_{Total}\, k_3\, Ap\, c_4\, k_4 + k_1\, k_3\, K_{Total} \end{aligned} \tag{1.2.2}$$

**i.e. By the assumption of ACR in Ap, it is fixed at a given value (by definition of ACR) for a range of $K_{Total}$ values. Thus in the resulting polynomial from KCon, $\epsilon$ (feasible roots of the resulting polynomial in $\epsilon$) will have to necessarily change as $K_{Total}$ changes. However equation ACon above, which was independent of $K_{Total}$, has fixed roots of $\epsilon$ which are consequently independent of $K_{Total}$. As stated earlier the common feasible root of these two equations would define steady state concentrations of the system. Since the roots in $\epsilon$ change in equation KCon with changing $K_{Total}$ and remain fixed in equation ACon - we have a contradiction. Thus Ap cannot be fixed exactly for a range of $K_{Total}$.**

**Thus Ap cannot exhibit ACR.**

## App

**We begin by solving the remaining equations to obtain expressions for substrates A, Ap at steady state in terms of concentrations of App and the ratio $\epsilon$. We do this by using the inbuilt solve command in *Maple*.**

$$Sol := solve(\{dA, dAp\}, \{A, Ap\}) = \left\{A = \frac{App\, c_3\, c_4\, k_3\, k_4}{c_1\, c_2\, \epsilon^2\, k_1\, k_2}, Ap = \frac{App\, c_3\, k_3}{c_2\, \epsilon\, k_2}\right\}$$

$$assign(Sol)$$

**Thus at steady state - analyzing the total conservation of the substrate using these expressions leads to (we have used the *numer* and *collect* command to isolate the numerator and reorganize the resulting $A_{Total}$ expression respectively),**

$$collect(numer(-ACon), \epsilon)$$

$$\begin{aligned}&\left(App^2\, c_1\, c_2^{\,2}\, c_3\, k_1\, k_2^{\,2} - App\, c_1\, c_2^{\,2}\, c_3\, k_1\, k_2^{\,2}\, A_{Total} + App\, c_1\, c_2^{\,2}\, c_3\, k_1\, k_2^{\,2}\, P_{Total}\right. \\ &\quad \left. + App\, c_1\, c_2^{\,2}\, c_3\, k_1\, k_2\, k_3\, P_{Total} + App\, c_1\, c_2^{\,2}\, k_1\, k_2^{\,2} - c_1\, c_2^{\,2}\, k_1\, k_2^{\,2}\, A_{Total}\right) \epsilon^3 \\ &+ \left(App^2\, c_1\, c_2\, c_3^{\,2}\, k_1\, k_2\, k_3 + App^2\, c_1\, c_2\, c_3\, c_4\, k_1\, k_2\, k_3 - App\, c_1\, c_2\, c_3\, c_4\, k_1\, k_2\, k_3\, A_{Total}\right. \\ &\quad + App\, c_1\, c_2\, c_3\, c_4\, k_1\, k_2\, k_3\, P_{Total} + App\, c_1\, c_2\, c_3\, c_4\, k_2\, k_3\, k_4\, P_{Total} \\ &\quad \left. + App\, c_1\, c_2\, c_3\, k_1\, k_2\, k_3\right) \epsilon^2 + \left(App^2\, c_1\, c_3^{\,2}\, c_4\, k_1\, k_3^{\,2} + App^2\, c_2\, c_3^{\,2}\, c_4\, k_2\, k_3\, k_4\right. \\ &\quad \left. + App\, c_2\, c_3\, c_4\, k_2\, k_3\, k_4\right) \epsilon + App^2\, c_3^{\,2}\, c_4^{\,2}\, k_3^{\,2}\, k_4\end{aligned} \tag{1.3.1}$$

**Here we proceed by contradiction - if we assume that App exhibits ACR for a range of $K_{Total}$ values - then this expression simplifies to simply a single variate cubic polynomial in $\epsilon$ (roots of which represents steady state solutions of the system). The key insight here is that since this expression does not contain $K_{Total}$, changing $K_{Total}$ values would not affect the roots of this expresssion, implying $\epsilon$ can only take certain values at steady states with ACR. This is a contradiction as the expression for $K_{Total}$ cannot accomodate a non-changing App and $\epsilon$ with changing $K_{Total}$, as shown below**

$$collect(numer(KCon), \epsilon)$$

$$\begin{aligned}&-\epsilon^3\, P_{Total}\, c_2^{\,2}\, k_2^{\,2}\, k_1 + c_2\, k_2\, \left(App\, c_2\, c_3\, k_1\, k_2\, K_{Total} - App\, c_2\, c_3\, k_1\, k_3\, P_{Total}\right. \\ &\quad \left. + c_2\, k_1\, k_2\, K_{Total}\right) \epsilon^2 + c_2\, k_2\, \left(App\, c_3\, c_4\, k_1\, k_3\, K_{Total} - P_{Total}\, App\, c_3\, c_4\, k_3\, k_4\right) \epsilon\end{aligned} \tag{1.3.2}$$

**i.e. By the assumption of ACR in App, it is fixed at a given value (by definition of ACR) for a range of $K_{Total}$ values. Thus in the resulting polynomial from KCon, $\epsilon$ (feasible roots of the resulting polynomial in $\epsilon$) will have to necessarily change as $K_{Total}$ changes. However equation ACon above, which was independent of $K_{Total}$, has fixed roots of $\epsilon$ which are consequently independent of $K_{Total}$. As stated earlier the common feasible root of these two**

**equations would define steady state concentrations of the system. Since the roots in $\epsilon$ change in equation KCon with changing $K_{Total}$ and remain fixed in equation ACon - we have a contradiction. Thus App cannot be fixed exactly for a range of $K_{Total}$.**

**Thus App cannot exhibit ACR.**

Due to the structure of the DSP network, $K_{Total}$ (App) and $P_{Total}$ (A) are behaviorly equivalent (acheivable by a simple transformation of variables) and thus by extension a similar proof (omitted here) will rule out the presence of ACR in any of the substrates with increasing $P_{Total}$.

## Absence of ACR in the fully modified (and unmodified) substrate with increasing $A_{Total}$

In this sub-section we show that the fully modified and the fully unmodified substrate forms (A and App) are capable of exhibiting ACR with changing total substrate concentrations.

**We begin by solving the remaining equations to obtain expressions for substrates A, Ap at steady state in terms of concentrations of App and the ratio $\epsilon$. We do this by using the inbuilt solve command in *Maple*.**

$$Sol := solve(\{dA, dAp\}, \{A, Ap\}) =$$

$$Sol := \left\{A = \frac{App\, c_3\, c_4\, k_3\, k_4}{c_1\, c_2\, \epsilon^2\, k_1\, k_2}, Ap = \frac{App\, c_3\, k_3}{c_2\, \epsilon\, k_2}\right\} \quad \textbf{(2.1)}$$

$$assign(Sol)$$

**Thus at steady state - analyzing the total conservation of the substrate using these expressions leads to (we have used the *numer* and *collect* command to isolate the numerator and reorganize the resulting $A_{Total}$ expression respectively),**

$$collect(numer(KCon), \epsilon)$$

$$-\epsilon^3\, P_{Total}\, c_2^2\, k_2^2\, k_1 + c_2\, k_2 \left(App\, c_2\, c_3\, k_1\, k_2\, K_{Total} - App\, c_2\, c_3\, k_1\, k_3\, P_{Total} + c_2\, k_1\, k_2\, K_{Total}\right) \epsilon^2 + c_2\, k_2 \left(App\, c_3\, c_4\, k_1\, k_3\, K_{Total} - P_{Total}\, App\, c_3\, c_4\, k_3\, k_4\right) \epsilon \quad \textbf{(2.2)}$$

**Similar to the earlier proof, we proceed with a proof by contradiction. If we assume that App presents with ACR for a range of $A_{Total}$ values. Then the above expression simplifies to a single variate polynomial in $\epsilon$ (the roots of which represents steady state solutions of the system). The key insight here is that since this expression does not contain $A_{Total}$, changing $A_{Total}$ values would not affect the roots of this expresssion, implying $\epsilon$ can only take certain values at steady states with ACR. This is a contradiction as the expression for $A_{Total}$ cannot accomodate a non-**

**changing App and $\epsilon$ with changing $A_{Total}$.**

$collect(numer(ACon), \epsilon)$

$$\left(-App^2 c_1 c_2^2 c_3 k_1 k_2^2 + App c_1 c_2^2 c_3 k_1 k_2^2 A_{Total} - App c_1 c_2^2 c_3 k_1 k_2^2 P_{Total} - App c_1 c_2^2 c_3 k_1 k_2 k_3 P_{Total} - App c_1 c_2^2 k_1 k_2^2 + c_1 c_2^2 k_1 k_2^2 A_{Total}\right) \epsilon^3 + \left(-App^2 c_1 c_2 c_3^2 k_1 k_2 k_3 - App^2 c_1 c_2 c_3 c_4 k_1 k_2 k_3 + App c_1 c_2 c_3 c_4 k_1 k_2 k_3 A_{Total} - App c_1 c_2 c_3 c_4 k_1 k_2 k_3 P_{Total} - App c_1 c_2 c_3 c_4 k_2 k_3 k_4 P_{Total} - App c_1 c_2 c_3 k_1 k_2 k_3\right) \epsilon^2 + \left(-App^2 c_1 c_3^2 c_4 k_1 k_3^2 - App^2 c_2 c_3^2 c_4 k_2 k_3 k_4 - App c_2 c_3 c_4 k_2 k_3 k_4\right) \epsilon - App^2 c_3^2 c_4^2 k_3^2 k_4 \quad (2.3)$$

**Thus App cannot exhibit ACR.**

Due to the structure of the DSP network A and App are topologically equivalent with respect to variation in $A_{Total}$ concentrations. Thus by a similar proof (omitted here), A is incapable of exhibiting ACR with changing $A_{Total}$.

## Presence and features of ACR in the partially modified substrate form with increasing $A_{Total}$

In the earlier sub-sections we have shown how ACR is not possible in A and App with changing $A_{Total}$. In this sub-section we show that the remaining substrate form Ap is capable of exhibiting ACR with changing $A_{Total}$ and also elucidate the associated features of such the network and it's steady states when it accomodates ACR.

**We begin by solving the remaining equations to obtain expressions for substrates A, App at steady state in terms of concentrations of App and the ratio $\epsilon$. We do this by using the inbuilt solve command in *Maple*.**

$$Sol := solve(\{dA, dApp\}, \{A, App\}) = \left\{A = \frac{Ap\, c_4\, k_4}{c_1\, \epsilon\, k_1}, App = \frac{\epsilon\, Ap\, c_2\, k_2}{c_3\, k_3}\right\}$$

$assign(Sol)$

**Thus at steady state - analyzing the total conservation of the substrate using these expressions leads to (we have used the *numer* and *collect* command to isolate the numerator and reorganize the resulting $A_{Total}$ expression respectively),**

$collect(numer(KCon), \epsilon)$

$$\left(Ap\,c_2\,k_1\,k_2\,K_{Total} - Ap\,c_2\,k_1\,k_3\,P_{Total} - k_1\,k_3\,P_{Total}\right)\epsilon + Ap\,c_4\,k_1\,k_3\,K_{Total} - P_{Total}\,k_3\,Ap\,c_4\,k_4 + k_1\,k_3\,K_{Total} \tag{3.1}$$

**Now if Ap were to exhibit ACR with changing $A_{Total}$, the above expression which is independent of $A_{Total}$, should be satisfied for differing values of the ratio $\epsilon$. This independence in $\epsilon$ can only be guarenteed when the coeffecient of $\epsilon$ and the constant term are both zero.**

**This condition will provide a strict requirement on the kinetics (going beyond the granular symmetry assumed in case 1 symmetry earlier) and will also establish the concentration of Ap as functions of the kinetics of the network on the ACR branch as follows.**

$$Ap = solve\left(\left(Ap\,c_2\,k_1\,k_2\,K_{Total} - Ap\,c_2\,k_1\,k_3\,P_{Total} - k_1\,k_3\,P_{Total}\right), Ap\right) =$$

$$Ap = \frac{k_3\,P_{Total}}{c_2\left(k_2\,K_{Total} - k_3\,P_{Total}\right)}$$

$$Ap = solve\left(Ap\,c_4\,k_1\,k_3\,K_{Total} - P_{Total}\,k_3\,Ap\,c_4\,k_4 + k_1\,k_3\,K_{Total}, Ap\right) =$$

$$Ap = -\frac{k_1\,K_{Total}}{c_4\left(K_{Total}\,k_1 - k_4\,P_{Total}\right)}$$

**The two expressions for Ap obtained above (as functions of kinetics and total enzyme amounts) must be strictly equal and positive (since Ap is strictly positive).**

**Note: We quickly note an insight here that with case 1 symmetry in the kinetics, these two expressions are equivalent, however without the granular symmetry of case 1, these two expressions can still be satisfied, by suitable choice of kinetic constants and total enzyme concentrations.**

**Taking this in to account - ACon = 0 (the conservation expression for the total substrate concentration) similarly simplifies to**

$$collect(numer(ACon), \epsilon)$$

$$-Ap^2\,c_1\,c_2^2\,k_1\,k_2^2\,\epsilon^3 + \left(-Ap^2\,c_1\,c_2\,c_3\,k_1\,k_2\,k_3 - Ap^2\,c_1\,c_2\,c_4\,k_1\,k_2\,k_3 + Ap\,c_1\,c_2\,c_3\,k_1\,k_2\,k_3\,A_{Total} - Ap\,c_1\,c_2\,c_3\,k_1\,k_2\,k_3\,P_{Total} - Ap\,c_1\,c_2\,c_3\,k_1\,k_3^2\,P_{Total} - Ap\,c_1\,c_2\,k_1\,k_2\,k_3\right)\epsilon^2 + \left(-Ap^2\,c_1\,c_3\,c_4\,k_1\,k_3^2 - Ap^2\,c_2\,c_3\,c_4\,k_2\,k_3\,k_4 + Ap\,c_1\,c_3\,c_4\,k_1\,k_3^2\,A_{Total} - Ap\,c_1\,c_3\,c_4\,k_1\,k_3^2\,P_{Total} - Ap\,c_1\,c_3\,c_4\,k_3^2\,k_4\,P_{Total} - Ap\,c_1\,c_3\,k_1\,k_3^2 + c_1\,c_3\,k_1\,k_3^2\,A_{Total}\right)\epsilon - Ap^2\,c_3\,c_4^2\,k_3^2\,k_4 - Ap\,c_3\,c_4\,k_3^2\,k_4 \tag{3.2}$$

**Thus at a fixed value for Ap (deteremined from the equations earlier) - the above equation becomes a single variate polynomial in $\epsilon$.**

**We make two observations here,**

**1. The above polynomial in $\epsilon$ has necessarily a negative real root (ascertained by noting the signs of the leading coeffecient and the constant) for positive kinetic constants.**
**2. The product of roots of the polynomial is negative.**

**Thus, if there exists an $\epsilon$ positive root, signifying a distinct ACR steady state, satisfying the polynomial there will always exist another positive real root root signifying the presence of another ACR steady state. Thus if ACR is to be present - it is necessarily present on two such branches for any given $A_{Total}$.**

**"Hence we draw our first conclusion regarding the ACR steady states - that should there exist an ACR steady state for Ap, there are necesarily two such branches with changing $A_{Total}$"**

## Sufficiency of necessary conditions

**Here, before proceeding to ascertain further features of the system while exhibiting ACR, we show that the necessary kinetic conditions obtained earlier are sufficient to obtain ACR in the system at some finite positive $A_{Total}$**

**We make two further observations at this stage,**

**1. That at sufficiently large ATotal concentrations, the sum of the roots of the polynomial ( in $\epsilon$, ACon = 0) earlier is postive.**
**2. The discriminant of the cubic polynomial is also positive at some finite but sufficiently large $A_{Total}$ as shown below**

$$coeff\left(simplify(discrim(numer(ACon),\epsilon)),A_{Total},4\right)$$

$$c_3^{\,4}\,k_1^{\,4}\,c_1^{\,4}\left(-Ap\,c_4-1\right)^2 k_3^{\,6}\,k_2^{\,2}\,Ap^2\,c_2^{\,2} \qquad \textbf{(3.1.1)}$$

**The resulting polynomial for the discriminant is a quartic in $A_{Total}$. We isolate the leading coeffecient of this polynomial as shown above using the inbuilt Maple command *coeff*. Now since this expression is positive for all positive kinetic constants, the discriminant is itself positive at some finite but sufficiently large $A_{Total}$ concentration.**

**Thus the above two observations conclude that at some sufficiently large $A_{Total}$ the polynomial admits two $\epsilon$ roots, each denoting a steady state on an ACR branch.**

**We now prove the existence of an intersection of the non-ACR branch and ACR branch.**

**We proceed with a proof by contraditction. Assuming that no intersection exists - there is no**

**common $\epsilon$ root for both equation 3.1 and 3.2. We should note here that while Ap is not fixed at its ACR concentration on the non-ACR branch (by definition) the kinetic parameters should none the less satisfy the necessary constraint elucidated above for the presence of ACR in the network; i.e.,**

$$\frac{k_3 P_{Total}}{c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)} = -\frac{k_1 K_{Total}}{c_4 \left(K_{Total} k_1 - k_4 P_{Total}\right)}$$

**Inorder to establish this equivalency (in kinetic constants and total concentrations of enzymes), we solve for k1 from this equation. This is an artibitrary choice and this expression can be satisfied by solving for any of the other constants involved.**

$$k_1 = simplify\left(solve\left(\frac{k_3 P_{Total}}{c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)} = -\frac{k_1 K_{Total}}{c_4 \left(K_{Total} k_1 - k_4 P_{Total}\right)}, k_1\right)\right)$$

$$k_1 := \frac{c_4 k_3 k_4 P_{Total}^{\,2}}{\left(-k_3 \left(c_2 - c_4\right) P_{Total} + c_2 k_2 K_{Total}\right) K_{Total}} \qquad \textbf{(3.3)}$$

**With this algebraic manipulation, equation 3.1 (KCon=0) from earlier can be rewritten as shown below,**

$$simplify\left(numer\left(eval\left(KCon, k_1 = simplify\left(solve\left(\frac{k_3 P_{Total}}{c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right)} = -\frac{k_1 K_{Total}}{c_4 \left(K_{Total} k_1 - k_4 P_{Total}\right)}, k_1\right)\right)\right)\right)\right)$$

$$k_3 \left(\left(-Ap\, c_2 - 1\right) k_3 P_{Total} + K_{Total}\, Ap\, c_2 k_2\right) \left(\epsilon P_{Total} - K_{Total}\right) \qquad \textbf{(3.4)}$$

**We can see the birth and characteristic of the ACR branches more clearly now. There are two possible ways in which this conservation can be satisfied for any steady state. Either we are on an ACR branch - in which case the expression $\left(-Ap\, c_2 - 1\right) k_3 P_{Total} + Ap\, c_2 k_2 K_{Total}$ is equal to 0 (as we have noted earlier for the ACR branch), or $\epsilon$ is fixed as $\frac{K_{Total}}{P_{Total}}$.**

**Thus if we isolate for just the non-ACR branch (where $\epsilon$ is fixed as $\frac{K_{Total}}{P_{Total}}$), we observe that the remaining conservation equation for the substrate simplifies to a univariate quadratic polynomial in Ap. In this instance we don't resubstitute the k1 as we did earlier - but allow it to be a variable as such for the sake of transparancy in the expression.**

$$T := collect\left(numer\left(eval\left(ACon, \epsilon = \frac{K_{Total}}{P_{Total}}\right)\right), Ap\right)$$

$$\begin{aligned} T := {} & \left(-c_1 c_2^2 k_1 k_2^2 K_{Total}^3 - c_1 c_2 c_3 k_1 k_2 k_3 K_{Total}^2 P_{Total}\right. \\ & - c_1 c_2 c_4 k_1 k_2 k_3 K_{Total}^2 P_{Total} - c_1 c_3 c_4 k_1 k_3^2 K_{Total} P_{Total}^2 \\ & \left. - c_2 c_3 c_4 k_2 k_3 k_4 K_{Total} P_{Total}^2 - c_3 c_4^2 k_3^2 k_4 P_{Total}^3\right) Ap^2 \\ & + \left(c_1 c_2 c_3 k_1 k_2 k_3 A_{Total} K_{Total}^2 P_{Total} - c_1 c_2 c_3 k_1 k_2 k_3 K_{Total}^2 P_{Total}^2\right. \\ & - c_1 c_2 c_3 k_1 k_3^2 K_{Total}^2 P_{Total}^2 + c_1 c_3 c_4 k_1 k_3^2 A_{Total} K_{Total} P_{Total}^2 \\ & - c_1 c_3 c_4 k_1 k_3^2 K_{Total} P_{Total}^3 - c_1 c_3 c_4 k_3^2 k_4 K_{Total} P_{Total}^3 \\ & \left. - c_1 c_2 k_1 k_2 k_3 K_{Total}^2 P_{Total} - c_1 c_3 k_1 k_3^2 K_{Total} P_{Total}^2 - c_3 c_4 k_3^2 k_4 P_{Total}^3\right) Ap \\ & + c_1 c_3 k_1 k_3^2 A_{Total} K_{Total} P_{Total}^2 \end{aligned} \tag{3.5}$$

**This quadratic polynomial has necessarily only one positive solution for Ap for every single value of $A_{Total}$ (This can be quickly verified by noting the sign of the coeffecients of the polynomial). The range of the feasible root of Ap, as $A_{Total}$ changes, would represent the range of Ap on the non-ACR branch. Allowing $A_{Total}$ to vary between 0 to +infinity, one can note that the range of Ap also continously changes from 0 to +infinity (monotonically since it's a quadratic polynomial).**

**At some point within this range - Ap reaches the ACR value (provided by the kinetic constants).**

**Thus this proves the existence of an intersection between an ACR branch and the non-ACR branch, and that the non-ACR branch in the system has the unique feature that the ratio of unbound kinase to phosphatase is exactly fixed by the ratio of the total amounts of kinase and phosphatase at all total substrate concentrations.**

# Enzymatic networks generate inbuilt safe guards against overexpression: A study in Erk

**Vaidhiswaran Ramesh[1], J Krishnan[1,2,*]**

**Affiliations:**
1 - Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, Imperial College London, London, SW7 2AZ, UK
2 - Institute for Systems and Synthetic Biology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

* - Corresponding author (j.krishnan@ic.ac.uk)

This supplementary material provides detailed model descriptions and proofs for the analytical arguments presented in the main text. The document is organized into many fice folders (and many sub-folders), as described below. A PDF version of the file is also provided for easy accessibility.

**Organization of the results**

Read me file (this file)
Folder 1: Simple enzymatic networks
Folder 2: Multisite modification (double site ordered phosphorylation model with common enzymes)
Folder 3: Multisite modification - enzyme combinations (double site ordered phosphorylation model with different enzyme combinations acting on each modification site)
Folder 4: Beyond multisite modification (coupled covalent modifications and cascaded enzymatic networks)
Folder 5: Biphasic responses in Erk Model

Each folder if presented with a Read_Me file that serves as a preamble to indicate the nature of results within and each file is accompanied with detailed instructions on how to run the code in Maple.

**Key insight relevant to biphasic dose response**

In the following documents, in order to establish the presence or absence of biphasic responses in steady state concentration of a given variable with a dose parameter, we use the mathematical definition of the behavior. This main tenet of the behavior forms an important part of the proofs that follow and is highlighted here.

If a system is capable of exhibiting biphasic response in a variable 'x' as a dose parameter 'p' is changed, then it requires that there exists a feasible steady state of the system satisfying the model description (the system of ODEs and the conservation equation), which also simultaneously satisfies the following mathematical expression

$$\frac{\partial x}{\partial p} = 0$$

This arises from the fact that, for a biphasic behavior to exist, there has to naturally exist a peak concentration of the variable 'x' where the gradient of the steady state curve is zero. The absence of such a point indicates the absence of biphasic response in the system.

An enzymatic network has two natural 'doses' from the perspective of the maximally modified substrate; the total substrate in the system and the total amount of kinase (the enzyme effecting the modifications). Enzymatic networks are capable of exhibiting biphasic responses in the modified form (or the maximally modified form) with change in both doses. Both of these biphasic dose responses are of interest as they belay simple intuitive expectations of the system to behave monotonically with change in dose. These different biphasic responses shown in the substrate are simply labelled as enzyme biphasic or substrate biphasic from here on within this text (refering to scenarios where the enzyme or the substrate is the dose, respectively).

## Simple enzymatic network motifs

In this folder we detail analytical results pertaining to simple enzymatic network motifs (covalent modification cycle and protein-protein interaction model), specifically their capacity to exhibit biphasic responses (in steady state dose response). The study of these networks and the results provide an emerging synthesis on the different kinds of biphasic responses that can be seen in signaling systems and the minimal ingredients required to generate the same.

# Covalent modification network

## [single modification/demotification of a protein]

In this file we prove that the simple covalent modification network with a kinase and a phosphatase effecting the modifications and demodification (phosphorylation and dephosphorylation) is incapable of presenting any biphasic dose response behavior in the modified substrate form. More specifically, we show the

1. absence of enzyme biphasic response in concentration of Ap with changing total kinase concentration
2. absence of substrate biphasic response in concentration of Ap with changing total substrate concentration

Detailed definitions of these two different classes of biphasic behavior are described in the manuscript (see Read_Me file on page 1). We reiterate that should a biphasic behavior exist in the steady state concentration of the modified substrate form (Ap) with respect to any parameter 'p' (total enzyme amount or the total substrate amount), then there exists steady state concentrations of the variables where the following mathematical expression is necessarily satisfied.

$$\frac{\partial Ap}{\partial p} = 0$$

This takes the form $\frac{\partial Ap}{\partial K_{Total}} = 0$ for enzyme biphasic and $\frac{\partial Ap}{\partial A_{Total}} = 0$ for substrate biphasic response.

In this file, we begin by describing the mathematical model for a covalent modification network. Then by solving the system of ODEs at steady state, we obtain correlations between the steady state concentrations of the different variables and using these we proceed to show the absence of both biphasic dose response behaviors (subsection 1 and 2).

**Covalent modification network:** We first describe the mathematical model of a covalent modification network.

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart : with(LinearAlgebra) : with(Student[LinearAlgebra]) : with(VectorCalculus) :
```

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_2 \cdot ApP - k_{b1} \cdot A \cdot K + k_{ub1} \cdot AK :$$
$$dAp := k_1 \cdot AK + k_{ub2} \cdot ApP - k_{b2} \cdot Ap \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApP := k_{b2} \cdot Ap \cdot P - \left(k_{ub2} + k_2\right) \cdot ApP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dP := -k_{b2} \cdot Ap \cdot P + \left(k_{ub2} + k_2\right) \cdot ApP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + ApP + AK - A_{Total} :$$
$$KCon := K + AK - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

Now we begin solving the system of equations to obtain correlations between the steady state concentrations of the variables. Primarily to obtain steady state correlations of variables as a function of concentrations of Ap and the free kinase concentration K. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$
```
assign(solve({dApP, dA}, {ApP, A})) :
```

Once this is done, we again solve for the steady state of phosphatase and the substrate involved in the covalent modification cycle using the conservation expression for the enzyme (PCon),

```
P := solve(PCon, P) :
```

Now having solved for the steady state of the system in terms of Ap and K, the only two equations that remain (which define the steady state of the system) are the conservation equations for the kinase and the substrate (KCon and PCon) respectively (see below).

```
ACon
```

$$\frac{Ap\,P_{Total}\,k_2\,k_{b2}\left(k_{ub1}+k_1\right)}{\left(k_{b2}\,Ap+k_2+k_{ub2}\right)K\,k_1\,k_{b1}}+Ap+\frac{k_{b2}\,Ap\,P_{Total}}{k_{b2}\,Ap+k_2+k_{ub2}}+\frac{Ap\,P_{Total}\,k_2\,k_{b2}}{\left(k_{b2}\,Ap+k_2+k_{ub2}\right)k_1}-A_{Total} \tag{1}$$

*KCon*

$$K+\frac{Ap\,P_{Total}\,k_2\,k_{b2}}{\left(k_{b2}\,Ap+k_2+k_{ub2}\right)k_1}-K_{Total} \tag{2}$$

## 1. Enzyme biphasic

Now in order to show the absence of enzyme biphasic response in Ap with total enzyme concentration, we proceed with a proof by contradiction. As mentioned earlier, for enzyme biphasic in Ap to exist, for some steady state of the system, $\frac{\partial Ap}{\partial K_{Total}}=0$ must be satisfied. Thus, we begin with the assumption that such a biphasic exists, satisfying the condition.

Now, differentiating ACon with respect to $K_{Total}$ provides the following

$$\frac{dACon}{dK_{Total}}=0=\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}+\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}}=0=\frac{dKCon}{\partial K_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}$$

With the assumption that there exists a biphasic behavior in the substrate with changing total substrate amounts, these equations simplify as shown below

$$0=\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$1=\frac{\partial K}{\partial K_{Total}}$$

***Note***

This reduction was possible since $\frac{\partial KCon}{\partial Ap}$ and $\frac{\partial ACon}{\partial Ap}$ have rational denominators (as shown below), allowing the product with $\frac{\partial ACon}{\partial Ap}$ to be zero.

$$simplify\left(diff\left(KCon,Ap\right)\right)=\frac{P_{Total}\,k_2\,k_{b2}\left(k_{ub2}+k_2\right)}{\left(k_{b2}\,Ap+k_2+k_{ub2}\right)^2\,k_1}$$

$$simplify\left(diff\left(ACon,Ap\right)\right)=$$

$$\frac{1}{\left(k_{b2}\,Ap+k_2+k_{ub2}\right)^2\,K\,k_1\,k_{b1}}\left(Ap^2\,K\,k_1\,k_{b1}\,{k_{b2}}^2+2\left(k_{ub2}+k_2\right)\left(K\left(\left(Ap+\frac{P_{Total}}{2}\right)k_1\right.\right.\right.$$

$$+ \frac{k_2 P_{Total}}{2}\Bigg) k_{b1} + \frac{P_{Total} k_2 \left(k_{ub1} + k_1\right)}{2}\Bigg) k_{b2} + K k_1 k_{b1} \left(k_{ub2} + k_2\right)^2\Bigg)$$

The above equation indicates that $\frac{\partial K}{\partial K_{Total}}$ is equal to 1. From this, we can observe that $\frac{\partial ACon}{\partial K}$ must be zero to satisfy the first expression (obtained from differentiating ACon). However as seen below, we can see that this is not possible for any feasible steady state of the system.

$$simplify(diff(ACon, K)) = -\frac{Ap\, P_{Total} k_2 k_{b2} \left(k_{ub1} + k_1\right)}{\left(k_{b2} Ap + k_2 + k_{ub2}\right) K^2 k_1 k_{b1}}$$

**This is a contradiction and thus enzyme biphasic behavior in Ap with total enzyme concentration ($K_{Total}$) is not possible.**

## 2. Substrate biphasic

Now in order to show the absence of substrate biphasic response in Ap with total substrate concentration, we proceed with a proof by contradiction. As mentioned earlier, for substrate biphasic in Ap to exist, for some steady state of the system, $\frac{\partial Ap}{\partial A_{Total}} = 0$ must be satisfied. Thus we begin with the assumption that such a biphasic exists, satisfying the condition.

This implies that differentiating ACon with respect to $A_{Total}$ provides the following

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}} + \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}}$$

$$\frac{dKCon}{dA_{Total}} = 0 = \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}}$$

With the assumption that there exists a biphasic behavior in the substrate with changing total substrate amounts, these equations simplify as shown below

$$1 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

$$0 = \frac{\partial K}{\partial A_{Total}}$$

***Note***

This reduction was possible since $\frac{\partial KCon}{\partial Ap}$ and $\frac{\partial ACon}{\partial Ap}$ have rational denominators (as shown below), allowing the product with $\frac{\partial ACon}{\partial Ap}$ to be zero.

$$simplify(diff(KCon, Ap)) = \frac{P_{Total} k_2 k_{b2} \left(k_{ub2} + k_2\right)}{\left(k_{b2} Ap + k_2 + k_{ub2}\right)^2 k_1}$$

$$simplify(diff(ACon, Ap)) = \frac{1}{\left(k_{b2}\,Ap + k_2 + k_{ub2}\right)^2 K\,k_1\,k_{b1}} \left( Ap^2\,K\,k_1\,k_{b1}\,{k_{b2}}^2 + 2\left( K\left( \left( Ap + \frac{P_{Total}}{2} \right) k_1 + \frac{k_2\,P_{Total}}{2} \right) k_{b1} + \frac{P_{Total}\,k_2\left(k_{ub1} + k_1\right)}{2} \right) \left(k_{ub2} + k_2\right) k_{b2} + K\,k_1\,k_{b1}\left(k_{ub2} + k_2\right)^2 \right)$$

The first expression indicates that $\frac{\partial K}{\partial A_{Total}}$ cannot be equal to 0 (since $\frac{\partial ACon}{\partial K}$ itself is always positive, as shown below). However the second expression (obtained through differentiating KCon) contradicts this.

$$simplify(diff(ACon, K)) = -\frac{Ap\,P_{Total}\,k_2\,k_{b2}\left(k_{ub1} + k_1\right)}{\left(k_{b2}\,Ap + k_2 + k_{ub2}\right) K^2\,k_1\,k_{b1}}$$

**This is a contradiction and thus substrate biphasic behavior in Ap with total substrate concentration ($A_{Total}$) is not possible.**

# Protein-protein interaction model (PPI)

**[adopted from Kholodenko et al 2015., Interface.]**

In this manuscript we analytically show the propensity of the PPI model (first described by Kholodenko et al., 2015) to exhibit biphasic behavior in the substrate form as the total enzyme concentration changes (enzyme biphasic).

To recap, this has been computationally established in the manuscript by Kholodenko et al., 2015, and here we show that this is analytically guaranteed in for the entire range of kinetic parameters due to the network structure of the system.

We further build on the computational result of Kholodenko et al., 2015 to show that

1. The PPI system does accommodate a biphasic behavior in the modified substrate form for the entire range of kinetic parameters due to network structure.
2. We show that the full model (described using mass action kinetics without any assumptions on enzymatic regimes such as the ones assumed in Kholodenko et al., 2015) is also capable of enzyme biphasic response for any kinetic parameter choice while being incapable of substrate biphasic response in the modified substrate form irrespective of kinetic parameters.

## Protein-protein interaction model [model from Kholodenko et al., 2015]

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(Student[LinearAlgebra]) : with(VectorCalculus) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature as described in Kholodenko et al. , 2015. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_{b1}\cdot A\cdot B + k_{ub1}\cdot AB :$$

$$dB := -k_{b1}\cdot A\cdot B + k_{ub1}\cdot AB - \frac{k_2\cdot B}{K_2 + B} + \frac{k_3\cdot A\cdot Bs}{K_3 + Bs} :$$

$$dBs := \frac{k_2\cdot B}{K_2 + B} - \frac{k_3\cdot A\cdot Bs}{K_3 + Bs} :$$

$$dAB := k_{b1}\cdot A\cdot B - k_{ub1}\cdot AB :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A_{Total} - A - AB :$$

$$BCon := B_{Total} - AB - B - Bs :$$

Now we begin by solving the system of equations to obtain correlations between the steady state concentrations of the variables. Primarily to obtain steady state correlations of variables as a function of concentrations of B and the Bs. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AB := simplify(solve(dAB, AB)) = \frac{k_{b1}\, A\, B}{k_{ub1}}$$

$$A := simplify((solve(dB, A))) :$$

$$A = \frac{k_2\, B\, (K_3 + Bs)}{(K_2 + B)\, k_3\, Bs}$$

$$AB = \frac{k_{b1}\, k_2\, B^2\, (K_3 + Bs)}{(K_2 + B)\, k_3\, Bs\, k_{ub1}}$$

Note for positive concentrations of B and Bs, A and AB are positive as well. Thus this results in us solving all ODEs and conservation equations describing the system, except for the conservation expressions ACon and BCon.

$$ACon$$

$$A_{Total} - \frac{k_2 B \left(K_3 + Bs\right)}{\left(K_2 + B\right) k_3 Bs} - \frac{k_{b1} k_2 B^2 \left(K_3 + Bs\right)}{\left(K_2 + B\right) k_3 Bs k_{ub1}} \tag{1.1}$$

$BCon$

$$B_{Total} - \frac{k_{b1} k_2 B^2 \left(K_3 + Bs\right)}{\left(K_2 + B\right) k_3 Bs k_{ub1}} - B - Bs \tag{1.2}$$

If we differentiate both these with respect to the total enzyme concentration ($A_{Total}$) in the system, we get

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial B} \cdot \frac{\partial B}{\partial A_{Total}} + \frac{\partial ACon}{\partial Bs} \cdot \frac{\partial Bs}{\partial A_{Total}}$$
$$\frac{dBCon}{dA_{Total}} = 0 = \frac{\partial BCon}{\partial B} \cdot \frac{\partial B}{\partial A_{Total}} + \frac{\partial BCon}{\partial Bs} \cdot \frac{\partial Bs}{\partial A_{Total}}$$

Now, in order for there to exist a biphasic response in B with total enzyme concentration, $\frac{\partial B}{\partial A_{Total}}$ must equal zero (the gradient at the peak concentration achieved by B in the dose response curve is zero). Thus, we can simplify these expressions further as,

**Note**

This simplification was possible since $\frac{\partial ACon}{\partial B}$ and $\frac{\partial BCon}{\partial B}$ have non-zero denominators (see below)

$$simplify(diff(ACon, B)) = -\frac{\left(\left(2 k_{b1} B + k_{ub1}\right) K_2 + k_{b1} B^2\right) \left(K_3 + Bs\right) k_2}{\left(K_2 + B\right)^2 k_3 Bs k_{ub1}}$$

$$simplify(diff(BCon, B)) = \frac{1}{\left(K_2 + B\right)^2 k_3 Bs k_{ub1}} \left(\left(\left(-k_2 k_{b1} - k_3 k_{ub1}\right) Bs - K_3 k_2 k_{b1}\right) B^2 - 2 \left(\left(k_2 k_{b1} + k_3 k_{ub1}\right) Bs + K_3 k_2 k_{b1}\right) K_2 B - Bs K_2^2 k_3 k_{ub1}\right)$$

$$1 = \frac{\partial ACon}{\partial Bs} \cdot \frac{\partial Bs}{\partial A_{Total}}$$
$$0 = \frac{\partial BCon}{\partial Bs} \cdot \frac{\partial Bs}{\partial A_{Total}}$$

Now since, $\frac{\partial ACon}{\partial Bs}$ is always positive (see below), $\frac{\partial Bs}{\partial A_{Total}}$ has to be non-zero. Which implies, in order to achieve a biphasic response in B with total enzyme concentration, $\frac{\partial BCon}{\partial Bs}$ must equal zero. This is in fact a sufficiency condition as well, since for any concentration of B satisfying such a condition, we can suitably find a total substrate and enzyme concentration using ACon and BCon at which they are steady state concentrations as well.

$simplify(diff(BCon, Bs)) =$

$$\frac{-\left(K_2+B\right) k_3\, Bs^2\, k_{ub1} + k_{b1}\, k_2\, B^2\, K_3}{\left(K_2+B\right) k_3\, Bs^2\, k_{ub1}}$$

Solving this for the concentration of Bs, we obtain the following correlation between the concentration of Bs and B, denoting the point at a biphasic peak achieved in the dose response curve of B with $A_{Total}$.

$Bs := simplify(solve(diff(BCon, Bs), Bs)[1])$

$$Bs := \frac{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}\, B}{k_3\, k_{ub1}\left(K_2+B\right)} \tag{1.3}$$

Now thus, all that remains is to find a suitable total substrate and enzyme amount where this peak can exist. This however is trivial to show, as for every choice of concentration of B, there exists a unique Bs, A, AB (each identified using the relevant correlations between the steady state species concentrations)

$ACon =$

$$A_{Total} - \frac{k_2\, k_{ub1}\left(K_3 + \frac{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}\, B}{k_3\, k_{ub1}\left(K_2+B\right)}\right)}{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}} - \frac{k_{b1}\, k_2\left(K_3 + \frac{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}\, B}{k_3\, k_{ub1}\left(K_2+B\right)}\right) B}{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}}$$

$BCon =$

$$B_{Total} - \frac{k_{b1}\, k_2\left(K_3 + \frac{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}\, B}{k_3\, k_{ub1}\left(K_2+B\right)}\right) B}{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}} - B - \frac{\sqrt{k_3\, k_{ub1}\left(K_2+B\right) K_3\, k_2\, k_{b1}}\, B}{k_3\, k_{ub1}\left(K_2+B\right)}$$

And thus implies, that for every choice of underlying kinetics and for every feasible concentration of B at the biphasic peak, there exits total amount of substrate and enzyme, where it is possible to observe a enzyme biphasic response with total enzyme concentration (i.e. the behavior to guaranteed to occur for any given underlying kinetics, provided there is flexibility on the total concentrations of the substrate).

Below, we illustrate a parameter set where the biphasic behavior is analytically predicted (and this is shown and confirmed computationally in Figure 2)

$k_{b1} := 0.000662316 : k_{ub1} := 0.000701878 : k_2 := 53.6473 : K_2 := 1880.36 : k_3 := 2.98182 : K_3 := 11.0657 :$

$B := 19.428458 :$

$simplify(AB) = 9.482758847$
$simplify(A) = 0.5172407784$
$simplify(Bs) = 6.109563569$

$A_{Total} := solve(ACon) = 9.999999625$
$B_{Total} := solve(BCon) = 35.02078042$

Thus in this subsection, we have shown how the PPI model illustrated by Kholodenko et al., 2015, which is a simple variation of a covalent modification system involving just an additional inactive complex formation with the enzyme (albeit with specific regimes of enzyme action - see next subsection) is capable of exhibiting enzyme biphasic dose response robustly in the entirety of kinetic parameter space.

## Protein-protein interaction model [Full system]

Note: In this subsection, we (re)model PPI network using mass action kinetic description. This kinetic description for all reactions assumes no limitations on the regime of enzyme action (unlike the PPI model described in Kholodenko et al., 2015 and used in the previous subsection) and thus is a more accurate representation of the variation of the covalent modification system it is. Using this model, we show that the PPI network is incapable of exhibiting substrate biphasic response, while (as shown in the previous subsection) it is capable of exhibiting robust enzyme biphasic responses in the modified substrate form.

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(Student[LinearAlgebra]) : with(VectorCalculus) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2}\right) \cdot ApK :$$
$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dK := -k_{b1}\, A\, K + \left(k_{ub1} + k_1\right) AK - k_{b2}\, Ap\, K + \left(k_{ub2}\right) ApK :$$
$$dP := -k_{b4}\, Ap\, P + \left(k_{ub4} + k_4\right) ApP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, KCon and PCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A_{Total} - A - ApP - ApK - AK - Ap :$$
$$KCon := K_{Total} - AK - K - ApK :$$
$$PCon := P_{Total} - P - ApP :$$

Now we begin by solving the system of equations to obtain correlations between the steady state concentrations of the variables. Primarily to obtain steady state correlations of variables as a function of concentrations of Ap and the ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$
$$assign\left(solve(\{dA, dApK, dApP\}, \{A, ApK, ApP\})\right)$$

We now introduce a new ratio, $\epsilon = \frac{K}{P}$ (ratio of the free enzymes), Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$K := \epsilon \cdot P :$$
$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_{ub2}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

After this simplification, we solve the conservation equation of the phosphatase for the steady state concentration of the enzyme as shown below,

$$P := solve(simplify(PCon), P) :$$

Thus this results in us solving all ODEs and conservation equations describing the system, except for the conservation expressions ACon and KCon.

$$A = \frac{Ap\, k_4\, c_4}{\epsilon\, k_1\, c_1}$$
$$AK = \frac{Ap\, k_4\, c_4\, P_{Total}}{k_1 \left(Ap\, c_4 + 1\right)}$$
$$ApK = \frac{c_2\, Ap\, \epsilon\, P_{Total}}{Ap\, c_4 + 1}$$
$$ApP = \frac{c_4\, Ap\, P_{Total}}{Ap\, c_4 + 1}$$

The steady state of the system is now defined by feasible solutions (of Ap and $\epsilon$) of the expressions ACon and BCon (Note for positive concentrations of Ap and $\epsilon$, the remaining species concentrations are positive as well).

*ACon*

$$A_{Total} - \frac{Ap\, k_4\, c_4}{\epsilon\, k_1\, c_1} - \frac{c_4\, Ap\, P_{Total}}{Ap\, c_4 + 1} - \frac{c_2\, Ap\, \epsilon\, P_{Total}}{Ap\, c_4 + 1} - \frac{Ap\, k_4\, c_4\, P_{Total}}{k_1 \left(Ap\, c_4 + 1\right)} - Ap \qquad \textbf{(2.1)}$$

*KCon*

$$K_{Total} - \frac{Ap\, k_4\, c_4\, P_{Total}}{k_1 \left(Ap\, c_4 + 1\right)} - \frac{\epsilon\, P_{Total}}{Ap\, c_4 + 1} - \frac{c_2\, Ap\, \epsilon\, P_{Total}}{Ap\, c_4 + 1} \qquad \textbf{(2.2)}$$

## Proof of absence of substrate biphasic response

Now, to show the absence of a substrate biphasic response in Ap with total substrate concentration, we proceed with a proof by contradiction. In order for there to exist a substrate biphasic, the following must be necessarily true at some steady state in the system (the gradient at the peak concentration achieved by Ap in the dose response curve is zero)

$$\frac{\partial Ap}{\partial A_{Total}} = 0$$

If we differentiate both conservation expressions with respect to the total enzyme concentration ($A_{Total}$) in the system, we get

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$
$$\frac{dKCon}{dA_{Total}} = 0 = \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

Since we begin with the assumption that there exists such a biphasic , we can simplify these expressions further as

$$1 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$
$$0 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

**Note**

This simplification was possible since $\frac{\partial ACon}{\partial Ap}$ and $\frac{\partial KCon}{\partial Ap}$ have non-zero denominators (see below)

$$simplify(diff(ACon, Ap)) = \frac{1}{\epsilon\, k_1\, c_1 \left(Ap\, c_4 + 1\right)^2} \left( -Ap^2\, {c_4}^3\, k_4 - Ap \left(Ap\, c_1\, \epsilon\, k_1 + 2\, k_4\right) {c_4}^2 + \left( -2\, c_1 \left( \left( Ap + \frac{P_{Total}}{2} \right) k_1 + \frac{k_4\, P_{Total}}{2} \right) \epsilon - k_4 \right) c_4 - c_1\, k_1\, \epsilon \left(c_2\, \epsilon\, P_{Total} + 1\right) \right)$$

$$simplify(diff(KCon, Ap)) = -\frac{\left( \left( -\epsilon\, k_1 + k_4 \right) c_4 + c_2\, \epsilon\, k_1 \right) P_{Total}}{k_1 \left(Ap\, c_4 + 1\right)^2}$$

Now since, $\frac{\partial KCon}{\partial \epsilon}$ is always negative (see below), $\frac{\partial \epsilon}{\partial A_{Total}}$ $\frac{\partial ACon}{\partial \epsilon}$ has a non-zero denominator.

$$simplify(diff(KCon, \text{epsilon})) = -\frac{P_{Total} \left(Ap\, c_2 + 1\right)}{Ap\, c_4 + 1}$$

$$simplify(diff(ACon, \text{epsilon})) = \frac{Ap \left( -c_2\, P_{Total}\, \epsilon^2\, k_1\, c_1 + Ap\, k_4\, {c_4}^2 + k_4\, c_4 \right)}{\epsilon^2\, k_1\, c_1 \left(Ap\, c_4 + 1\right)}$$

**This is a contradiction and thus substrate biphasic behavior in Ap with total substrate concentration ($A_{Total}$) is not possible.**

## Enzyme biphasic

The full model description of the PPI system is capable of exhibiting enzyme biphasic response. In this subsection, we prove the existence of the behavior irrespective of kinetic parameter choice.

We begin with the assumption that there exists an enzyme biphasic response in Ap to total kinase concentration. We know that in order for there to exist an enzyme biphasic, the following must be necessarily true at some steady state in the system (the gradient at the peak concentration achieved by Ap in the dose response curve is zero)

$$\frac{\partial Ap}{\partial K_{Total}} = 0$$

If we differentiate both conservation expressions with respect to the total enzyme concentration ($A_{Total}$) in the system, we get

$$\frac{dACon}{dK_{Total}} = 0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$
$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

Since we begin with the assumption that there exists such a biphasic, we can simplify these expressions further as

$$0 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$
$$1 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

**Note**

This simplification was possible since $\frac{\partial ACon}{\partial Ap}$ and $\frac{\partial BCon}{\partial Ap}$ have non-zero denominators (see below)

$$simplify(diff(ACon, Ap)) = \frac{1}{\epsilon\, k_1\, c_1 \left(Ap\, c_4 + 1\right)^2} \left( -Ap^2\, {c_4}^3\, k_4 - Ap \left(Ap\, c_1\, \epsilon\, k_1 + 2\, k_4\right) {c_4}^2 + \left( -2 \left( \left( Ap + \frac{P_{Total}}{2} \right) k_1 + \frac{k_4\, P_{Total}}{2} \right) c_1\, \epsilon - k_4 \right) c_4 - c_1\, k_1\, \epsilon \left( c_2\, \epsilon\, P_{Total} + 1 \right) \right)$$

$$simplify(diff(KCon, Ap)) \quad -\frac{\left( \left( -\epsilon\, k_1 + k_4 \right) c_4 + c_2\, \epsilon\, k_1 \right) P_{Total}}{k_1 \left(Ap\, c_4 + 1\right)^2} \qquad \textbf{(2.2.1.1)}$$

Now since, $\frac{\partial KCon}{\partial \epsilon}$ is always negative (see below), $\frac{\partial \epsilon}{\partial A_{Total}}$ has to be non-zero. Which implies, in order to achieve a biphasic response in Ap with total enzyme concentration, $\frac{\partial ACon}{\partial \epsilon}$

must equal zero. This is in fact a sufficiency condition as well, since for any concentration of $\epsilon$ and Ap satisfying such a condition, we can suitably find a total substrate and enzyme concentration using ACon and KCon at which they are steady state concentrations as well.

$$simplify(diff(KCon, \epsilon)) = -\frac{P_{Total}\left(Ap\, c_2 + 1\right)}{Ap\, c_4 + 1}$$

Solving $\frac{\partial ACon}{\partial \epsilon} = 0$ for the concentration of Ap, we obtain the following correlation between the concentration of Ap and $\epsilon$, denoting the point at a biphasic peak achieved in the dose response curve of Ap with $K_{Total}$.

$$Ap := simplify(solve(diff(ACon, \epsilon), Ap)[2]) = \frac{c_1 c_2 \epsilon^2 k_1 P_{Total} - c_4 k_4}{c_4^2 k_4}$$

Now thus, all that remains is to find a suitable total substrate and enzyme amount where this peak can exist. This however is trivial to show, as for every choice of concentration of $\epsilon$ (when Ap is positive here), there exists a unique AK, ApK, ApP (each identified using the relevant correlations between the steady state species concentrations)

$\epsilon$ **(2.2.1)**

$$ACon = A_{Total} - \frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 \epsilon k_1 c_1} - \frac{\left(c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4\right) P_{Total}}{c_4 k_4 \left(\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1\right)} - \frac{c_2 \left(c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4\right) \epsilon P_{Total}}{c_4^2 k_4 \left(\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1\right)} - \frac{\left(c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4\right) P_{Total}}{c_4 k_1 \left(\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1\right)} - \frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4^2 k_4}$$

$$KCon = K_{Total} - \frac{\left(c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4\right) P_{Total}}{c_4 k_1 \left(\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1\right)} - \frac{\epsilon P_{Total}}{\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1} - \frac{c_2 \left(c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4\right) \epsilon P_{Total}}{c_4^2 k_4 \left(\frac{c_2 P_{Total} \epsilon^2 k_1 c_1 - k_4 c_4}{c_4 k_4} + 1\right)}$$

And thus implies, that for every choice of underlying kinetics and for every feasible concentration of $\epsilon$ (and Ap) at the biphasic peak, there exits total amount of substrate and enzyme, where it is possible to observe a enzyme biphasic response with total enzyme concentration (i.e. the behavior to guaranteed to occur for any given underlying kinetics, provided there is flexibility on the total concentrations of the substrate).

Below, we illustrate a parameter set where the biphasic behavior is analytically predicted (and this is shown and confirmed computationally in figure 2)

$k_1 := 1 : k_4 := 1 : k_{b1} := 1 : k_{b2} := 20 : k_{b4} := 1 : k_{ub1} := 1 : k_{ub2} := 1 : k_{ub4} := 1 : c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_{ub2}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$

$P_{Total} := 1 :$

$Ap := 1.139478030 :$ epsilon $:= 2 :$

$simplify(AK) = 0.3629514268$
$simplify(A) = 0.5697390150$
$simplify(ApK) = 29.03611414$
$simplify(ApP) = 0.3629514266$

$A_{Total} := solve(ACon) = 31.47123404$
$K_{Total} := solve(KCon) = 30.67316271$

**Thus, in this subsection we have shown how for every choice of underlying kinetics the PPI model (without any assumptions on regime of enzyme action) is capable of exhibiting enzyme biphasic in Ap dose response with changing total kinase concentration.**

# Double site phosphorylation system (DSP) [common kinase and common phosphatase]

## Features and requirements of obtaining biphasic response in the maximally modified substrate

In this file we analytically study the presence of enzyme and substrate biphasic in the maximally modified substrate form. In doing so we establish the following key results with regard to the behavior.

**Enzyme biphasic response (biphasic behavior in the dose response curve of App as $K_{Total}$ changes)**

1. The maximally modified substrate is incapable of exhibiting enzyme biphasic response irrespective of parameter values.

**Substrate biphasic response (biphasic behavior in the dose response curve of App as $A_{Total}$ changes)**

1. Substrate biphasic response is possible in App for any kinetic regime (i.e. for any choice of underlying kinetics the system is capable of exhibiting substrate biphasic dose response at some total concentration of enzymes)

We note that the key signature of biphasic behavior in the dose response curve of the system is the presence of a steady state of the system that satisfies the following condition.

$$\frac{dApp}{dK_{Total}} = 0 \text{ (for enzyme biphasic)}$$

$$\frac{dApp}{dA_{Total}} = 0 \text{ (for substrate biphasic)}$$

**Model DSP:** We first describe the model of double site phosphorylation with common kinase and common phosphatase enzyme action.

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - \left(k_{ub3} + k_3\right) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dK := -k_{b1}\, A\, K + \left(k_{ub1} + k_1\right) AK - k_{b2}\, Ap\, K + \left(k_{ub2} + k_2\right) ApK :$$
$$dP := -k_{b3}\, App\, P + \left(k_{ub3} + k_3\right) AppP - k_{b4}\, Ap\, P + \left(k_{ub4} + k_4\right) ApP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK - K_{Total} :$$

We now solve the system described at steady state to obtain expression linking the steady state concentrations of the various species. Here we use the Maple command *solve* to solve the equations for a given variable as shown below. We pursue this to finally obtain the steady state concentrations of most species in terms of App and a ratio $\epsilon$ (defined below).

$$assign(solve(\{dAK, dApK, dAppP, dApP\}, \{AK, ApK, AppP, ApP\}))$$
$$assign(solve(\{dA, dAp\}, \{A, Ap\}))$$

Further we now introduce a ratio, $\epsilon$ = K/P (defined as the ratio of the free enzymes). Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$K := \epsilon \cdot P :$$
$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

Once this is done, we again solve for the steady state of the phosphatase using the conservation expression for the enzyme (PCon).

$$P := simplify(solve(PCon, P)) = \frac{P_{Total}\,\epsilon\,k_2\,c_2}{c_2\,k_2\,(App\,c_3+1)\,\epsilon + c_4\,c_3\,App\,k_3}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration App and $\epsilon$

$$A = \frac{k_4\,c_4\,c_3\,App\,k_3}{\epsilon^2\,k_2\,c_2\,k_1\,c_1}$$

$$Ap = \frac{c_3\,App\,k_3}{\epsilon\,k_2\,c_2}$$

$$AK = \frac{k_4\,c_4\,c_3\,App\,k_3\,P_{Total}}{k_1\,(c_2\,k_2\,(c_3\,App+1)\,\epsilon + c_4\,c_3\,App\,k_3)}$$

$$ApK = \frac{c_2\,c_3\,App\,k_3\,\epsilon\,P_{Total}}{c_2\,k_2\,(c_3\,App+1)\,\epsilon + c_4\,c_3\,App\,k_3}$$

$$AppP = \frac{c_3\,App\,P_{Total}\,\epsilon\,k_2\,c_2}{c_2\,k_2\,(c_3\,App+1)\,\epsilon + c_4\,c_3\,App\,k_3}$$

$$ApP = \frac{c_4\,c_3\,App\,k_3\,P_{Total}}{c_2\,k_2\,(c_3\,App+1)\,\epsilon + c_4\,c_3\,App\,k_3}$$

Note that when App and $\epsilon$ are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of App and $\epsilon$. We now have two expressions, ACon (or ACon_Red if we are working in the enzyme limiting regime - see below) and KCon - the conservation of the substrate and kinase, whose solution for the variables define the steady state of the system.

# Enzyme biphasic response

## Impossibility of enzyme biphasic behavior in the limiting enzyme regime

In this subsection we show that the DSP model is incapable of enzyme biphasic behavior in the maximally modified substrate form when the enzymes are limiting (where the total substrate concentration is much greater than the total enzyme concentration). Thus in such a regime, the conservation expressions for the substrates can be written as shown below (ACon_Red = 0)

$$ACon\_Red := A + Ap + App - A_{Total}:$$

As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system,

$$\frac{dApp}{dK_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ACon_Red = 0 (see below) whose solutions to the variables App and $\epsilon$ define the steady state of the system.

$$ACon\_Red = \frac{k_4 c_4 c_3 App k_3}{\epsilon^2 k_2 c_2 k_1 c_1} + \frac{c_3 App k_3}{\epsilon k_2 c_2} + App - A_{Total}$$

$$KCon = \frac{\epsilon^2 P_{Total} k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} - K_{Total}$$

If we differentiate both these with respect to the total kinase in the system, we get

$$\frac{dACon\_Red}{dK_{Total}} = 0 = \frac{\partial ACon\_Red}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial ACon\_Red}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

Now in order to show the absence of an enzyme biphasic response in this regime of enzyme action, we use a proof by contradiction. If we are to assume that there exists a biphasic response, then $\frac{dApp}{dK_{Total}} = 0$ for some steady state.

This results in the above expressions being simplified as follows.

$$0 = \frac{\partial ACon\_Red}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

$$1 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon\_Red}{\partial App}, \frac{\partial KCon}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial K_{Total}}$ can be zero.

$$simplify(diff(ACon\_Red, App)) = \frac{1}{\left(Ap c_2 + Bp d_2 + 1\right)^2 K k_1 c_1 p_1} \left(\left(\left(Ap^2 k_1 c_2^2 + \left(\left(P_{Total} + 2 Ap\right) k_1 + P_{Total} k_2\right) \left(Bp d_2 + 1\right) c_2 + k_1 \left(Bp d_2 + 1\right)^2\right) K c_1 + P_{Total} c_2 k_2 \left(Bp d_2 + 1\right)\right) p_1 - d_2 Bp P_{Total} p_2 c_2 K k_1 c_1\right)$$

$$simplify(diff(KCon, App)) = -\frac{\left(c_2 d_1 \left(p_1 + p_2\right) Ap^2 + 2 p_2 c_2 Ap + p_2 \left(Bp d_2 + 1\right)\right) d_2 Bp P_{Total}}{\left(Ap c_2 + Bp d_2 + 1\right)^2 Ap^2 p_1 d_1}$$

However we know that $\frac{\partial ACon\_Red}{\partial \epsilon}$ is non-zero (see below).

$$simplify(diff(ACon\_Red, \epsilon)) = -\frac{c_3\, App\, k_3 \left(c_1\, \epsilon\, k_1 + 2\, c_4\, k_4\right)}{\epsilon^3\, k_2\, c_2\, k_1\, c_1}$$

This implies that $\frac{\partial \epsilon}{\partial K_{Total}}$ is zero. However this leads to a contradiction as $1 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$. (Note that the denominator of $\frac{\partial KCon}{\partial \epsilon}$ is non-zero as well).
Hence our assumption must be wrong.

**Thus enzyme biphasic behavior in App with $K_{Total}$ is impossible when total substrate concentrations are higher than total enzyme concentrations.**

## Presence and necessary analytical condition (involving catalytic constants) for biphasic behavior to exist

In this subsection, we analytically show the presence of enzyme biphasic (in the full system) and extract kinetic constraints that enable (and preclude) biphasic behavior. As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system.

$$\frac{\partial App}{\partial K_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ACon = 0 (see below) whose solutions to the variables App and $\epsilon$ define the steady state of the system.

$$ACon = \frac{k_4\, c_4\, c_3\, App\, k_3}{\epsilon^2\, k_2\, c_2\, k_1\, c_1} + \frac{c_3\, App\, k_3}{\epsilon\, k_2\, c_2} + App + \frac{k_4\, c_4\, c_3\, App\, k_3\, P_{Total}}{k_1 \left(c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3\right)} + \frac{c_2\, c_3\, App\, k_3\, \epsilon\, P_{Total}}{c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3} + \frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2}{c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3} + \frac{c_4\, c_3\, App\, k_3\, P_{Total}}{c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3} - A_{Total}$$

$$KCon = \frac{\epsilon^2\, P_{Total}\, k_2\, c_2}{c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3} + \frac{k_4\, c_4\, c_3\, App\, k_3\, P_{Total}}{k_1 \left(c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3\right)} + \frac{c_2\, c_3\, App\, k_3\, \epsilon\, P_{Total}}{c_2\, k_2 \left(c_3\, App + 1\right) \epsilon + c_4\, c_3\, App\, k_3} - K_{Total}$$

If we differentiate both these with respect to the total kinase concentration in the system, we get

$$\frac{dACon}{dK_{Total}} = 0 = \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

Now in order to show the presence of an enzyme biphasic response and study it's features, we begin with exploring the necessary features that the system must satisfy for the behavior to exist. We begin with the basic tenet that for the behavior there should exist a steady state of the system where $\frac{dApp}{dK_{Total}} = 0$ is satisfied.
At this point then, the above expressions simply as follows

$$0 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$
$$1 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}, \frac{\partial KCon}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial K_{Total}}$ can be zero.

$$simplify(diff(ACon, App)) = \left(k_2^2 c_2^3 \left(App^2 c_3^2 k_2 + \left(\left(2\, App + P_{Total}\right) k_2 + P_{Total} k_3\right) c_3 + k_2\right) k_1 c_1 \epsilon^4 + k_2^2 c_2^2 \left(App^2 c_3^2 k_1 + 2\, App\, k_1 \left(c_4\, App + 1\right) c_3 + 2\, App\, c_4 k_1 + P_{Total} c_4 k_1 + P_{Total} c_4 k_4 + k_1\right) k_3 c_1 c_3 \epsilon^3 + 2 k_2 c_2 c_4 \left(\frac{k_2 k_4 \left(c_3\, App + 1\right)^2 c_2}{2} + App\, c_1 c_3 k_1 k_3 \left(c_3\, App + \frac{1}{2} c_4\, App + 1\right)\right) k_3 c_3 \epsilon^2 + \left(2 k_2 k_4 \left(c_3\, App + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) c_4^2\, App\, k_3^2 c_3^2 \epsilon + App^2 c_3^3 c_4^3 k_3^3 k_4\right) \Big/ \left(k_2 \left(c_2 k_2 \left(c_3\, App + 1\right) \epsilon + c_4 c_3\, App\, k_3\right)^2 c_2 k_1 c_1 \epsilon^2\right)$$

$$simplify(diff(KCon, App)) = -\frac{c_3 \epsilon k_2 \left(k_1 c_2 k_2 \epsilon^2 - k_1 k_3 \left(c_2 - c_4\right) \epsilon - k_4 c_4 k_3\right) c_2 P_{Total}}{\left(c_2 k_2 \left(c_3\, App + 1\right) \epsilon + c_4 c_3\, App\, k_3\right)^2 k_1}$$

We know that the denominator of $\frac{\partial KCon}{\partial \epsilon}$ is non-zero, thus this implies that $\frac{\partial \epsilon}{\partial K_{Total}}$ has to be non-zero in order to satisfy the second expression above.

This insight then informs us that in order to satisfy the first expression, $\frac{\partial ACon}{\partial \epsilon}$ must be equal to zero.

$$Condition := simplify(diff(ACon, \epsilon))$$
$$Condition := -\left(App\, k_3 \left(k_2 c_2^2 \left(\left(App^2 c_3^2 k_1 + App\left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total}\left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App\, P_{Total} c_3 c_4 k_1 k_3\right) c_1 \epsilon^3 + 2 k_2 \left(k_2 k_4 \left(c_3\, App + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) c_2 \left(c_3\, App + 1\right) c_4 \epsilon^2 + \left(4 k_2 k_4 \left(c_3\, App + 1\right) c_2\right. \right. \quad \textbf{(1.2.1)}$$

$$+ App\, c_1 c_3 k_1 k_3\Big)\, c_4^2\, App\, k_3 c_3 \epsilon + 2\, App^2 c_3^2 c_4^3 k_3^2 k_4\Big)\, c_3\Big) \Big/ \Big(k_2 \Big(c_2 k_2 \Big(c_3\, App + 1\Big)\epsilon + c_4 c_3\, App\, k_3\Big)^2 c_2 k_1 c_1 \epsilon^3\Big)$$

In order for this expression to equal zero, the numerator of it (which a polynomial in $\epsilon$/App) to be equal to zero. Isolating and studying it as a polynomial in App (as shown below), we can observe that all except the coefficient of the second exponent are negative.

$$\begin{aligned} &collect(numer(Condition), App) \\ &-k_3 \Big(c_1 c_2^2 c_3^2 \epsilon^3 k_1 k_2^2 + 2\, c_1 c_2 c_3^2 c_4 \epsilon^2 k_1 k_2 k_3 + 2\, c_2^2 c_3^2 c_4 \epsilon^2 k_2^2 k_4 + c_1 c_3^2 c_4^2 \epsilon k_1 k_3^2 \\ &\quad + 4\, c_2 c_3^2 c_4^2 \epsilon k_2 k_3 k_4 + 2\, c_3^2 c_4^3 k_3^2 k_4\Big)\, c_3\, App^3 - k_3 \Big(-c_1 c_2^2 c_3 c_4 \epsilon^3 k_1 k_2 k_3 P_{Total} \\ &\quad + c_1 c_2^2 c_3 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + 2\, c_1 c_2^2 c_3 \epsilon^3 k_1 k_2^2 + 2\, c_1 c_2 c_3 c_4 \epsilon^2 k_1 k_2 k_3 \\ &\quad + 4\, c_2^2 c_3 c_4 \epsilon^2 k_2^2 k_4 + 4\, c_2 c_3 c_4^2 \epsilon k_2 k_3 k_4\Big)\, c_3\, App^2 - k_3 \Big(c_1 c_2^2 c_4 \epsilon^3 k_1 k_2^2 P_{Total} \\ &\quad + c_1 c_2^2 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + c_1 c_2^2 \epsilon^3 k_1 k_2^2 + 2\, c_2^2 c_4 \epsilon^2 k_2^2 k_4\Big)\, c_3\, App \end{aligned} \tag{1.2.2}$$

Thus for the polynomial to equal zero, this coefficient must be positive. Isolating the coefficient further as shown below, reveals that $k_1 k_3 - k_2 k_4$ must be greater than zero for this to be true.

$$\begin{aligned} &simplify\Big(-k_3 \Big(-c_1 c_2^2 c_3 c_4 \epsilon^3 k_1 k_2 k_3 P_{Total} + c_1 c_2^2 c_3 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + 2\, c_1 c_2^2 c_3 \epsilon^3 k_1 k_2^2 \\ &\quad + 2\, c_1 c_2 c_3 c_4 \epsilon^2 k_1 k_2 k_3 + 4\, c_2^2 c_3 c_4 \epsilon^2 k_2^2 k_4 + 4\, c_2 c_3 c_4^2 \epsilon k_2 k_3 k_4\Big)\, c_3\, App^2\Big) \\ &k_2 c_2 \Big(\Big(P_{Total}\Big(k_1 k_3 - k_2 k_4\Big) c_4 - 2\, k_1 k_2\Big) c_1 c_2 \epsilon^2 - 2\, c_4 \Big(c_1 k_1 k_3 + 2\, c_2 k_2 k_4\Big)\epsilon \\ &\quad - 4\, c_4^2 k_3 k_4\Big)\, App^2 k_3 c_3^2 \epsilon \end{aligned} \tag{1.2.3}$$

Thus, this becomes a necessary condition for the presence of biphasic dose response in the maximally modified substrate (App) with total enzyme concentration ($K_{Total}$).

**Necessary condition for enzyme biphasic dose response:** $k_1 k_3 - k_2 k_4 > 0$

It is worth noting that from a different study (C. Conradi and M. Mincheva, "Catalytic constants enable the emergence of bistability in dual phosphorylation," J. R. Soc. Interface, vol. 11, no. 95, 2014, doi: 10.1098/rsif.2014.0158) it was established that if $k_1 k_3 - k_2 k_4 < 0$ then multi-stationarity is guaranteed for some finite positive total enzyme, substrate concentrations.

We now make a comment regarding the contrast in the necessary and sufficient condition for obtaining biphasic behavior (result here) and the condition enabling bistability in the model (Conradi et al., 2014). While the conditions contrast it is possible to obtain both biphasic and bistability for the same underlying kinetic system (but different choice of total concentrations of substrates and enzymes). This is depicted in the main text in figure N-2.

**Comment on guarantees of biphasic behavior:**

In this discussion we make the augment to show that should the necessary condition above be satisfied, then there exists total concentrations of substrate ($A_{Total}$) and phosphatase ($P_{Total}$).

In order to make the argument, we recapitulate that the steady state of the system is defined by the solutions to the expressions ACon and KCon. However since we are

$$ACon$$

$$\frac{k_4 c_4 c_3 App k_3}{\epsilon^2 k_2 c_2 k_1 c_1} + \frac{c_3 App k_3}{\epsilon k_2 c_2} + App + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{c_3 App P_{Total} \epsilon k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{c_4 c_3 App k_3 P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} - A_{Total} \tag{1.2.4}$$

$$KCon$$

$$\frac{\epsilon^2 P_{Total} k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} - K_{Total} \tag{1.2.5}$$

However, since we have complete flexibility for the total amounts, any feasible concentration of App and $\epsilon$ can be made to satisfy ACon and KCon by choosing $A_{Total}$ and $K_{Total}$.

Thus in order to have a steady state which presents with a biphasic response, it is thus enough to find a feasible concentration of App and $\epsilon$ that satisfies the necessary condition for biphasic response from earlier ( $\frac{dApp}{dK_{Total}} = 0$ ).
We recall that this line or reasoning resulted in requiring us to find a steady state of App and $\epsilon$ that satisfies $\frac{\partial ACon}{\partial \epsilon} = 0$. This resulted in the following expression.

$$collect\left(\frac{-numer(Condition)}{App}, App\right)$$

$$k_3 \left(c_1 c_2^2 c_3^2 \epsilon^3 k_1 k_2^2 + 2 c_1 c_2 c_3^2 c_4 \epsilon^2 k_1 k_2 k_3 + 2 c_2^2 c_3^2 c_4 \epsilon^2 k_2^2 k_4 + c_1 c_3^2 c_4^2 \epsilon k_1 k_3^2 + 4 c_2 c_3^2 c_4^2 \epsilon k_2 k_3 k_4 + 2 c_3^2 c_4^3 k_3^2 k_4\right) c_3 App^2 + k_3 \left(-c_1 c_2^2 c_3 c_4 \epsilon^3 k_1 k_2 k_3 P_{Total} + c_1 c_2^2 c_3 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + 2 c_1 c_2^2 c_3 \epsilon^3 k_1 k_2^2 + 2 c_1 c_2 c_3 c_4 \epsilon^2 k_1 k_2 k_3 + 4 c_2^2 c_3 c_4 \epsilon^2 k_2^2 k_4 + 4 c_2 c_3 c_4^2 \epsilon k_2 k_3 k_4\right) c_3 App + k_3 \left(c_1 c_2^2 c_4 \epsilon^3 k_1 k_2^2 P_{Total} + c_1 c_2^2 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + c_1 c_2^2 \epsilon^3 k_1 k_2^2 + 2 c_2^2 c_4 \epsilon^2 k_2^2 k_4\right) c_3 \tag{1.2.6}$$

This is a second order polynomial in App (after we have factored an App from the expression). Now we can observe (as we have early) that the leading coefficient and the constant term are always positive, while the coefficient of the first exponent can be negative (this is the coefficient from which we discerned the necessary condition).

$$simplify\left(-c_1 c_2^2 c_3 c_4 \epsilon^3 k_1 k_2 k_3 P_{Total} + c_1 c_2^2 c_3 c_4 \epsilon^3 k_2^2 k_4 P_{Total} + 2 c_1 c_2^2 c_3 \epsilon^3 k_1 k_2^2 + 2 c_1 c_2 c_3 c_4 \epsilon^2 k_1 k_2 k_3 + 4 c_2^2 c_3 c_4 \epsilon^2 k_2^2 k_4 + 4 c_2 c_3 c_4^2 \epsilon k_2 k_3 k_4\right)$$

$$-c_3 \epsilon \left(\left(P_{Total} \left(k_1 k_3 - k_2 k_4\right) c_4 - 2 k_1 k_2\right) c_1 c_2 \epsilon^2 - 2 c_4 \left(c_1 k_1 k_3 + 2 c_2 k_2 k_4\right) \epsilon - 4 c_4^2 k_3 k_4\right) k_2 c_2 \tag{1.2.7}$$

Now we can observe that if the necessary condition is satisfied ($k_1 k_3 - k_2 k_4 > 0$), the phosphatase concentration can be sufficiently increased to make the whole expression further negative. This implies that if the necessary condition is satisfied then, by varying (increasing) the total phosphatase concentration, the coefficient can be made more and more negative. Thus, this implies that for some arbitrary choice of $\epsilon$ there exists feasible (positive) values of App for which the expression is guaranteed to be 0.

Thus this implies that should the required necessary condition be satisfied, then for a high enough $P_{Total}$, there exits an App and $\epsilon$ (which provides $A_{Total}$, $K_{Total}$ from there on), which satisfies, $\frac{\partial ACon}{\partial \epsilon} = 0$ and thus satisfies the necessary requirement of the enzyme biphasic response of App with $K_{Total}$. Thus ensuring the sufficiency of the necessary condition to guarantee the existence enzyme biphasic response for some total concentration of substrate and enzymes.

## Substrate biphasic response

In this subsection, we analytically show the presence of and study the features of substrate biphasic (in the full system) and extract kinetic constraints that enable (and preclude) biphasic behavior. As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system.

$$\frac{dApp}{dA_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ACon = 0 (see below) whose solutions to the variables App and $\epsilon$ define the steady state of the system.

$$\begin{aligned} ACon = {} & \frac{k_4 c_4 c_3 App\, k_3}{\epsilon^2 k_2 c_2 k_1 c_1} + \frac{c_3 App\, k_3}{\epsilon k_2 c_2} + App + \frac{k_4 c_4 c_3 App\, k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3\right)} \\ & + \frac{c_2 c_3 App\, k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3} + \frac{c_3 App\, P_{Total} \epsilon k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3} \\ & + \frac{c_4 c_3 App\, k_3 P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3} - A_{Total} \end{aligned}$$

$$\begin{aligned} KCon = {} & \frac{\epsilon^2 P_{Total} k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3} + \frac{k_4 c_4 c_3 App\, k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3\right)} \\ & + \frac{c_2 c_3 App\, k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\, k_3} - K_{Total} \end{aligned}$$

If we differentiate both these with respect to the total substrate concentration in the system, we get

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

$$\frac{dKCon}{dA_{Total}} = 0 = \frac{\partial KCon}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

Now in order to show the presence of a substrate biphasic response and study it's features, we begin with

exploring the necessary features that the system must satisfy for the behavior to exist. We begin with the basic tenet that for the behavior there should exist a steady state of the system where $\frac{\partial App}{\partial A_{Total}} = 0$ is satisfied.
At this point then, the above expressions simply as follows

$$1 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$
$$0 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}, \frac{\partial KCon}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$$simplify(diff(ACon, App)) =$$
$$\left(\left(App^2 c_3^2 k_2 + \left(\left(2\,App + P_{Total}\right) k_2 + P_{Total} k_3\right) c_3 + k_2\right) c_2^3 c_1 k_2^2 k_1 \epsilon^4 + k_3 \left(App^2 c_3^2 k_1 + 2\,App\,k_1 \left(c_4 App + 1\right) c_3 + 2\,App\,c_4 k_1 + P_{Total} c_4 k_1 + P_{Total} c_4 k_4 + k_1\right) c_2^2 c_1 k_2^2 c_3 \epsilon^3 + 2\left(\frac{k_2 k_4 \left(c_3 App + 1\right)^2 c_2}{2} + App\,c_1 c_3 k_1 k_3 \left(c_3 App + \frac{1}{2} c_4 App + 1\right)\right) k_3 c_4 c_2 k_2 c_3 \epsilon^2 + App\,k_3^2 c_4^2 \left(2\,k_2 k_4 \left(c_3 App + 1\right) c_2 + App\,c_1 c_3 k_1 k_3\right) c_3^2 \epsilon + App^2 c_3^3 c_4^3 k_3^3 k_4\right) \Big/ \left(c_2 c_1 k_2 \epsilon^2 k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\,k_3\right)^2\right)$$
$$simplify(diff(KCon, App)) = -\frac{\epsilon \left(k_1 c_2 k_2 \epsilon^2 - k_1 k_3 \left(c_2 - c_4\right) \epsilon - k_4 c_4 k_3\right) c_3 c_2 P_{Total} k_2}{\left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App\,k_3\right)^2 k_1}$$

Observing, the second expression above (from differentiation of the total kinase concentration) we can see that either $\frac{\partial KCon}{\partial \epsilon}$ or $\frac{\partial \epsilon}{\partial A_{Total}}$ must be equal to zero. However, $\frac{\partial \epsilon}{\partial A_{Total}}$ cannot be zero as, if it was indeed zero, then since the denominator of $\frac{\partial ACon}{\partial \epsilon}$ is non-zero, there would be a contradiction with the first expression (from differentiation of the total substrate concentration). Thus, $\frac{\partial KCon}{\partial \epsilon}$ must be equal to zero.

$$Condition := simplify(diff(KCon, \text{epsilon}))$$
$$Condition := \frac{1}{\left(App\left(\epsilon k_2 c_2 + c_4 k_3\right) c_3 + \epsilon k_2 c_2\right)^2 k_1} \left(\left(App^2 c_4 k_3 \left(k_1 k_3 - k_2 k_4\right) c_3^2 + App\left(k_1 c_2 k_2 \epsilon^2 + 2\,k_3 \left(k_1 \epsilon - \frac{k_4}{2}\right) c_4\right) k_2 c_3 + \epsilon^2 k_2^2 c_2 k_1\right) c_2 P_{Total}\right) \quad \textbf{(2.1)}$$

Now for this expression to be zero, the numerator (which is a function of App and epsilon) must be equal to zero. isolating this, we get the following expression

$$simplify(numer(Condition))$$

$$\left(App^2 c_4 k_3 \left(k_1 k_3 - k_2 k_4\right) c_3^2 + App \left(k_1 c_2 k_2 \epsilon^2 + 2 k_3 \left(k_1 \epsilon - \frac{k_4}{2}\right) c_4\right) k_2 c_3 + \epsilon^2 k_2^2 c_2 k_1\right) c_2 P_{Total} \tag{2.2}$$

We recap here that, so long as this expression is zero for a feasible steady state concentrations of App and epsilon, we can find total concentrations of substrates and enzyme where the biphasic response is guaranteed. Keeping this in mind, we rewrite the expression as a polynomial function of epsilon as shown below

$$collect(numer(Condition), \text{epsilon})$$

$$\left(App\, k_1 c_2 k_2^2 c_3 + k_2^2 c_2 k_1\right) c_2 P_{Total} \epsilon^2 + 2\, App\, k_1 k_3 c_4 k_2 c_3 c_2 P_{Total} \epsilon + \left(App^2 c_3^2 c_4 k_1 k_3^2 - App^2 c_3^2 c_4 k_2 k_3 k_4 - App\, c_3 c_4 k_2 k_3 k_4\right) c_2 P_{Total} \tag{2.3}$$

This allows us to observe that so long as the constant term (which is itself a function of App) is negative, there is guaranteed to exist a steady state of the system where the expression is zero and thus where the system accommodate a biphasic dose response with total substrate concentration ($A_{Total}$).

Isolating the constant we can observe the grouping of catalytic constants from the enzyme biphasic responses' necessary condition appear here too.

$$simplify\left(\left(App^2 c_3^2 c_4 k_1 k_3^2 - App^2 c_3^2 c_4 k_2 k_3 k_4 - App\, c_3 c_4 k_2 k_3 k_4\right) c_2 P_{Total}\right)$$

$$App\, P_{Total} c_3 c_2 k_3 \left(App \left(k_1 k_3 - k_2 k_4\right) c_3 - k_2 k_4\right) c_4 \tag{2.4}$$

Now, if this catalytic constant grouping $(k_1 k_3 - k_2 k_4)$ is negative, then the constant term is negative irrespective of the steady state concentration of App and thus the whole expression from earlier is zero, guaranteeing substrate biphasic response. However, if the grouping $(k_1 k_3 - k_2 k_4)$ is positive, then the constant term can still be negative, however in this case App (at the biphasic peak) is necessarily to be less than $\frac{k_2 k_4}{c_3 \cdot \left(k_1 k_3 - k_2 k_4\right)}$

This is however still feasible since, App and epsilon can take any positive concentration and we can find a suitable total concentration of substrate and enzyme that will accommodate it as steady state concentrations (using ACon and KCon which we are yet to solve).

$$ACon = \frac{k_4 c_4 c_3 App k_3}{\epsilon^2 k_2 c_2 k_1 c_1} + \frac{c_3 App k_3}{\epsilon k_2 c_2} + App + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{c_3 App P_{Total} \epsilon k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{c_4 c_3 App k_3 P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} - A_{Total}$$

$$KCon =$$

$$\frac{\epsilon^2 P_{Total} k_2 c_2}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{c_2 k_2 \left(c_3 App + 1\right) \epsilon + c_4 c_3 App k_3} - K_{Total}$$

Thus, what this implies is that, irrespective of the underlying grouping (or the sign of the catalytic constant grouping $k_1 k_3 - k_2 k_4$), substrate biphasic behavior is guaranteed to exist for some total concentration of substrate and enzyme in the system.

However, we wish to note here that depending on the sign of the catalytic constant grouping $k_1 k_3 - k_2 k_4$ it is either more likely or less likely to observe substrate biphasic response in App with total amount of substrate in the system. Further, depending on the sign of the catalytic constant grouping (if $k_1 k_3 - k_2 k_4 > 0$), the peak concentration of the biphasic response capable of being observed is capped at $\frac{k_2 k_4}{c_3 \cdot \left(k_1 k_3 - k_2 k_4\right)}$ which is not the case when $(k_1 k_3 - k_2 k_4 < 0)$ when App can take any value at biphasic peak (given that the total amounts are completely flexible.

$$T := simplify\Big(\Big(\Big(k_{b2}\ \Big(\Big(\Big(\big(Ap\, k_{b4} + k_4 + k_{ub4}\big)\, k_3 + k_2\ \big(k_{ub4} + k_4\big)\Big)\, k_1 - k_{b4}\, Ap\, k_2\, k_4\Big)\, P + App\, k_1\, k_2\ \big(k_{ub4} + k_4\big)\Big)\, K^2 + \Big(k_{b4}\ \Big(k_3\ \big(Ap\, k_{b2} + k_2 + k_{ub2}\big)\, k_1 - \big(\big(-k_{ub2} - k_2\big)\, k_3 + k_{b2}\, Ap\, k_2\big)\, k_4\Big)\, P + \Big(\big(Ap\ (A + Ap + App)\, k_{b4} + \big(k_{ub4} + k_4\big)\ (A + Ap)\big)\, k_{b2} + \big(k_{ub2} + k_2\big)\ \big((Ap + App)\, k_{b4} + k_{ub4} + k_4\big)\Big)\, k_3\, k_1 - k_{b2}\, k_{b4}\, Ap\, k_2\, k_4\ (A + Ap + App)\Big)\, P\, K + A\, k_{b4}\, P^2\, k_3\, k_4\ \big(k_{ub2} + k_2\big)\Big)\, k_{b3} + k_2\ \big(k_{ub3} + k_3\big)\, K\, k_{b2}\ \Big(k_1\ \big(Ap\, k_{b4} + k_4 + k_{ub4}\big)\, K - k_4\, Ap\, P\, k_{b4}\Big)\Big)\, k_{b1} - \big(k_2\, Ap\, K\, k_{b2} - P\, k_3\ \big(Ap\, k_{b2} + k_2 + k_{ub2}\big)\big)\, k_{b4}\, k_{b3}\ \big(k_{ub1} + k_1\big)\, P\, k_4\Big)$$

$$T := -\frac{1}{\epsilon\, k_1\, k_3}\Big(\big(k_{ub4} + k_4\big)\, P^2\ \big(k_{ub2} + k_2\big)\ \Big(-k_2^{\,2}\, k_1^{\,2}\, c_2^{\,2}\, c_1\, Ap\, \epsilon^4 + k_1\, c_1\, c_2\ \Big(-k_3\ \Big(P\, c_3\ \big(Ap\, c_4 + 1\big)\, k_3 + k_2\ \big(Ap^2\, c_2\, c_4 + 2\, Ap\, c_4 + P\, c_3 + 1\big)\Big)\, k_1 + k_4\, Ap\, c_4\, k_2\ \big(Ap\, c_2\, k_2 + P\, k_3\, c_3\big)\Big)\, \epsilon^3 + k_1\ \Big(-c_1\, c_3\, k_3\ \big(Ap\, c_2 + 1\big)\ \big(1 + (P + Ap)\, c_4\big)\, k_1 + k_4\ \Big(-P\, c_1\, c_3\, k_3 + Ap\ \Big(\big(1 + (P + Ap)\, c_3\big)\, c_1 + c_3\Big)\, k_2\, c_2\Big)\, c_4\Big)\, k_3\, \epsilon^2 + k_4\, c_3\ \Big(-k_3\ \Big(1 + \big(Ap^2\, c_4 + 2\, Ap\big)\, c_2\Big)\, k_1 + c_2\, c_4\, Ap^2\, k_2\, k_4\Big)\, k_3\, c_4\, \epsilon - c_4^{\,2}\, Ap\, k_4^{\,2}\, k_3^{\,2}\, c_3\Big)\ \big(k_{ub1} + k_1\big)\ \big(k_{ub3} + k_3\big)\Big) \quad \textbf{(1)}$$

$$simplify\big(solve\big(PCon, P_{Total}\big)\big)$$

$$\frac{\big(K\, k_2\ \big(App\, c_3 + 1\big)\, c_2 + App\, P\, c_3\, c_4\, k_3\big)\, P}{K\, c_2\, k_2} \quad \textbf{(2)}$$

$$k_1 := 5 : k_2 := 2 : k_3 := 10 : k_4 := 1 : k_{b1} := 1 : k_{b2} := 100 : k_{b3} := 1 : k_{b4} := 200 : k_{ub1} := 100 : k_{ub2} := 1 : k_{ub3} := 100 : k_{ub4} := 1 :$$

$$c_1 := \frac{k_{b1}}{k_{ub1} + k_1} : c_2 := \frac{k_{b2}}{k_{ub2} + k_2} : c_3 := \frac{k_{b3}}{k_{ub3} + k_3} : c_4 := \frac{k_{b4}}{k_{ub4} + k_4} :$$

$$simplify(T)$$

$$\frac{1}{\epsilon}\left(960000\,P^2\left(\frac{55\,Ap\,\epsilon^4}{36}+\left(\frac{2200\,Ap^2}{3}+\left(P+\frac{275}{6}\right)Ap+\frac{P}{80}+\frac{11}{48}\right)\epsilon^3+\left(Ap^2+\left(P-\frac{107}{12}\right)Ap+\frac{3\,P}{80}+\frac{1}{3200}\right)\epsilon^2+\left(2100\,Ap^2+\frac{175}{4}\,Ap+\frac{21}{32}\right)\epsilon+\frac{105\,Ap}{8}\right)\right) \tag{3}$$

$$Ap$$

$$\frac{c_3\,App\,P\,k_3}{K\,k_2\,c_2} \tag{4}$$

$$collect\left(numer\left(simplify\left(\frac{k_4\,c_4\,c_3\,App\,k_3}{\epsilon^2\,k_2\,c_2\,k_1\,c_1}+\frac{c_3\,App\,k_3}{\epsilon\,k_2\,c_2}+App+\frac{k_4\,c_4\,c_3\,App\,k_3\,P_{Total}}{k_1\left(c_2\,k_2\left(c_3\,App+1\right)\epsilon+c_4\,c_3\,App\,k_3\right)}+\frac{c_2\,c_3\,App\,k_3\,\epsilon\,P_{Total}}{c_2\,k_2\left(c_3\,App+1\right)\epsilon+c_4\,c_3\,App\,k_3}+\frac{c_3\,App\,P_{Total}\,\epsilon\,k_2\,c_2}{c_2\,k_2\left(c_3\,App+1\right)\epsilon+c_4\,c_3\,App\,k_3}+\frac{c_4\,c_3\,App\,k_3\,P_{Total}}{c_2\,k_2\left(c_3\,App+1\right)\epsilon+c_4\,c_3\,App\,k_3}-A_{Total}\right)\right),\ App\right)$$

$$\left(c_1\,c_2^2\,c_3\,\epsilon^3\,k_1\,k_2^2+c_1\,c_2\,c_3^2\,\epsilon^2\,k_1\,k_2\,k_3+c_1\,c_2\,c_3\,c_4\,\epsilon^2\,k_1\,k_2\,k_3+c_1\,c_3^2\,c_4\,\epsilon\,k_1\,k_3^2+c_2\,c_3^2\,c_4\,\epsilon\,k_2\,k_3\,k_4+c_3^2\,c_4^2\,k_3^2\,k_4\right)App^2+\left(-c_1\,c_2^2\,c_3\,\epsilon^3\,k_1\,k_2^2\,A_{Total}+c_1\,c_2^2\,c_3\,\epsilon^3\,k_1\,k_2^2\,P_{Total}+c_1\,c_2^2\,c_3\,\epsilon^3\,k_1\,k_2\,k_3\,P_{Total}-c_1\,c_2\,c_3\,c_4\,\epsilon^2\,k_1\,k_2\,k_3\,A_{Total}+c_1\,c_2\,c_3\,c_4\,\epsilon^2\,k_1\,k_2\,k_3\,P_{Total}+c_1\,c_2\,c_3\,c_4\,\epsilon^2\,k_2\,k_3\,k_4\,P_{Total}+c_1\,c_2^2\,\epsilon^3\,k_1\,k_2^2+c_1\,c_2\,c_3\,\epsilon^2\,k_1\,k_2\,k_3+c_2\,c_3\,c_4\,\epsilon\,k_2\,k_3\,k_4\right)App-c_1\,c_2^2\,\epsilon^3\,k_1\,k_2^2\,A_{Total} \tag{5}$$

# Double site phosphorylation system (DSP) [enzyme combinations - same/different acting on the two modification sites]

## Features and requirements to obtain substrate and enzyme biphasic behavior in the maximally modified substrate

In this folder we detail analytical results pertaining to biphasic response in the maximally modified substrate in various models of the DSP (with common or different enzymes acting on the two modification sites). Specifically, we study each system's capacity to exhibit biphasic responses in the with changing total concentration of substrate (substrate biphasic) and total concentration of enzyme (enzyme biphasic).

A summary of the results is presented in the table below.

**nd Enzyme biphasic dose responses in the double site ordered modification system with common/different enz**

| System | Substrate Biphasic | Enzyme Biphasic |
|---|---|---|
| Common Kinase Seperate Phophatase | Present | Not possible |
| Separate Kinase Common Phosphatase | Not possible | Present with $K2_{Total}$ Not possible with $K1_{Total}$ |
| Separate Kinase Separate Phosphatase | Not possible | Not possible |

# Double site phosphorylation system (DSP) [common kinase and separate phosphatase]

## Features and requirements of obtaining biphasic response in the maximally modified substrate

In this file we analytically study the presence of enzyme and substrate biphasic in the maximally modified substrate form. In doing so we establish the following key results with regard to the behavior.

**Enzyme biphasic response (biphasic behavior in the dose response curve of App as $K_{Total}$ changes)**

1. The system is incapable of exhibiting enzyme biphasic dose response with increasing amounts of total kinase.

**Substrate biphasic response (biphasic behavior in the dose response curve of App as $A_{Total}$ changes)**

1. Substrate biphasic response is possible in App for any kinetic regime (i.e. for any choice of underlying kinetics the system is capable of exhibiting substrate biphasic dose response at some total concentration of enzymes)

We note that the key signature of biphasic behavior in the dose response curve of the system is the presence of a steady state of the system that satisfies the following condition.

$$\frac{dApp}{dK_{Total}} = 0 \text{ (for enzyme biphasic)}$$

$$\frac{dApp}{dA_{Total}} = 0 \text{ (for substrate biphasic)}$$

**Model DSP:** We first describe the model of double site phosphorylation with common kinase and different phosphatase acting on each modification site.

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart: with(LinearAlgebra): with(VectorCalculus): with(Student[LinearAlgebra]):
```

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP1 + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K:$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP2 + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP1) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P1:$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP2 - k_{b3} \cdot App \cdot P2:$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK:$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK:$$
$$dAppP2 := k_{b3} \cdot App \cdot P2 - (k_{ub3} + k_3) \cdot AppP2:$$
$$dApP1 := k_{b4} \cdot Ap \cdot P1 - (k_{ub4} + k_4) \cdot ApP1:$$

$$dK := -k_{b1}\, A\, K + (k_{ub1} w + k_1)\, AK - k_{b2}\, Ap\, K + (k_{ub2} + k_2)\, ApK:$$
$$dP1 := -k_{b4}\, Ap\, P1 + (k_{ub4} + k_4)\, ApP1:$$
$$dP2 := -k_{b3}\, App\, P2 + (k_{ub3} + k_3)\, AppP2:$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, P1Con, P2Con and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP2 + ApP1 - A_{Total}:$$
$$P1Con := P1 + ApP1 - P1_{Total}:$$
$$P2Con := P2 + AppP2 - P2_{Total}:$$
$$KCon := K + AK + ApK - K_{Total}:$$

We now solve the system described at steady state to obtain expression linking the steady state concentrations of the various species. Here we use the Maple command *solve*

```
assign(solve({dAK, dApK, dAppP2, dApP1}, {AK, ApK, AppP2, ApP1})):
assign(solve({dA, dApp}, {A, Ap}))
```

Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$k_{b1} := c_1 \cdot (k_1 + k_{ub1}): k_{b2} := c_2 \cdot (k_2 + k_{ub2}): k_{b3} := c_3 \cdot (k_3 + k_{ub3}): k_{b4} := c_4 \cdot (k_4 + k_{ub4}):$$

Once this is done, we again solve for the steady states of the two phosphatases using the conservation expression for the enzymes (P1Con and P2Con).

```
P1 := solve(P1Con, P1):
P2 := solve(P2Con, P2):
```

This results in the following expressions for the steady state concentrations of the various species in terms of

the steady state substrate concentration App and K.

$$A = \frac{P1_{Total}\, k_4\, c_4\, App\, P2_{Total}\, k_3\, c_3}{\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right) K \left(App\, c_3 + 1\right) k_1 c_1}$$

$$App = App$$

$$AK = \frac{P1_{Total}\, k_4\, c_4\, App\, P2_{Total}\, k_3\, c_3}{\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right) \left(App\, c_3 + 1\right) k_1}$$

$$ApK = \frac{App\, P2_{Total}\, k_3\, c_3}{\left(App\, c_3 + 1\right) k_2}$$

$$AppP2 = \frac{c_3\, App\, P2_{Total}}{App\, c_3 + 1}$$

$$ApP1 = \frac{c_4\, App\, P2_{Total}\, k_3\, c_3\, P1_{Total}}{\left(App\, c_3 + 1\right)\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right)}$$

Note that when App and  We now have two expressions, ACon and KCon - the conservation of the substrate and kinase, whose solution for the variables define the steady state of the system.

## Enzyme biphasic

In this subsection, we analytically show the absence of enzyme biphasic (in the full system). As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system.

$$\frac{dApp}{dK_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ACon = 0 (see below) whose solutions to the variables App and K define the steady state of the system.

$$\begin{aligned} ACon = {} & \frac{P1_{Total}\, k_4\, c_4\, App\, P2_{Total}\, k_3\, c_3}{\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right) K \left(App\, c_3 + 1\right) k_1 c_1} + \frac{App\, P2_{Total}\, k_3\, c_3}{\left(App\, c_3 + 1\right) K k_2 c_2} + App \\ & + \frac{P1_{Total}\, k_4\, c_4\, App\, P2_{Total}\, k_3\, c_3}{\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right) \left(App\, c_3 + 1\right) k_1} + \frac{App\, P2_{Total}\, k_3\, c_3}{\left(App\, c_3 + 1\right) k_2} + \frac{c_3\, App\, P2_{Total}}{App\, c_3 + 1} \\ & + \frac{c_4\, App\, P2_{Total}\, k_3\, c_3\, P1_{Total}}{\left(App\, c_3 + 1\right)\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right)} - A_{Total} \end{aligned}$$

$$KCon = K + \frac{P1_{Total}\, k_4\, c_4\, App\, P2_{Total}\, k_3\, c_3}{\left(\frac{App\, P2_{Total}\, c_3\, c_4\, k_3}{App\, c_3 + 1} + K k_2 c_2\right)\left(App\, c_3 + 1\right) k_1} + \frac{App\, P2_{Total}\, k_3\, c_3}{\left(App\, c_3 + 1\right) k_2} - K_{Total}$$

If we differentiate both these with respect to the total kinase concentration in the system, we get

$$\frac{dACon}{dK_{Total}} = 0 = \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$
$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial App} \cdot \frac{\partial App}{\partial K_{Total}} + \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

Now in order to show the absence of an enzyme biphasic response we use a proof by contradiction. We begin with the assumption that there exista a biphasic beahvior. In which case there should exist a steady state of the system where $\frac{dApp}{dK_{Total}} = 0$ is satisfied.

At this point then, the above expressions simply as follows

$$0 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$
$$1 = \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}, \frac{\partial KCon}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial K_{Total}}$ can be zero.

$$simplify(diff(ACon, App)) =$$
$$\left(k_2^2\, c_1 \left(App^2 k_2 c_3^2 + \left(\left(P2_{Total} + 2\, App\right) k_2 + P2_{Total} k_3\right) c_3 + k_2\right) K^3 k_1 \left(App\, c_3 + 1\right)^2 c_2^3 + \left(2\, K App^3 c_1 c_4 k_1 k_2 c_3^2 + App\left(\left(\left(4\, App\, c_4 k_1 + c_4 k_1 P1_{Total} + 2\, c_4 k_1 P2_{Total} + k_4 P1_{Total} c_4 + k_1\right) k_2 + 2\, P2_{Total} c_4 k_1 k_3\right) c_1 K + k_4 P1_{Total} c_4 k_2\right) c_3 + k_2 \left(c_1 \left(2\, App\, c_4 k_1 + c_4 k_1 P1_{Total} + k_4 P1_{Total} c_4 + k_1\right) K + k_4 P1_{Total} c_4\right)\right) k_3 k_2 P2_{Total} K c_3 \left(App\, c_3 + 1\right) c_2^2 + App\, k_3^2 \left(App^3 c_4 k_2 c_3^2 + App\left(\left(2\, App\, c_4 + c_4 P2_{Total} + 2\right) k_2 + P2_{Total} c_4 k_3\right) c_3 + k_2 \left(App\, c_4 + 2\right)\right) c_4 P2_{Total}^2 c_1 K k_1 c_3^2 c_2 + App^2 c_1 c_3^3 c_4^2 k_1 k_3^3 P2_{Total}^3\right) \Big/ \left(c_2 k_2 c_1 K k_1 \left(App\, P2_{Total} c_3 c_4 k_3 + K k_2 c_2 \left(App\, c_3 + 1\right)\right)^2 \left(App\, c_3 + 1\right)^2\right)$$
$$simplify(diff(KCon, App)) =$$
$$\left(c_3 P2_{Total} k_3 \left(App^2 \left(K c_2 \left(K c_2 k_1 + k_4 P1_{Total} c_4\right) k_2^2 + 2\, K P2_{Total} c_2 c_4 k_1 k_3 k_2 + P2_{Total}^2 c_4^2 k_1 k_3^2\right) c_3^2 + 2\, App\left(\left(K c_2 k_1 + k_4 P1_{Total} c_4\right) k_2 + P2_{Total} c_4 k_1 k_3\right) c_2 K k_2 c_3 + K c_2 \left(K c_2 k_1 + k_4 P1_{Total} c_4\right) k_2^2\right)\right) \Big/ \left(\left(App\, c_3 + 1\right)^2 k_1 \left(App\left(K k_2 c_2 + P2_{Total} c_4 k_3\right) c_3 + K k_2 c_2\right)^2 k_2\right)$$

We know that the denominator of

$\frac{\partial KCon}{\partial K}$ is non-zero (see below), thus this implies that $\frac{\partial K}{\partial K_{Total}}$ has to be non-zero in order to satisfy the second expression above.

$simplify(diff(KCon, K)) =$

$$\frac{1}{k_1 \left(App\ P2_{Total}\, c_3\, c_4\, k_3 + K\, k_2\, c_2 \left(App\ c_3 + 1\right)\right)^2} \left(K^2\, k_1\, k_2^2 \left(App\ c_3 + 1\right)^2 c_2^2 + 2\, c_3\, P2_{Total}\, App \left(App\ c_3 + 1\right) k_3 \left(K\, k_1 - \frac{P1_{Total}\, k_4}{2}\right) k_2\, c_4\, c_2 + App^2\, c_3^2\, c_4^2\, k_1\, k_3^2\, P2_{Total}^2\right)$$

This insight then informs us that in order to satisfy the first expression, $\frac{\partial ACon}{\partial K}$ must be equal to zero.

$Condition := simplify(diff(ACon, K))$

$$Condition := -\left(App\, k_3 \left(k_2^2\, K \left(c_1 \left(P1_{Total} \left(k_4 + k_1\right) c_4 + k_1\right) K + 2\, k_4\, P1_{Total}\, c_4\right) \left(App\ c_3 + 1\right)^2 c_2^2 + 2\, App \left(K\, c_1\, k_1 + \frac{k_4\, P1_{Total}\, c_4}{2}\right) k_3\, k_2\, c_4\, P2_{Total}\, c_3 \left(App\ c_3 + 1\right) c_2 + App^2\, c_1\, c_3^2\, c_4^2\, k_1\, k_3^2\, P2_{Total}^2\right) P2_{Total}\, c_3\right) \Big/ \left(c_2\, k_2\, c_1\, K^2\, k_1 \left(App\ P2_{Total}\, c_3\, c_4\, k_3 + K\, k_2\, c_2 \left(App\ c_3 + 1\right)\right)^2 \left(App\ c_3 + 1\right)\right) \quad \textbf{(1.1)}$$

However, from above we can see that this is not possible as the expression is always negative irrespective of choice of parameters or steady state concentrations of App and K.

**Thus we have a contradiction, indicating that the system is incapable of exhibiting enzyme biphasic response in steady state concentration of App with total kinase concentration.**

## Substrate biphasic

As seen in the main text figure N-5, substrate biphasic response is need possible in the ordered double site with common kinase and separate phosphatase. In this section we illustrate analytically the features of such biphasic responses and characterize the kinetic parameter dependency of the behavior.

In order to do this, we begin with the assumption that there exists a substrate biphasic response in App. i. e. there exists a steady state of the system where $\frac{dApp}{dA_{Total}} = 0$ is satisfied.

We now have two remaining conservations, KCon = 0 & ACon = 0 (see below) whose solutions to the variables App and K define the steady state of the system.

$ACon =$

$$\frac{P1_{Total}\, k_4\, c_4\, App\ P2_{Total}\, k_3\, c_3}{\left(\frac{App\ P2_{Total}\, c_3\, c_4\, k_3}{App\ c_3 + 1} + K\, k_2\, c_2\right) K \left(App\ c_3 + 1\right) k_1\, c_1} + \frac{App\ P2_{Total}\, k_3\, c_3}{\left(App\ c_3 + 1\right) K\, k_2\, c_2} + App$$

$$+\frac{P1_{Total}\,k_4\,c_4\,App\,P2_{Total}\,k_3\,c_3}{\left(\frac{App\,P2_{Total}\,c_3\,c_4\,k_3}{App\,c_3+1}+K\,k_2\,c_2\right)\left(App\,c_3+1\right)k_1}+\frac{App\,P2_{Total}\,k_3\,c_3}{\left(App\,c_3+1\right)k_2}+\frac{c_3\,App\,P2_{Total}}{App\,c_3+1}$$
$$+\frac{c_4\,App\,P2_{Total}\,k_3\,c_3\,P1_{Total}}{\left(App\,c_3+1\right)\left(\frac{App\,P2_{Total}\,c_3\,c_4\,k_3}{App\,c_3+1}+K\,k_2\,c_2\right)}-A_{Total}$$

$$KCon=K+\frac{P1_{Total}\,k_4\,c_4\,App\,P2_{Total}\,k_3\,c_3}{\left(\frac{App\,P2_{Total}\,c_3\,c_4\,k_3}{App\,c_3+1}+K\,k_2\,c_2\right)\left(App\,c_3+1\right)k_1}+\frac{App\,P2_{Total}\,k_3\,c_3}{\left(App\,c_3+1\right)k_2}-K_{Total}$$

If we differentiate both these with respect to the total substrate concentration in the system, we get

$$\frac{dACon}{dA_{Total}}=0=\frac{\partial ACon}{\partial A_{Total}}+\frac{\partial ACon}{\partial App}\cdot\frac{\partial App}{\partial A_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$
$$\frac{dKCon}{dA_{Total}}=0=\frac{\partial KCon}{\partial App}\cdot\frac{\partial App}{\partial A_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

At this point then, the above expressions simply as follows (evaluated at the biphasic peak where $\frac{dApp}{dA_{Total}}=0$

$$1=\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$
$$0=\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}$, $\frac{\partial KCon}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$simplify(diff(ACon, App)) =$
$$\left(\left(App^2\,k_2\,c_3^{\,2}+\left(\left(P2_{Total}+2\,App\right)k_2+P2_{Total}\,k_3\right)c_3+k_2\right)k_1\,c_1\,k_2^{\,2}\,K^3\left(App\,c_3+1\right)^2c_2^{\,3}\right.$$
$$+\left(2\,K\,App^3\,c_1\,c_4\,k_1\,k_2\,c_3^{\,2}+App\left(\left(\left(4\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+2\,c_4\,k_1\,P2_{Total}+k_4\,P1_{Total}\,c_4\right.\right.\right.$$
$$\left.\left.+k_1\right)k_2+2\,P2_{Total}\,c_4\,k_1\,k_3\right)c_1\,K+k_4\,P1_{Total}\,c_4\,k_2\Big)c_3+\left(c_1\left(2\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}\right.\right.$$
$$\left.\left.+k_4\,P1_{Total}\,c_4+k_1\right)K+k_4\,P1_{Total}\,c_4\right)k_2\Big)k_3\,P2_{Total}\,c_3\,k_2\,K\left(App\,c_3+1\right)c_2^{\,2}$$
$$+App\,k_3^{\,2}\,k_1\,P2_{Total}^{\,2}\,c_3^{\,2}\,c_1\,K\left(App^3\,c_4\,k_2\,c_3^{\,2}+App\left(\left(2\,App\,c_4+c_4\,P2_{Total}+2\right)k_2\right.\right.$$
$$\left.\left.+P2_{Total}\,c_4\,k_3\right)c_3+k_2\left(App\,c_4+2\right)\right)c_4\,c_2+App^2\,c_1\,c_3^{\,3}\,c_4^{\,2}\,k_1\,k_3^{\,3}\,P2_{Total}^{\,3}\Big)\Big/$$
$$\left.\left(k_1\,c_1\,k_2\left(App\,P2_{Total}\,c_3\,c_4\,k_3+K\,k_2\,c_2\left(App\,c_3+1\right)\right)^2K\left(App\,c_3+1\right)^2c_2\right)\right.$$

$simplify(diff(KCon, App)) =$
$$\left(P2_{Total}\left(App^2\left(K\,c_2\left(K\,c_2\,k_1+k_4\,P1_{Total}\,c_4\right)k_2^{\,2}+2\,K\,P2_{Total}\,c_2\,c_4\,k_1\,k_3\,k_2+P2_{Total}^{\,2}\,c_4^{\,2}\,k_1\,k_3^{\,2}\right)c_3^{\,2}\right.\right.$$

$$+ 2\,App\,k_2\,c_2\left(\left(Kc_2k_1 + k_4\,P1_{Total}\,c_4\right)k_2 + P2_{Total}\,c_4\,k_1\,k_3\right)Kc_3 + Kc_2\left(Kc_2k_1 + k_4\,P1_{Total}\,c_4\right)k_2^2\right)k_3\,c_3\Big)\Big/\Big(\left(App\,c_3+1\right)^2\left(App\left(Kk_2\,c_2 + P2_{Total}\,c_4\,k_3\right)c_3 + Kk_2\,c_2\right)^2 k_2\,k_1\Big)$$

We know that the denominator of $\frac{\partial ACon}{\partial K}$ is non-zero (see below), thus this implies that $\frac{\partial K}{\partial A_{Total}}$ has to be non-zero in order to satisfy the first expression above.

$simplify(diff(ACon, K))$

$$-\Big(App\,k_3\Big(k_2^2\,K\left(App\,c_3+1\right)^2\left(\left(P1_{Total}\left(k_4+k_1\right)c_4 + k_1\right)c_1K + 2\,k_4\,P1_{Total}\,c_4\right)c_2^2 + 2\,App\,k_3\,P2_{Total}\,c_3\,k_2\,c_4\left(Kc_1k_1 + \frac{k_4\,P1_{Total}\,c_4}{2}\right)\left(App\,c_3+1\right)c_2 + App^2\,c_1\,c_3^2\,c_4^2\,k_1\,k_3^2\,P2_{Total}^2\Big)P2_{Total}\,c_3\Big)\Big/\Big(k_1\,c_1\,k_2\left(App\,P2_{Total}\,c_3\,c_4\,k_3 + Kk_2\,c_2\left(App\,c_3+1\right)\right)^2K^2\left(App\,c_3+1\right)c_2\Big) \tag{2.1}$$

Hence in order for the biphasic response to exist, $\frac{\partial KCon}{\partial K}$ has to be zero. This expression is shown below (stored as the variable *Condition*).

$Condition := simplify(diff(KCon, K))$

$$Condition := \Big(K^2k_1k_2^2\left(App\,c_3+1\right)^2c_2^2 + 2\,App\left(App\,c_3+1\right)P2_{Total}\left(Kk_1 - \frac{P1_{Total}\,k_4}{2}\right)k_2\,c_4\,k_3\,c_3\,c_2 + App^2\,c_3^2\,c_4^2\,k_1\,k_3^2\,P2_{Total}^2\Big)\Big/\Big(\left(App\,P2_{Total}\,c_3\,c_4\,k_3 + Kk_2\,c_2\left(App\,c_3+1\right)\right)^2k_1\Big) \tag{2.2}$$

Now writing the numerator of *Conditon* as a polynomial in K (as shown below), reveals that irrespective of the steady state concentrations of App the parameter values, the coeffecient of the first and second exponent of K are positive.

$collect(numer(Condition), K)$

$$\left(App^2\,c_2^2\,c_3^2\,k_1\,k_2^2 + 2\,App\,c_2^2\,c_3\,k_1\,k_2^2 + c_2^2\,k_1\,k_2^2\right)K^2 + \left(2\,App^2\,c_2\,c_3^2\,c_4\,k_1\,k_2\,k_3\,P2_{Total} + 2\,App\,c_2\,c_3\,c_4\,k_1\,k_2\,k_3\,P2_{Total}\right)K - App^2\,c_2\,c_3^2\,c_4\,k_2\,k_3\,k_4\,P1_{Total}\,P2_{Total} + App^2\,c_3^2\,c_4^2\,k_1\,k_3^2\,P2_{Total}^2 - P1_{Total}\,k_4\,c_4\,App\,P2_{Total}\,k_3\,c_3\,k_2\,c_2 \tag{2.3}$$

The constant term however depending on the concentration of App can be negative or positive. Now this polynomial (*Condition*) has to be zero for the biphasic to exist, implying the constant term has to necessarily be negative (which will guarentee the existence of a positive concentration of K that is a root of the polynomial). This requirement of negativity for the constant term provides us with the following condition involving the concentration of App

$$App < solve\left(-App^2\,c_2\,c_3^2\,c_4\,k_2\,k_3\,k_4\,P1_{Total}\,P2_{Total} + App^2\,c_3^2\,c_4^2\,k_1\,k_3^2\,P2_{Total}^2 - P1_{Total}\,k_4\,c_4\,App\,P2_{Total}\,k_3\,c_3\,k_2\,c_2, App\right)[2]$$

$$App < -\frac{c_2 k_2 k_4 P1_{Total}}{c_3 \left(c_2 k_2 k_4 P1_{Total} - P2_{Total} c_4 k_1 k_3\right)} \tag{2.4}$$

Depending on whether $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3$ is positive or negative, this condition is trvially satified or isn't.
Now, if this catalytic constant grouping ( $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3$) is negative, then the constant term is negative irrespective of the steady state concentration of App and thus the whole expression from earlier guarentees substrate biphasic response at some total concentration. However, if the grouping ( $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3$) is positive, then the constant term can still be negative, however in this case App (at the biphasic peak) is necessarily to be less than

$$\frac{c_2 k_2 k_4 P1_{Total}}{c_3 \left(-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3\right)}.$$

This is however still feasible since, App and epsilon can take any positive concentration and we can find a suitable total concentration of substrate and enzyme that will accommodate it as steady state concentrations (using ACon and KCon which we are yet to solve).

$$\begin{aligned} ACon = {} & \frac{P1_{Total} k_4 c_4 App\, P2_{Total} k_3 c_3}{\left(\frac{App\, P2_{Total} c_3 c_4 k_3}{App\, c_3 + 1} + K k_2 c_2\right) K \left(App\, c_3 + 1\right) k_1 c_1} + \frac{App\, P2_{Total} k_3 c_3}{\left(App\, c_3 + 1\right) K k_2 c_2} + App \\ & + \frac{P1_{Total} k_4 c_4 App\, P2_{Total} k_3 c_3}{\left(\frac{App\, P2_{Total} c_3 c_4 k_3}{App\, c_3 + 1} + K k_2 c_2\right) \left(App\, c_3 + 1\right) k_1} + \frac{App\, P2_{Total} k_3 c_3}{\left(App\, c_3 + 1\right) k_2} + \frac{c_3\, App\, P2_{Total}}{App\, c_3 + 1} \\ & + \frac{c_4\, App\, P2_{Total} k_3 c_3 P1_{Total}}{\left(App\, c_3 + 1\right) \left(\frac{App\, P2_{Total} c_3 c_4 k_3}{App\, c_3 + 1} + K k_2 c_2\right)} - A_{Total} \end{aligned}$$

$$KCon = K + \frac{P1_{Total} k_4 c_4 App\, P2_{Total} k_3 c_3}{\left(\frac{App\, P2_{Total} c_3 c_4 k_3}{App\, c_3 + 1} + K k_2 c_2\right) \left(App\, c_3 + 1\right) k_1} + \frac{App\, P2_{Total} k_3 c_3}{\left(App\, c_3 + 1\right) k_2} - K_{Total}$$

Thus, what this implies is that, irrespective of the underlying grouping (or the sign of the catalytic constant grouping $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3$), substrate biphasic behavior is guaranteed to exist for some total concentration of substrate and enzyme in the system.

However, we wish to note here that depending on the sign of the catalytic constant grouping $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3$ it is either more likely or less likely to observe substrate biphasic response in App with total amount of substrate in the system. Further, depending on the sign of the catalytic constant grouping (if $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3 > 0$), the peak concentration of the biphasic response capable of being observed is capped at $\frac{c_2 k_2 k_4 P1_{Total}}{c_3 \left(-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3\right)}$ which is not the case when ( $-c_2 k_2 k_4 P1_{Total} + P2_{Total} c_4 k_1 k_3 < 0$) when App can take any value at biphasic peak (given that the total amounts are completely flexible.

**Additionally, now we show how the parameter values used to generate main text figure N-5, satisfy the requirements.**

$k_1 := 3.5 : k_2 := 50 : k_3 := 100 : k_4 := 100 :$
$k_{b1} := 20 : k_{b2} := 75 : k_{b3} := 50 : k_{b4} := 30 :$
$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 : P1_{Total} := 1 : P2_{Total} := 1 :$
$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$

$App := 2 :$
$K := solve(Condition)[1] = 2.195303469$

$ACon = 6.526495239 - A_{Total}$
$KCon = 5.586607136 - K_{Total}$

# Double site phosphorylation system (DSP) [separate kinase and common phosphatase]

## Features and requirements of obtaining biphasic response in the maximally modified substrate

In this file we analytically prove the absence of enzyme and substrate biphasic in the maximally modified substrate form (App), with changes in $K1_{Total}$ and $A_{Total}$ respectively.

Note: Enzyme biphasic response in the dose response behavior of App with $K2_{Total}$ is possible and was computationally demonstrated in figure 2 for a choice of kinetic parameters.

We also recapitulate that the key signature of biphasic behavior in the dose response curve of the system is the presence of a steady state of the system that satisfies the following condition.

$\frac{\partial App}{\partial K_{Total}} = 0$ (for enzyme biphasic, $K1_{Total}$ or $K2_{Total}$ depending on the enzyme dose considered)

$\frac{\partial App}{\partial A_{Total}} = 0$ (for substrate biphasic)

**Model DSP with separate kinase and common phosphatse:** We first describe the model of double site phosphorylation with different kinase and common phosphatase acting on each modification site.

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK1 - k_{b1} \cdot A \cdot K1 :$$
$$dAp := k_1 \cdot AK1 + k_3 \cdot AppP + k_{ub2} \cdot (ApK2) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K2) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK2 + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK1 := k_{b1} \cdot A \cdot K1 - \left(k_{ub1} + k_1\right) \cdot AK1 :$$
$$dApK2 := k_{b2} \cdot Ap \cdot K2 - \left(k_{ub2} + k_2\right) \cdot ApK2 :$$
$$dAppP := k_{b3} \cdot App \cdot P - \left(k_{ub3} + k_3\right) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dK1 := -k_{b1}\, A\, K1 + \left(k_{ub1} + k_1\right) AK1 :$$
$$dK2 := -k_{b2}\, Ap\, K2 + \left(k_{ub2} + k_2\right) ApK2 :$$
$$dP := -k_{b3}\, App\, P + \left(k_{ub3} + k_3\right) AppP - k_{b4}\, Ap\, P + \left(k_{ub4} + k_4\right) ApP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon, K1Con and K2Con for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK1 + ApK2 + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$K1Con := K1 + AK1 - K1_{Total} :$$
$$K2Con := K2 + ApK2 - K2_{Total} :$$

We now solve the system described at steady state to obtain expression linking the steady state concentrations of the various species. Here we use the Maple command *solve* to solve the equations for a given variable as shown below. We pursue this to finally obtain the steady state concentrations of most species in terms of App , K1 and a ratio $\epsilon$ (defined below).

$$AK1 := solve(dAK1, AK1) = \frac{k_{b1}\, A\, K1}{k_{ub1} + k_1}$$
$$assign(solve(\{dApK2, dAppP, dApP\}, \{ApK2, AppP, ApP\}))$$
$$assign(solve(\{dA, dAp\}, \{A, Ap\}))$$

Further we now introduce a ratio, $\epsilon$ = K2/P (defined as the ratio of the free enzymes). Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy

tractability of the expressions obtained.

$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) : k_{b2} := c_2 \cdot (k_2 + k_{ub2}) : k_{b3} := c_3 \cdot (k_3 + k_{ub3}) : k_{b4} := c_4 \cdot (k_4 + k_{ub4}) :$
$K2 := \text{epsilon} \cdot P :$

Once this is done, we again solve for the steady state of the phosphatase using the conservation expression for the enzyme (PCon).

$P := simplify(solve(PCon, P))$

$$P := \frac{P_{Total}\, \epsilon\, k_2\, c_2}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3} \tag{1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration App and $\epsilon$

$$A := simplify(A) = \frac{App\, P_{Total}\, k_3\, k_4\, c_3\, c_4}{(c_2\, k_2\, (App\, c_3 + 1)\, \epsilon + c_4\, App\, k_3\, c_3)\, c_1\, K1\, k_1}$$

$$Ap = \frac{App\, k_3\, c_3}{\epsilon\, k_2\, c_2}$$

$$AK1 = \frac{App\, P_{Total}\, k_3\, k_4\, c_3\, c_4}{(c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3)\, k_1}$$

$$ApK2 = \frac{c_2\, App\, k_3\, c_3\, \epsilon\, P_{Total}}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3}$$

$$AppP = \frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3}$$

$$ApP = \frac{c_4\, App\, k_3\, c_3\, P_{Total}}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3}$$

Note that when App, K1 $\epsilon$ are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of App, K1 and $\epsilon$. We now have three expressions, ACon, K1Con and K2Con - the conservation of the substrate and kinases, whose solution for the variables define the steady state of the system.

$$ACon = \frac{App\, P_{Total}\, k_3\, k_4\, c_3\, c_4}{(c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3)\, c_1\, K1\, k_1} + \frac{App\, k_3\, c_3}{\epsilon\, k_2\, c_2} + App + \frac{App\, P_{Total}\, k_3\, k_4\, c_3\, c_4}{(c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3)\, k_1} + \frac{c_2\, App\, k_3\, c_3\, \epsilon\, P_{Total}}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3} + \frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3} + \frac{c_4\, App\, k_3\, c_3\, P_{Total}}{c_2\, k_2\, (c_3\, App + 1)\, \epsilon + c_4\, App\, k_3\, c_3} - A_{Total}$$

$K1Con =$

$$K2Con = \frac{K1 + \frac{App\, P_{Total}\, k_3\, k_4\, c_3\, c_4}{\left(c_2\, k_2\, \left(c_3\, App + 1\right) \epsilon + c_4\, App\, k_3\, c_3\right) k_1} - K1_{Total}}{c_2\, k_2\, \left(App\, c_3 + 1\right) \epsilon + c_4\, App\, k_3\, c_3} + \frac{\epsilon^2\, P_{Total}\, k_2\, c_2}{c_2\, k_2\, \left(App\, c_3 + 1\right) \epsilon + c_4\, App\, k_3\, c_3} + \frac{c_2\, App\, k_3\, c_3\, \epsilon\, P_{Total}}{c_2\, k_2\, \left(App\, c_3 + 1\right) \epsilon + c_4\, App\, k_3\, c_3} - K2_{Total}$$

## Substrate biphasic (Absence)

Now in order to show the absence of substrate biphasic response in the maximally modified substrate form (App) with changing $A_{Total}$, we use a proof by contradiction. Thus we begin with the assumption that substrate biphasic exists, meaning there exists a steady state of the system where $\frac{\partial App}{\partial A_{Total}}$ equals zero.

Now if we were to differentiate ACon, K1Con and K2Con with $A_{Total}$, we obtain the following (note that K2Con is a function of App and $\epsilon$ only)

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}}$$
$$\frac{dK1Con}{dA_{Total}} = 0 = \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}} + \frac{\partial K1Con}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial K1Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$
$$\frac{dK2Con}{dA_{Total}} = 0 = \frac{\partial K2Con}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial K2Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

Now, since we assume that there exists substrate biphasic, we can simplify these expressions further as shown below

$$1 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}}$$
$$0 = \frac{\partial K1}{\partial A_{Total}} + \frac{\partial K1Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$
$$0 = \frac{\partial K2Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}$, $\frac{\partial K2Con}{\partial App}$ and $\frac{\partial K1Con}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$simplify(diff(ACon, App))$

$$\left(k_2^2\, K1\, c_2^3\, k_1 \left(\left(1 + App^2\, c_3^2 + \left(P_{Total} + 2\, App\right) c_3\right) k_2 + P_{Total}\, c_3\, k_3\right) c_1\, \epsilon^3 \right. \qquad \textbf{(1.1.1)}$$

$$+ k_3 k_2^2 c_2^2 c_3 \left(K1\ App^2 c_1 c_3^2 k_1 + 2\ K1\ App\ c_1 k_1 \left(App\ c_4 + 1\right) c_3 + K1 \left(\left(2\ App\ c_4 + P_{Total} c_4 + 1\right) k_1 + c_4 k_4 P_{Total}\right) c_1 + c_4 k_4 P_{Total}\right) \epsilon^2 + 2\ c_4 k_3^2 k_2 K1 \left(App\ c_3 + \frac{1}{2} App\ c_4 + 1\right) App\ c_2 k_1 c_1 c_3^2 \epsilon + App^2 K1\ c_1 c_3^3 c_4^2 k_1 k_3^3\Big) \Big/ \left(\epsilon k_2 K1 \left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 c_2 k_1 c_1\right)$$

$$simplify(diff(K2Con, App)) = -\frac{P_{Total} \left(\left(\epsilon k_2 - k_3\right) c_2 + c_4 k_3\right) c_2 \epsilon^2 c_3 k_2}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2}$$

$$simplify(diff(K1Con, App)) = \frac{P_{Total} k_3 k_4 c_3 c_4 c_2 \epsilon k_2}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 k_1}$$

We know that the denominator of $\frac{\partial K2Con}{\partial \epsilon}$ is non-zero (see below), thus this implies that $\frac{\partial \epsilon}{\partial A_{Total}}$ has to be zero in order to satisfy the first expressionabove.

$$simplify(diff(K2Con, \epsilon)) = \frac{\left(\epsilon^2 k_2^2 c_2 \left(App\ c_3 + 1\right) + 2\ App\ c_3 c_4 \epsilon k_2 k_3 + App^2 c_3^2 c_4 k_3^2\right) P_{Total} c_2}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2}$$

This insight then informs us that in order to satisfy the second expression, $\frac{\partial K1}{\partial A_{Total}}$ must be equal to zero. (Again this simplification is possible since the denominator of $\frac{\partial K1Con}{\partial \epsilon}$ is non-zero (see below).

$$simplify(diff(K1Con, \text{epsilon})) = -\frac{App\ P_{Total} k_3 k_4 c_3 c_4 c_2 k_2 \left(App\ c_3 + 1\right)}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 k_1}$$

Put together, the fact that $\frac{\partial K1}{\partial A_{Total}}$ and $\frac{\partial \epsilon}{\partial A_{Total}}$ are both zero provides a contradiction with the requirements of the first expression (obtained upon differentiation of the total substrate concentration equation).

## Note

This assertion is possible since the denominators of $\frac{\partial ACon}{\partial K1}$ and $\frac{\partial ACon}{\partial \epsilon}$ are both non-zero (see below).

$$simplify(diff(ACon, K1)) = -\frac{App\ P_{Total} k_3 k_4 c_3 c_4}{K1^2 \left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right) k_1 c_1}$$

$$simplify(diff(ACon, \text{epsilon})) = -\Big(k_3 App \Big(k_2 c_2^2 \Big(\Big(K1 \left(App^2 k_1 c_3^2 + App \left(c_4 k_4 P_{Total} + 2 k_1\right) c_3 + P_{Total} \left(k_4 + k_1\right) c_4\right.$$

$$+ k_1\Big) c_1 + k_4 P_{Total} c_4 \left(App\, c_3 + 1\right)\Big) k_2 - K1\, P_{Total}\, App\, c_1 c_3 c_4 k_1 k_3\Big) \epsilon^2$$
$$+ 2\, K1\, App\, c_1 c_2 c_3 c_4 k_1 k_2 k_3 \left(App\, c_3 + 1\right) \epsilon + App^2\, K1\, c_1 {c_3}^2 {c_4}^2 k_1 {k_3}^2\Big) c_3\Big) \Big/$$
$$\left(\epsilon^2 k_2 K1 \left(c_2 k_2 \left(App\, c_3 + 1\right) \epsilon + c_4\, App\, k_3 c_3\right)^2 c_2 k_1 c_1\right)$$

**Thus we have a contradiction, indicating that the system is incapable of exhibiting substrate biphasic response in steady state concentration of App with total substrate concentration.**

## Biphasic with K1-Total (Absence)

Now in order to show the absence of enzyme biphasic response in the maximally modified substrate form (App) with changing $K1_{Total}$, we use a proof by contradiction. Thus we begin with the assumption that enzyme biphasic exists, meaning there exists a steady state of the system where $\frac{\partial App}{\partial K1_{Total}}$ equals zero.

Now if we were to differentiate ACon, K1Con and K2Con with $K1_{Total}$, we obtain the following (note that K2Con is a function of App and $\epsilon$ only)

$$\frac{dACon}{dK1_{Total}} = 0 = \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial K1_{Total}} + \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial K1_{Total}}$$
$$\frac{dK1Con}{dK1_{Total}} = 0 = \frac{\partial K1Con}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial App} \cdot \frac{\partial App}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}}$$
$$\frac{dK2Con}{dK1_{Total}} = 0 = \frac{\partial K2Con}{\partial App} \cdot \frac{\partial App}{\partial K1_{Total}} + \frac{\partial K2Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}}$$

Now, since we assume that there exists substrate biphasic, we can simplify these expressions further as shown below

$$0 = \frac{\partial ACon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial K1_{Total}}$$
$$1 = \frac{\partial K1}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}}$$
$$0 = \frac{\partial K2Con}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial K1_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}$, $\frac{\partial K2Con}{\partial App}$ and $\frac{\partial K1Con}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$simplify(diff(ACon, App))$

$$\left(k_2^2\, K1\, {c_2}^3 k_1 \left(\left(1 + App^2 {c_3}^2 + \left(P_{Total} + 2\, App\right) c_3\right) k_2 + P_{Total} c_3 k_3\right) c_1 \epsilon^3\right. \qquad \textbf{(2.1.1)}$$

$$+ k_3 k_2^2 c_2^2 c_3 \left(K1\ App^2 c_1 c_3^2 k_1 + 2\ K1\ App\ c_1 k_1 \left(App\ c_4 + 1\right) c_3 + K1 \left(\left(2\ App\ c_4 + P_{Total} c_4 + 1\right) k_1 + c_4 k_4 P_{Total}\right) c_1 + c_4 k_4 P_{Total}\right) \epsilon^2 + 2\ c_4 k_3^2 k_2 K1 \left(App\ c_3 + \frac{1}{2} App\ c_4 + 1\right) App\ c_2 k_1 c_1 c_3^2 \epsilon + App^2 K1\ c_1 c_3^3 c_4^2 k_1 k_3^3\Big) \Big/ \left(\epsilon k_2 K1 \left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 c_2 k_1 c_1\right)$$

$$simplify(diff(K2Con, App)) = -\frac{P_{Total} \left(\left(\epsilon k_2 - k_3\right) c_2 + c_4 k_3\right) c_2 \epsilon^2 c_3 k_2}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2}$$

$$simplify(diff(K1Con, App)) = \frac{P_{Total} k_3 k_4 c_3 c_4 c_2 \epsilon k_2}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 k_1}$$

Now from equation 3 above (obtained from differentiating K2Con), we can observe that $\frac{\partial \epsilon}{\partial K1_{Total}}$ is zero, (since $\frac{\partial K2Con}{\partial \epsilon}$ cannot be equal to zero, see note below).

This implies that (from analyzing equation 1 above, obtained from differentiating ACon), $\frac{\partial K1}{\partial K1_{Total}}$ is also equal to zero since $\frac{\partial ACon}{\partial \epsilon}$ has a non-zero denominator, and $\frac{\partial ACon}{\partial K1}$ is nonzero always (see note below).

***Note***

$$simplify(diff(K2Con, \text{epsilon})) = \frac{c_2 \left(\epsilon^2 k_2^2 c_2 \left(App\ c_3 + 1\right) + 2\ App\ c_3 c_4 \epsilon k_2 k_3 + App^2 c_3^2 c_4 k_3^2\right) P_{Total}}{\left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2}$$

$$simplify(diff(ACon, K1)) = -\frac{App\ P_{Total} k_3 k_4 c_3 c_4}{c_1 k_1 \left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right) K1^2}$$

$$simplify(diff(ACon, \text{epsilon})) = -\Big(k_3 App\ c_3 \left(\left(\left(K1 \left(App^2 k_1 c_3^2 + App \left(c_4 k_4 P_{Total} + 2 k_1\right) c_3 + P_{Total} \left(k_4 + k_1\right) c_4 + k_1\right) c_1 + k_4 P_{Total} c_4 \left(App\ c_3 + 1\right)\right) k_2 - K1\ P_{Total} App\ c_1 c_3 c_4 k_1 k_3\right) k_2 c_2^2 \epsilon^2 + 2\ K1\ App\ c_1 c_2 c_3 c_4 k_1 k_2 k_3 \left(App\ c_3 + 1\right) \epsilon + App^2 K1\ c_1 c_3^2 c_4^2 k_1 k_3^2\right)\Big) \Big/ \left(c_1 \epsilon^2 k_1 \left(c_2 k_2 \left(App\ c_3 + 1\right) \epsilon + c_4 App\ k_3 c_3\right)^2 K1\ k_2 c_2\right)$$

Thus, from the above inferences ( $\frac{\partial K1}{\partial K1_{Total}}$ and $\frac{\partial \epsilon}{\partial K1_{Total}}$ are both zero), we find that a contradiction in equation 2 (obtained from differentiating K1Con)

**Thus contradiction. The conditions can't be satisfied implying that a biphasic response in App is not possible with total enzyme concentration $K1_{Total}$.**

## Biphasic with K2-Total (Presence)

$$T := collect(simplify(numer(diff(ACon, \text{epsilon}) - diff(ACon, K1)\cdot diff(K1Con, \text{epsilon}))), K1)$$

$$T := -\left(k_2^{\,2}\left(App\,c_3+1\right)c_2^{\,3}\,k_1\left(\left(c_1\left(App^2\,c_3^{\,2}+2\,App\,c_3+c_4\,P_{Total}+1\right)k_1\right.\right.\right. \tag{3.1}$$
$$\left.\left.\left.+k_4\,P_{Total}\,c_4\left(App\,c_3+1\right)c_1\right)k_2-P_{Total}\,App\,c_1\,c_3\,c_4\,k_1\,k_3\right)\epsilon^3$$
$$+c_3\,c_4\left(\left(c_1\left(3\,App^2\,c_3^{\,2}+6\,App\,c_3+c_4\,P_{Total}+3\right)k_1^{\,2}+k_4\,P_{Total}\,c_4\left(App\,c_3\right.\right.\right.$$
$$\left.\left.\left.+1\right)c_1\,k_1\right)k_2-P_{Total}\,App\,c_1\,c_3\,c_4\,k_1^{\,2}\,k_3\right)k_2\,k_3\,c_2^{\,2}\,App\,\epsilon^2$$
$$+3\,App^2\,c_1\,c_2\,c_3^{\,2}\,c_4^{\,2}\,k_1^{\,2}\,k_2\,k_3^{\,2}\left(App\,c_3+1\right)\epsilon+App^3\,c_1\,c_3^{\,3}\,c_4^{\,3}\,k_1^{\,2}\,k_3^{\,3}\right)c_3\,k_3\,App\,K1^2$$
$$-\left(k_2^{\,3}\left(App\,c_3+1\right)^2c_2^{\,3}\,k_1\,k_4\,P_{Total}\,c_4\,\epsilon^3+c_3\,c_4^{\,2}\,k_4\,P_{Total}\left(App\,c_3\right.\right.$$
$$\left.\left.+1\right)k_1\,k_2^{\,2}\,k_3\,c_2^{\,2}\,App\,\epsilon^2\right)c_3\,k_3\,App\,K1-App^2\,P_{Total}^{\,2}\,k_3^{\,2}\,k_4^{\,2}\,c_3^{\,2}\,c_4^{\,2}\,c_2^{\,2}\,k_2^{\,2}\left(App\,c_3$$
$$+1\right)\epsilon^2$$

$$simplify\left(-\left(k_2^{\,2}\left(App\,c_3+1\right)c_2^{\,3}\,k_1\left(\left(c_1\left(App^2\,c_3^{\,2}+2\,App\,c_3+c_4\,P_{Total}+1\right)k_1\right.\right.\right.\right.$$
$$\left.\left.\left.\left.+k_4\,P_{Total}\,c_4\left(App\,c_3+1\right)c_1\right)k_2-P_{Total}\,App\,c_1\,c_3\,c_4\,k_1\,k_3\right)\epsilon^3+c_3\,c_4\left(\left(c_1\left(3\,App^2\,c_3^{\,2}\right.\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\left.+6\,App\,c_3+c_4\,P_{Total}+3\right)k_1^{\,2}+k_4\,P_{Total}\,c_4\left(App\,c_3+1\right)c_1\,k_1\right)k_2\right.\right.$$
$$\left.\left.-P_{Total}\,App\,c_1\,c_3\,c_4\,k_1^{\,2}\,k_3\right)k_2\,k_3\,c_2^{\,2}\,App\,\epsilon^2+3\,App^2\,c_1\,c_2\,c_3^{\,2}\,c_4^{\,2}\,k_1^{\,2}\,k_2\,k_3^{\,2}\left(App\,c_3+1\right)\epsilon\right.$$
$$\left.+App^3\,c_1\,c_3^{\,3}\,c_4^{\,3}\,k_1^{\,2}\,k_3^{\,3}\right)\Big)$$

$$-k_1\left(c_2^{\,2}\left(\left(App^2\,c_3^{\,2}\,k_1+App\left(k_4\,P_{Total}\,c_4+2\,k_1\right)c_3+P_{Total}\left(k_1+k_4\right)c_4+k_1\right)k_2\right.\right. \tag{3.2}$$
$$\left.\left.-App\,P_{Total}\,c_3\,c_4\,k_1\,k_3\right)k_2\,\epsilon^2+2\,App\,c_2\,c_3\,c_4\,k_1\,k_2\,k_3\left(App\,c_3+1\right)\epsilon\right.$$
$$\left.+App^2\,c_3^{\,2}\,c_4^{\,2}\,k_1\,k_3^{\,2}\right)\left(c_2\,k_2\left(App\,c_3+1\right)\epsilon+c_4\,App\,k_3\,c_3\right)c_1$$

$$collect\Big(c_2^{\,2}\left(\left(App^2\,c_3^{\,2}\,k_1+App\left(k_4\,P_{Total}\,c_4+2\,k_1\right)c_3+P_{Total}\left(k_1+k_4\right)c_4+k_1\right)k_2\right.$$
$$\left.-App\,P_{Total}\,c_3\,c_4\,k_1\,k_3\right)k_2\,\epsilon^2+2\,App\,c_2\,c_3\,c_4\,k_1\,k_2\,k_3\left(App\,c_3+1\right)\epsilon+App^2\,c_3^{\,2}\,c_4^{\,2}\,k_1\,k_3^{\,2},$$
$$P_{Total}\Big)$$

$$c_2^{\,2}\left(\left(App\,c_4\,k_4\,c_3+\left(k_1+k_4\right)c_4\right)k_2-App\,c_3\,c_4\,k_1\,k_3\right)k_2\,\epsilon^2\,P_{Total}+c_2^{\,2}\left(App^2\,c_3^{\,2}\,k_1\right. \tag{3.3}$$
$$\left.+2\,App\,k_1\,c_3+k_1\right)k_2^{\,2}\,\epsilon^2+2\,App\,c_2\,c_3\,c_4\,k_1\,k_2\,k_3\left(App\,c_3+1\right)\epsilon+App^2\,c_3^{\,2}\,c_4^{\,2}\,k_1\,k_3^{\,2}$$

$$collect\Big(\left(App\,c_4\,k_4\,c_3+\left(k_1+k_4\right)c_4\right)k_2-App\,c_3\,c_4\,k_1\,k_3,App\Big)$$
$$\left(-c_3\,c_4\,k_1\,k_3+c_4\,k_4\,c_3\,k_2\right)App+\left(k_1+k_4\right)c_4\,k_2 \tag{3.4}$$

$$simplify\left(-c_3\,c_4\,k_1\,k_3+c_4\,k_4\,c_3\,k_2\right)$$
$$-c_3\,c_4\left(k_1\,k_3-k_2\,k_4\right) \tag{3.5}$$

$$Condition := k_1 k_3 - k_2 k_4$$

# Double site phosphorylation system (DSP) [separate kinase and separate phosphatase]

## Features and requirements of obtaining biphasic response in the maximally modified substrate

In this file we analytically prove the absence of enzyme and substrate biphasic in the maximally modified substrate form (App), with changes in $K1_{Total}$, $K2_{Total}$ and $A_{Total}$ respectively.

We note that the key signature of biphasic behavior in the dose response curve of the system is the presence of a steady state of the system that satisfies the following condition.

$\frac{\partial App}{\partial K_{Total}} = 0$ (for enzyme biphasic, $K1_{Total}$ or $K2_{Total}$ depending on the enzyme dose considered)

$\frac{\partial App}{\partial A_{Total}} = 0$ (for substrate biphasic)

**Model DSP with separate kinase and separate phosphatse:** We first describe the model of double site phosphorylation with different kinase and separate phosphatase acting on each modification site.

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart: with(LinearAlgebra): with(VectorCalculus): with(Student[LinearAlgebra]):$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_{b1}\cdot A\cdot K1 + k_{ub1}\cdot AK1 + k_4\cdot ApP1:$$
$$dApp := k_2\cdot ApK2 - k_{b3}\cdot App\cdot P2 + k_{ub3}\cdot AppP2:$$
$$dAp := k_1\cdot AK1 - k_{b2}\cdot Ap\cdot K2 - k_{b4}\cdot Ap\cdot P1 + k_{ub2}\cdot ApK2 + k_{ub4}\cdot ApP1 + k_3\cdot AppP2:$$

$$dAK1 := k_{b1}\cdot A\cdot K1 - \left(k_{ub1} + k_1\right)\cdot AK1:$$
$$dApK2 := k_{b2}\cdot Ap\cdot K2 - \left(k_{ub2} + k_2\right)\cdot ApK2:$$
$$dAppP2 := k_{b3}\cdot App\cdot P2 - \left(k_{ub3} + k_3\right)\cdot AppP2:$$
$$dApP1 := k_{b4}\cdot Ap\cdot P1 - \left(k_{ub4} + k_4\right)\cdot ApP1:$$

$$dP1 := -k_{b4}\cdot Ap\cdot P1 + \left(k_{ub4} + k_4\right)\cdot ApP1:$$
$$dP2 := -k_{b3}\cdot App\cdot P2 + \left(k_{ub3} + k_3\right)\cdot AppP2:$$
$$dK1 := -k_{b1}\cdot A\cdot K1 + \left(k_{ub1} + k_1\right)\cdot AK1:$$
$$dK2 := -k_{b2}\cdot Ap\cdot K2 + \left(k_{ub2} + k_2\right)\cdot ApK2:$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, P1Con, P2Con, K1Con and K2Con for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$P1Con := P1 + ApP1 - P1_{Total}:$$
$$P2Con := P2 + AppP2 - P2_{Total}:$$
$$K1Con := K1 + AK1 - K1_{Total}:$$
$$K2Con := K2 + ApK2 - K2_{Total}:$$
$$ACon := A_{Total} - Ap - A - App - AK1 - ApK2 - AppP2 - ApP1:$$

We now solve the system described at steady state to obtain expression linking the steady state concentrations of the various species. Here we use the Maple command *solve* to solve the equations for a given variable as shown below. We pursue this to finally obtain the steady state concentrations of most species in terms of App.

$$AK1 := solve(dAK1, AK1) = \frac{k_{b1}\,A\,K1}{k_{ub1} + k_1}$$
$$assign(solve(\{dApK2, dAppP2, dApP1\}, \{ApK2, AppP2, ApP1\})):$$
$$assign(solve(\{dA, dAp\}, \{A, Ap\})):$$

Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) : k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

Once this is done, we again solve for the steady state of the phosphatases using the conservation expression for the enzymes (P1Con and P2Con).

$$P2 := solve(P2Con, P2) = \frac{P2_{Total}}{App\, c_3 + 1}$$

$$P1 := solve(P1Con, P1) = \frac{P1_{Total}\left(App\, c_3 + 1\right) K2\, k_2\, c_2}{App\, K2\, c_2\, c_3\, k_2 + App\, c_3\, c_4\, k_3\, P2_{Total} + K2\, c_2\, k_2}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration App and $\epsilon$

$$A := simplify(A) = \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{c_1\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) K1\, k_1}$$

$$Ap = \frac{c_3\, App\, P2_{Total}\, k_3}{\left(App\, c_3 + 1\right) K2\, k_2\, c_2}$$

$$AK1 = \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) k_1}$$

$$ApK2 = \frac{c_3\, App\, P2_{Total}\, k_3}{\left(App\, c_3 + 1\right) k_2}$$

$$AppP2 = \frac{c_3\, App\, P2_{Total}}{App\, c_3 + 1}$$

$$ApP1 = \frac{c_4\, c_3\, App\, P2_{Total}\, k_3\, P1_{Total}}{App\, K2\, c_2\, c_3\, k_2 + c_4\, c_3\, App\, P2_{Total}\, k_3 + K2\, k_2\, c_2}$$

Note that when App, K1 and K2 steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of App, K1 and K2. We now have two expressions, ACon, K1Con and K2Con - the conservation of the substrate and kinase, whose solution for the variables define the steady state of the system.

$$ACon = A_{Total} - \frac{c_3\, App\, P2_{Total}\, k_3}{\left(App\, c_3 + 1\right) K2\, k_2\, c_2} - \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{c_1\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) K1\, k_1} - App - \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) k_1} - \frac{c_3\, App\, P2_{Total}\, k_3}{\left(App\, c_3 + 1\right) k_2} - \frac{c_3\, App\, P2_{Total}}{App\, c_3 + 1} - \frac{c_4\, c_3\, App\, P2_{Total}\, k_3\, P1_{Total}}{App\, K2\, c_2\, c_3\, k_2 + c_4\, c_3\, App\, P2_{Total}\, k_3 + K2\, k_2\, c_2}$$

$$K1Con = K1 + \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) k_1} - K1_{Total}$$

$$K2Con = K2 + \frac{c_3\, App\, P2_{Total}\, k_3}{\left(App\, c_3 + 1\right) k_2} - K2_{Total}$$

## Substrate biphasic

Now in order to show the absence of substrate biphasic response in the maximally modified substrate form (App) with changing $A_{Total}$, we use a proof by contradiction.

We begin with the assumption that substrate biphasic exists, meaning there exists a steady state of the system where $\frac{\partial App}{\partial A_{Total}}$ equals zero. Now if we were to differentiate ACon, K1Con and K2Con with $A_{Total}$, we obtain the following (note that K2Con is a function of App and K2 only).

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial ACon}{\partial K2} \cdot \frac{\partial K2}{\partial A_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}}$$

$$\frac{dK1Con}{dA_{Total}} = 0 = \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}} + \frac{\partial K1Con}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial K1Con}{\partial K2} \cdot \frac{\partial K2}{\partial A_{Total}}$$

$$\frac{dK2Con}{dA_{Total}} = 0 = \frac{\partial K2Con}{\partial App} \cdot \frac{\partial App}{\partial A_{Total}} + \frac{\partial K2Con}{\partial K2} \cdot \frac{\partial K2}{\partial A_{Total}}$$

Now, since we assume that there exists substrate biphasic, we can simplify these expressions further as shown below

$$1 = \frac{\partial ACon}{\partial K2} \cdot \frac{\partial K2}{\partial A_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial A_{Total}}$$

$$0 = \frac{\partial K1}{\partial A_{Total}} + \frac{\partial K1Con}{\partial K2} \cdot \frac{\partial K2}{\partial A_{Total}}$$

$$0 = \frac{\partial K2}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}$, $\frac{\partial K2Con}{\partial App}$ and $\frac{\partial K1Con}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

*simplify*(*diff*(*ACon*, *App*))

$$\left(-\left(App\, c_3 + 1\right)^2 k_2{}^2\, k_1\, c_1\, K1 \left(App^2\, k_2\, c_3{}^2 + \left(\left(P2_{Total} + 2\, App\right) k_2 + k_3\, P2_{Total}\right) c_3\right. \right. \quad \textbf{(1.1.1)}$$

$$\left. \left. + k_2\right) K2^3\, c_2{}^3 - 2\left(App\, c_3 + 1\right) k_2\, k_3 \left(K1\, App^3\, c_1\, c_3{}^2\, c_4\, k_1\, k_2\right.\right.$$

$$+\left(\left(\frac{K1\left(4\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+2\,c_4\,k_1\,P2_{Total}+k_4\,P1_{Total}\,c_4+k_1\right)c_1}{2}\right.\right.$$

$$\left.+\frac{k_4\,P1_{Total}\,c_4}{2}\right)k_2+K1\,k_3\,P2_{Total}\,c_1\,c_4\,k_1\Bigg)\,App\,c_3$$

$$\left.+\frac{k_2\left(K1\left(2\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+k_4\,P1_{Total}\,c_4+k_1\right)c_1+k_4\,P1_{Total}\,c_4\right)}{2}\right)$$

$$c_3\,P2_{Total}\,K2^2\,{c_2}^2-c_4\left(App^3\,c_4\,k_2\,{c_3}^2+\left(\left(2\,App\,c_4+c_4\,P2_{Total}+2\right)k_2\right.\right.$$

$$\left.\left.+k_3\,P2_{Total}\,c_4\right)App\,c_3+k_2\left(App\,c_4+2\right)\right){k_3}^2\,App\,k_1\,{c_3}^2\,c_1\,{P2_{Total}}^2\,K1\,K2\,c_2$$

$$\left.-App^2\,K1\,c_1\,{c_3}^3\,{c_4}^2\,k_1\,{k_3}^3\,{P2_{Total}}^3\right)\Big/\left(\left(App\,c_3+1\right)^2k_2\left(\left(App\,c_3\right.\right.\right.$$

$$\left.\left.\left.+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2c_2\,k_1\,c_1\,K1\,K2\right)$$

$$simplify\left(diff\left(K2Con,App\right)\right)=\frac{c_3\,P2_{Total}\,k_3}{\left(App\,c_3+1\right)^2k_2}$$

$$simplify\left(diff\left(K1Con,App\right)\right)=\frac{P1_{Total}\,k_4\,c_4\,c_3\,P2_{Total}\,k_3\,K2\,c_2\,k_2}{\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2k_1}$$

Thus $\frac{\partial K2}{\partial A_{Total}}$ has to be zero in order to satisfy the third expression above (obtained from differentiating K2Con).

This insight then informs us that in order to satisfy the second expression (obtained from differentiating K1Con), $\frac{\partial K1}{\partial A_{Total}}$ must be equal to zero. (Again this simplification is possible since the denominator of $\frac{\partial K1Con}{\partial K2}$ is non-zero (see below).

$$simplify\left(diff\left(K1Con,K2\right)\right)=-\frac{P1_{Total}\,k_4\,c_4\,c_3\,App\,P2_{Total}\,k_3\left(App\,c_3+1\right)k_2\,c_2}{\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2k_1}$$

Put together, the fact that $\frac{\partial K1}{\partial A_{Total}}$ and $\frac{\partial K2}{\partial A_{Total}}$ are both zero provides a contradiction with the requirements of the first expression (obtained upon differentiation of the total substrate concentration equation).

## Note

This assertion is possible since the denominators of $\frac{\partial ACon}{\partial K1}$ and $\frac{\partial ACon}{\partial K2}$ are both non-zero (see below).

$$simplify(diff(ACon, K1)) = \frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right) k_1\, c_1\, K1^2}$$

$$simplify(diff(ACon, K2)) = \left(\left(\left(\left(P1_{Total}\left(k_4 + k_1\right) c_4 + k_1\right) K1\, c_1 + k_4\, P1_{Total}\, c_4\right)\left(App\, c_3 + 1\right)^2 k_2{}^2\, K2^2\, c_2{}^2 + 2\, K1\, k_3\, P2_{Total}\, K2\, App\, c_1\, c_3\, c_4\, k_1\, k_2\left(App\, c_3 + 1\right) c_2 + App^2\, K1\, c_1\, c_3{}^2\, c_4{}^2\, k_1\, k_3{}^2\, P2_{Total}{}^2\right) k_3\, App\, c_3\, P2_{Total}\right) \Big/ \left(\left(App\, c_3 + 1\right) k_2 \left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right)^2 c_2\, k_1\, c_1\, K1\, K2^2\right)$$

**Thus we have a contradiction, indicating that the system is incapable of exhibiting substrate biphasic response in steady state concentration of App with total substrate concentration.**

## Enzyme biphasic

### *Enzyme biphasic in App with changing $K1_{Total}$*

Now in order to show the absence of enzyme biphasic response in the maximally modified substrate form (App) with changing $K1_{Total}$, we use a proof by contradiction.

We begin with the assumption that enzyme biphasic exists, meaning there exists a steady state of the system where $\frac{\partial App}{\partial K1_{Total}}$ equals zero. Now if we were to differentiate ACon, K1Con and K2Con with $K1_{Total}$, we obtain the following (note that K2Con is a function of App and K2 only).

$$\frac{dACon}{dK1_{Total}} = 0 = \frac{\partial ACon}{\partial App}\cdot\frac{\partial App}{\partial K1_{Total}} + \frac{\partial ACon}{\partial K2}\cdot\frac{\partial K2}{\partial K1_{Total}} + \frac{\partial ACon}{\partial K1}\cdot\frac{\partial K1}{\partial K1_{Total}}$$

$$\frac{dK1Con}{dK1_{Total}} = 0 = \frac{\partial K1\ Con}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial K1}\cdot\frac{\partial K1}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial App}\cdot\frac{\partial App}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial K2}\cdot\frac{\partial K2}{\partial K1_{Total}}$$

$$\frac{dK2Con}{dK1_{Total}} = 0 = \frac{\partial K2Con}{\partial App}\cdot\frac{\partial App}{\partial K1_{Total}} + \frac{\partial K2Con}{\partial K2}\cdot\frac{\partial K2}{\partial K1_{Total}}$$

Now, since we assume that there exists enzyme biphasic, we can simplify these expressions further as shown below

$$0 = \frac{\partial ACon}{\partial K2}\cdot\frac{\partial K2}{\partial K1_{Total}} + \frac{\partial ACon}{\partial K1}\cdot\frac{\partial K1}{\partial K1_{Total}}$$

$$1 = \frac{\partial K1}{\partial K1_{Total}} + \frac{\partial K1Con}{\partial K2}\cdot\frac{\partial K2}{\partial K1_{Total}}$$

$$0 = \frac{\partial K2}{\partial K1_{Total}}$$

#### *Note*

This simplification is possible since the functions

$\frac{\partial ACon}{\partial App}$, $\frac{\partial K2Con}{\partial App}$ and $\frac{\partial K1Con}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$simplify(diff(ACon, App))$

$$\left(-\left(App\,c_3+1\right)^2 k_2{}^2 k_1\,c_1\,K1\left(App^2\,k_2\,c_3{}^2+\left(\left(P2_{Total}+2\,App\right)k_2+k_3\,P2_{Total}\right)c_3+k_2\right)K2^3\,c_2{}^3-2\left(App\,c_3+1\right)k_2\,k_3\left(K1\,App^3\,c_1\,c_3{}^2\,c_4\,k_1\,k_2+\left(\left(\frac{1}{2}\left(K1\left(4\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+2\,c_4\,k_1\,P2_{Total}+k_4\,P1_{Total}\,c_4+k_1\right)c_1\right)+\frac{k_4\,P1_{Total}\,c_4}{2}\right)k_2+K1\,k_3\,P2_{Total}\,c_1\,c_4\,k_1\right)App\,c_3+\frac{1}{2}\left(k_2\left(K1\left(2\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+k_4\,P1_{Total}\,c_4+k_1\right)c_1+k_4\,P1_{Total}\,c_4\right)\right)c_3\,P2_{Total}\,K2^2\,c_2{}^2-c_4\left(App^3\,c_4\,k_2\,c_3{}^2+\left(\left(2\,App\,c_4+c_4\,P2_{Total}+2\right)k_2+k_3\,P2_{Total}\,c_4\right)App\,c_3+k_2\left(App\,c_4+2\right)\right)k_3{}^2\,App\,k_1\,c_3{}^2\,c_1\,P2_{Total}{}^2\,K1\,K2\,c_2-App^2\,K1\,c_1\,c_3{}^3\,c_4{}^2\,k_1\,k_3{}^3\,P2_{Total}{}^3\right)\Big/\left(\left(App\,c_3+1\right)^2 k_2\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2 c_2\,k_1\,c_1\,K1\,K2\right) \tag{2.1.1.1}$$

$$simplify(diff(K2Con, App)) = \frac{c_3\,P2_{Total}\,k_3}{\left(App\,c_3+1\right)^2 k_2}$$

$$simplify(diff(K1Con, App)) = \frac{P1_{Total}\,k_4\,c_4\,c_3\,P2_{Total}\,k_3\,K2\,c_2\,k_2}{\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2 k_1}$$

We can observe from the third expression above (obtained from differentiating K2Con), that $\frac{\partial K2}{\partial K1_{Total}}$ is equal to zero. Thus, resubstituting this in the other expressions we can further infer that, 1 $= \frac{\partial K1}{\partial K1_{Total}}$(from the second expression, obtained from differentiating K1Con).

Resubsituing this in the first expression, we get

$$0 = \frac{\partial ACon}{\partial K1}$$

### *Note*

This simplification was possible since the denominator of $\frac{\partial ACon}{\partial K2}$ is non-zero, allowing the product with $\frac{\partial K2}{\partial K1_{Total}}$ to be zero (see below)

$simplify(diff(ACon, K2)) =$

$$\left(App\left(\left(\left(P1_{Total}\left(k_4+k_1\right)c_4+k_1\right)K1\,c_1+c_4\,k_4\,P1_{Total}\right)\left(App\,c_3+1\right)^2 k_2{}^2\,K2^2\,c_2{}^2+2\,K1\,k_3\,P2_{Total}\,K2\,App\,c_1\,c_3\,c_4\,k_1\,k_2\left(App\,c_3+1\right)c_2+App^2\,K1\,c_1\,c_3{}^2\,c_4{}^2\,k_1\,k_3{}^2\,P2_{Total}{}^2\right)k_3\,c_3\,P2_{Total}\right)\Big/\left(c_2\,c_1\left(App\,c_3+1\right)\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2 K1\,k_1\,k_2\,K2^2\right) \tag{2.1.2.1}$$

However we can quickly observe from the expression for $\frac{\partial ACon}{\partial K1}$ (see below), that it can never be equal to zero for any feasible concentration of steady state of the system (when all the variables are positive).

$simplify(diff(ACon, K1))$

$$\frac{k_3\,P2_{Total}\,App\,c_3\,k_4\,P1_{Total}\,c_4}{k_1\,c_1\,K1^2\left(c_3\left(c_2\,k_2\,K2_{Total}-P2_{Total}\,k_3\left(c_2-c_4\right)\right)App+c_2\,k_2\,K2_{Total}\right)} \tag{2.1.1}$$

**Thus we have a contradiction, indicating that the system is incapable of exhibiting enzyme biphasic response in steady state concentration of App with total enzyme concentration ($K1_{Total}$).**

### *Enzyme biphasic in App with changing $K2_{Total}$*

Now in order to show the absence of substrate biphasic response in the maximally modified substrate form (App) with changing $A_{Total}$, we use a proof by contradiction.
Thus we begin with the assumption that substrate biphasic exists, meaning there exists a steady state of the system where $\frac{\partial App}{\partial A_{Total}}$ equals zero. Now if we were to differentiate ACon, K1Con and K2Con with $A_{Total}$, we obtain the following (note that K2Con is a function of App and K2 only).

$$\frac{dACon}{dK2_{Total}}=0=\frac{\partial ACon}{\partial App}\cdot\frac{\partial App}{\partial K2_{Total}}+\frac{\partial ACon}{\partial K2}\cdot\frac{\partial K2}{\partial K2_{Total}}+\frac{\partial ACon}{\partial K1}\cdot\frac{\partial K1}{\partial K2_{Total}}$$
$$\frac{dK1Con}{dK2_{Total}}=0=\frac{\partial K1Con}{\partial K1}\cdot\frac{\partial K1}{\partial K2_{Total}}+\frac{\partial K1Con}{\partial App}\cdot\frac{\partial App}{\partial K2_{Total}}+\frac{\partial K1Con}{\partial K2}\cdot\frac{\partial K2}{\partial K2_{Total}}$$
$$\frac{dK2Con}{dK2_{Total}}=0=\frac{\partial K2Con}{\partial K2_{Total}}+\frac{\partial K2Con}{\partial App}\cdot\frac{\partial App}{\partial K2_{Total}}+\frac{\partial K2Con}{\partial K2}\cdot\frac{\partial K2}{\partial K2_{Total}}$$

Now, since we assume that there exists substrate biphasic, we can simplify these expressions further as shown below

$$0 = \frac{\partial ACon}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial K2_{Total}}$$

$$0 = \frac{\partial K1}{\partial K2_{Total}} + \frac{\partial K1Con}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}}$$

$$1 = \frac{\partial K2}{\partial K2_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial App}$, $\frac{\partial K2Con}{\partial App}$ and $\frac{\partial K1Con}{\partial App}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial A_{Total}}$ can be zero.

$simplify(diff(ACon, App))$

$$\left(-\left(App\,c_3+1\right)^2 k_2{}^2 k_1\,c_1\,K1\left(App^2 k_2\,c_3{}^2+\left(\left(P2_{Total}+2\,App\right)k_2\right.\right.\right. \tag{2.2.1.1}$$

$$\left.+k_3\,P2_{Total}\right)c_3+k_2\Big)K2^3\,c_2{}^3-2\left(App\,c_3\right.$$

$$\left.+1\right)k_2\,k_3\left(K1\,App^3\,c_1\,c_3{}^2\,c_4\,k_1\,k_2\right.$$

$$+\left(\left(\frac{1}{2}\left(K1\left(4\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+2\,c_4\,k_1\,P2_{Total}+k_4\,P1_{Total}\,c_4\right.\right.\right.\right.$$

$$\left.\left.\left.\left.+k_1\right)c_1\right)+\frac{k_4\,P1_{Total}\,c_4}{2}\right)k_2+K1\,k_3\,P2_{Total}\,c_1\,c_4\,k_1\right)App\,c_3$$

$$+\frac{1}{2}\left(k_2\left(K1\left(2\,App\,c_4\,k_1+c_4\,k_1\,P1_{Total}+k_4\,P1_{Total}\,c_4+k_1\right)c_1\right.\right.$$

$$\left.\left.+k_4\,P1_{Total}\,c_4\right)\right)c_3\,P2_{Total}\,K2^2\,c_2{}^2-c_4\left(App^3\,c_4\,k_2\,c_3{}^2+\left(\left(2\,App\,c_4\right.\right.\right.$$

$$\left.\left.\left.+c_4\,P2_{Total}+2\right)k_2+k_3\,P2_{Total}\,c_4\right)App\,c_3+k_2\left(App\,c_4+2\right)\right)$$

$$\left.k_3{}^2\,App\,k_1\,c_3{}^2\,c_1\,P2_{Total}{}^2\,K1\,K2\,c_2-App^2\,K1\,c_1\,c_3{}^3\,c_4{}^2\,k_1\,k_3{}^3\,P2_{Total}{}^3\right)\Big/$$

$$\left(\left(App\,c_3+1\right)^2 k_2\left(\left(App\,c_3+1\right)K2\,k_2\,c_2\right.\right.$$

$$\left.\left.+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2 c_2\,k_1\,c_1\,K1\,K2\right)$$

$$simplify(diff(K2Con, App)) = \frac{c_3\,P2_{Total}\,k_3}{\left(App\,c_3+1\right)^2 k_2}$$

$$simplify(diff(K1Con, App)) = \frac{P1_{Total}\,k_4\,c_4\,c_3\,P2_{Total}\,k_3\,K2\,c_2\,k_2}{\left(\left(App\,c_3+1\right)K2\,k_2\,c_2+c_4\,c_3\,App\,P2_{Total}\,k_3\right)^2 k_1}$$

This thus allows us to compute $\frac{\partial K1}{\partial K2_{Total}}$ as shown below

$$\frac{\partial K1}{\partial K2_{Total}} = -\frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3 \left(App\, c_3 + 1\right) k_2\, c_2}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right)^2 k_1}$$

Substuting this into the expression obtained from differntiating ACon, we get

$$0 = \frac{\partial ACon}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \left( -\frac{P1_{Total}\, k_4\, c_4\, c_3\, App\, P2_{Total}\, k_3 \left(App\, c_3 + 1\right) k_2\, c_2}{\left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right)^2 k_1} \right)$$

which is evaulated below by Maple and stored as the expression called *Condition*

$$Condition := simplify(diff(ACon, K2) - (diff(K1Con, K2)) \cdot (diff(ACon, K1)))$$

$$\begin{aligned} Condition := {} & \left(c_3\, P2_{Total} \left(K2^3 \left(K1\, c_1 \left(c_4\, P1_{Total} + 1\right) k_1 + k_4\, P1_{Total}\, c_4 \left(K1\, c_1 + 1\right)\right) k_2{}^3\, k_1\, K1 \left(App\, c_3 + 1\right)^3 c_2{}^3 + K2^2\, k_2{}^2\, c_3\, P2_{Total} \left(K1^2\, c_1 \left(c_4\, P1_{Total} + 3\right) k_1{}^2 \right.\right.\right. \\ & + K1\, k_4\, P1_{Total}\, c_4 \left(K1\, c_1 + 1\right) k_1 + k_4{}^2\, P1_{Total}{}^2\, c_4 \left.\right) k_3\, App \left(App\, c_3 + 1\right)^2 c_4\, c_2{}^2 \\ & + 3\, K1^2\, k_3{}^2\, P2_{Total}{}^2\, K2\, App^2\, c_1\, c_3{}^2\, c_4{}^2\, k_1{}^2\, k_2 \left(App\, c_3 + 1\right) c_2 \\ & \left.\left. + App^3\, K1^2\, c_1\, c_3{}^3\, c_4{}^3\, k_1{}^2\, k_3{}^3\, P2_{Total}{}^3 \right) k_3\, App \right) \Big/ \left(K2^2\, k_2\, c_1\, k_1{}^2\, K1^2\, c_2 \left(App\, c_3 + 1\right) \left(\left(App\, c_3 + 1\right) K2\, k_2\, c_2 + c_4\, c_3\, App\, P2_{Total}\, k_3\right)^3\right) \end{aligned} \quad \textbf{(2.2.1)}$$

We can see however the expression Condition can never be equal to zero as it is a sum of positive parameters and variables K1, K2, and App, implying that the expression is purely positive for all feasible steady state concentrations.

**Thus we have a contradiction, indicating that the system is incapable of exhibiting enzyme biphasic response in steady state concentration of App with total enzyme concentration ($K2_{Total}$).**

# Enzymatic networks beyond multisite modification networks

In this folder we detail analytical results pertaining to biphasic dose response by substrates involved in enzymatic networks beyond covalent modification networks/multisite modification of proteins, such as coupled covalent modification network and cascaded enzymatic network.

# Two tier cascaded enzymatic network

## [modified form of one substrate acts as a kinase modifying another substrate]

In this file, we analytically establish the impossibility of different biphasic responses in the two tier cascaded enzymatic network. In such a network, there are two covalent modification cycles (undergoing phosphorylation and dephosphorylation); wherein the dephosphorylation is effected by either a common or separate phosphatases, while a kinase effects the phosphorylation of the first tier of covalent modification (A -> Ap) and the modified form of the first tier substrate (Ap) acts as the kinase for the modification of the second tier substrate (B -> Bp); See figure 2 in the main text for schematic.

In this manuscript we establish the following results regarding biphasic response in the modified substrates to various doses (substrate/enzyme amounts).

1. Common phosphatase network
   a. Enzyme biphasic response is not possible with kinase in either of the maximally modified substrate forms (Ap or Bp)
   b. Substrate biphasic response is not possible in the modified substrate of the first tier (Ap with $A_{Total}$)

Note: Please note that substrate-biphasic response in Bp with total substrate concentration $B_{Total}$ is possible and is shown computationally for a parameter point in the figure 2 in the main text.

2. Separate phosphatase network
   a. Enzyme biphasic response is not possible with kinase in either of the maximally modified substrate forms (Ap or Bp)
   b. Substrate biphasic response is not possible in the modified substrate of the first tier or the second tier (Ap with $A_{Total}$ or in Bp with $B_{Total}$)

These results are summarized in the following tabular column.

**Table 1: Substrate and Enzyme biphasic dose responses in the two tier cascaded enzymatic network**

| System | Substrate Biphasic | Enzyme Biphasic |
|---|---|---|
| Common Phosphatase | Not possible with $A_{Total}$<br>Possible with $B_{Total}$<br>(See figure 2) | Not possible |

| | | |
|---|---|---|
| Separate Phosphatases | Not possible with $A_{Total}$ or $B_{Total}$ | Not possible |

## Common phosphatase model

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_2 \cdot ApP - k_{b1} \cdot A \cdot K + k_{ub1} \cdot AK :$$
$$dAp := k_1 \cdot AK + k_{ub2} \cdot ApP - k_{b2} \cdot Ap \cdot P - p_{b1} \cdot B \cdot Ap + \left(p_{ub1} + p_1\right) \cdot BAp :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApP := k_{b2} \cdot Ap \cdot P - \left(k_{ub2} + k_2\right) \cdot ApP :$$

$$dB := -p_{b1} \cdot B \cdot Ap + p_{ub1} \cdot BAp + p_2 \cdot BpP :$$
$$dBp := p_1 \cdot BAp - p_{b2} \cdot Bp \cdot P + p_{ub2} \cdot BpP :$$

$$dBAp := p_{b1} \cdot B \cdot Ap - \left(p_{ub1} + p_1\right) \cdot BAp :$$
$$dBpP := p_{b2} \cdot Bp \cdot P - \left(p_{ub2} + p_2\right) \cdot BpP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dP := -k_{b2} \cdot Ap \cdot P + \left(k_{ub2} + k_2\right) \cdot ApP - p_{b2} \cdot Bp \cdot P + \left(p_{ub2} + p_2\right) \cdot BpP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + ApP + AK + BAp - A_{Total} :$$
$$KCon := K + AK - K_{Total} :$$
$$PCon := P + ApP + BpP - P_{Total} :$$
$$BCon := B + BpP + BAp + Bp - B_{Total} :$$

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of Ap and Bp. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$

We similarly solve for the other variables using the same command.

$$assign(solve(\{dApP, dBAp, dBpP\}, \{ApP, BAp, BpP\}))$$
$$assign(solve(\{dA, dB\}, \{A, B\}))$$

We introduce the following parameters ($c_1$, $c_2$, $d_1$, and $d_2$). This is done for the sake of brevity and easy

tractability of the expressions obtained.

$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$$
$$p_{b1} := d_1 \cdot \left(p_1 + p_{ub1}\right) : p_{b2} := d_2 \cdot \left(p_2 + p_{ub2}\right) :$$

Once this is done, we again solve for the steady state of the phosphatase using the conservation expression for the enzyme (PCon).

$$P := solve(PCon, P) = \frac{P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration Ap and Bp

$$A = \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K\, k_1\, c_1}$$
$$B = \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) Ap\, p_1\, d_1}$$
$$AK = \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1}$$
$$ApP = \frac{c_2\, Ap\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$
$$BAp = \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1}$$
$$BpP = \frac{d_2\, Bp\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$
$$P = \frac{P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$

Note that when Ap and Bp are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of Ap and Bp. We now have three expressions, ACon, BCon and KCon - the conservation of the substrates and kinase, whose solution for the variables define the steady state of the system.

## 1. Absence of enzyme-biphasic dose response in the modified forms (Ap and Bp) with total kinase concentration

In this subsection we show the absence of enzyme biphasic in either of the modified forms of the substrates with total kinase amounts. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

$$ACon$$

$$\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,k_1\,c_1}+Ap+\frac{c_2\,Ap\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-A_{Total} \quad \textbf{(1.1.1)}$$

*BCon*

$$\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)Ap\,p_1\,d_1}+\frac{d_2\,Bp\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}+Bp-B_{Total} \quad \textbf{(1.1.2)}$$

*KCon*

$$K+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}-K_{Total} \quad \textbf{(1.1.3)}$$

If we are to differentiate the expressions by total kinase amounts we would have the following expressions

$$\frac{dACon}{dK_{Total}}=0=\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$\frac{dBCon}{dK_{Total}}=0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}}=0=\frac{\partial KCon}{\partial K_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

***Absence of biphasic response in Ap with total kinase concentration***

We first begin by showing the absence of a biphasic response in the concentration of Ap as $K_{Total}$ changes. We show this with a proof by contradiction.

If we assume that a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Ap}{\partial K_{Total}}=0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0=\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$0=\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

$$1=\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial K}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Ap}, \frac{\partial BCon}{\partial Ap}, \frac{\partial KCon}{\partial Ap}$ have a non-zero

denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial K_{Total}}$ can be zero.

$$simplify(diff(ACon, Ap)) = \frac{1}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 K k_1 c_1 p_1}\left(\left(\left(Ap^2 k_1 c_2^2 + \left(\left(P_{Total} + 2\,Ap\right) k_1 + P_{Total} k_2\right)\left(Bp\,d_2 + 1\right) c_2 + k_1\left(Bp\,d_2 + 1\right)^2\right) K c_1 + P_{Total} c_2 k_2\left(Bp\,d_2 + 1\right)\right) p_1 - d_2\,Bp\,P_{Total} p_2 c_2 K k_1 c_1\right)$$

$$simplify(diff(BCon, Ap)) = -\frac{\left(c_2 d_1\left(p_1 + p_2\right) Ap^2 + 2\,p_2 c_2 Ap + p_2\left(Bp\,d_2 + 1\right)\right) d_2\,Bp\,P_{Total}}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 Ap^2 p_1 d_1}$$

$$simplify(diff(KCon, Ap)) = \frac{P_{Total} c_2 k_2\left(Bp\,d_2 + 1\right)}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 k_1}$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Bp}$ is never zero (as shown below), that $\frac{\partial Bp}{\partial K_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Bp)) = \frac{1}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 Ap\,p_1 d_1}\left(Ap^3 c_2^2 d_1 p_1 + 2\left(\left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2 P_{Total}}{2}\right) d_2 + p_1\right) d_1 c_2 Ap^2 + \left(\left(Bp^2 d_2^2 p_1 + \left(\left(2\,Bp + P_{Total}\right) p_1 + p_2 P_{Total}\right) d_2 + p_1\right) d_1 + c_2 d_2 p_2 P_{Total}\right) Ap + d_2 P_{Total} p_2\right)$$

Now since $\frac{\partial ACon}{\partial Bp}$ and $\frac{\partial KCon}{\partial Bp}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Bp)) = \frac{P_{Total} d_2\left(\left(Ap\left(\left(p_2 - p_1\right) k_1 - k_2 p_1\right) c_2 + p_2 k_1\right) K c_1 - c_2\,Ap\,k_2 p_1\right)}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 K k_1 c_1 p_1}$$

$$simplify(diff(KCon, Bp)) = -\frac{c_2\,Ap\,P_{Total} k_2 d_2}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 k_1}$$

the differentiated expressions further simplify to

$$0 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

$$0 = \frac{\partial Bp}{\partial K_{Total}}$$

$$1 = \frac{\partial K}{\partial K_{Total}}$$

This implies that $\frac{\partial ACon}{\partial K} = 0$ is a necessary condition for the biphasic behavior to exist. However we can also see that $\frac{\partial ACon}{\partial K} = 0$ is not possible for any feasible steady state of the system (see below)

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K^2\, k_1\, c_1}$$

**Thus a biphasic response in Ap cannot be possible with total kinase concentration ($K_{Total}$).**

***Absence of biphasic response in Bp with total kinase concentration***

In a similar manner we can show that a biphasic response in Bp can also not exist with total kinase.

We proceed similarly, with a proof by contradiction, starting with the assumption that there exists a biphasic response in Bp with total kinase.

$$\frac{\partial Bp}{\partial K_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$
$$0 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$
$$1 = \frac{\partial K}{\partial K_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial K_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Bp)) = \frac{P_{Total}\, d_2 \left(\left(Ap \left(\left(p_2 - p_1\right) k_1 - k_2\, p_1\right) c_2 + p_2\, k_1\right) K\, c_1 - c_2\, Ap\, k_2\, p_1\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K\, k_1\, c_1\, p_1}$$

$$simplify(diff(BCon, Bp)) = \frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 Ap\, p_1\, d_1} \left(Ap^3\, c_2^2\, d_1\, p_1 + 2 \left(\left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2\, P_{Total}}{2}\right) d_2 + p_1\right) d_1\, c_2\, Ap^2 + \left(\left(Bp^2\, d_2^2\, p_1 + \left(\left(2\, Bp + P_{Total}\right) p_1 + p_2\, P_{Total}\right) d_2 + p_1\right) d_1 + c_2\, d_2\, p_2\, P_{Total}\right) Ap + d_2\, P_{Total}\, p_2\right)$$

$$simplify(diff(KCon, Bp)) = -\frac{c_2\, Ap\, P_{Total}\, k_2\, d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1}$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Ap}$ is never zero (as shown below), that $\frac{\partial Ap}{\partial K_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Ap)) = -\frac{\left(c_2 d_1 \left(p_1 + p_2\right) Ap^2 + 2\, p_2 c_2\, Ap + p_2 \left(Bp\, d_2 + 1\right)\right) d_2\, Bp\, P_{Total}}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 Ap^2\, p_1\, d_1}$$

Now since $\frac{\partial ACon}{\partial Ap}$ and $\frac{\partial KCon}{\partial Ap}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Ap)) = \frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K k_1 c_1 p_1}\left(\left(\left(Ap^2 k_1 c_2^2 + \left(\left(P_{Total} + 2\, Ap\right) k_1 + P_{Total} k_2\right)\left(Bp\, d_2 + 1\right) c_2 + k_1 \left(Bp\, d_2 + 1\right)^2\right) K c_1 + P_{Total}\, c_2\, k_2 \left(Bp\, d_2 + 1\right)\right) p_1 - d_2\, Bp\, P_{Total}\, p_2\, c_2\, K k_1 c_1\right)$$

$$simplify(diff(KCon, Ap)) = \frac{P_{Total}\, c_2\, k_2 \left(Bp\, d_2 + 1\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1}$$

the differentiated expressions further simplify to

$$0 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

$$0 = \frac{\partial Ap}{\partial K_{Total}}$$

$$1 = \frac{\partial K}{\partial K_{Total}}$$

However we can also see that $\frac{\partial ACon}{\partial K} = 0$ is not possible for any feasible steady state of the system.

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K^2 k_1 c_1}$$

**Thus a biphasic response in Bp cannot be possible with total kinase concentration ($K_{Total}$).**

## 2. Absence of substrate-biphasic dose response in Ap with total substrate concentration ($A_{Total}$)

In this subsection we show the absence of substrate-biphasic in the modified form of the first tier substrate (Ap) with total substrate concentration ($A_{Total}$). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the

system.

$ACon$

$$\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,k_1\,c_1}+Ap+\frac{c_2\,Ap\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-A_{Total} \tag{1.2.1}$$

$BCon$

$$\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)Ap\,p_1\,d_1}+\frac{d_2\,Bp\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}+Bp-B_{Total} \tag{1.2.2}$$

$KCon$

$$K+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}-K_{Total} \tag{1.2.3}$$

If we are to differentiate the expressions by total substrate amount $A_{Total}$ we would have the following expressions

$$\frac{dACon}{dA_{Total}}=0=\frac{\partial ACon}{\partial A_{Total}}+\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

$$\frac{dBCon}{dA_{Total}}=0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}$$

$$\frac{dKCon}{dA_{Total}}=0=\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}$$

We now show the absence of a biphasic response in the concentration of Ap with $A_{Total}$ using a proof by contradiction.

Assuming that a biphasic response exists, there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Ap}{\partial A_{Total}}=0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$1=\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

$$0=\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}$$

$$0=\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Ap}, \frac{\partial BCon}{\partial Ap}, \frac{\partial KCon}{\partial Ap}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial K_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Ap)) = \frac{1}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 K k_1 c_1 p_1}\left(\left(\left(Ap^2 k_1 c_2^2 + \left(Bp\,d_2 + 1\right)\left(\left(P_{Total} + 2\,Ap\right)k_1 + P_{Total} k_2\right)c_2 + k_1\left(Bp\,d_2 + 1\right)^2\right)K c_1 + P_{Total} c_2 k_2\left(Bp\,d_2 + 1\right)\right)p_1 - d_2\,Bp\,P_{Total} p_2 c_2 K k_1 c_1\right)$$

$$simplify(diff(BCon, Ap)) = -\frac{Bp\left(c_2 d_1\left(p_1 + p_2\right)Ap^2 + 2\,p_2 c_2 Ap + p_2\left(Bp\,d_2 + 1\right)\right)d_2 P_{Total}}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 Ap^2 p_1 d_1}$$

$$simplify(diff(KCon, Ap)) = \frac{P_{Total} c_2 k_2\left(Bp\,d_2 + 1\right)}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 k_1}$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Bp}$ is never zero (as shown below), that $\frac{\partial Bp}{\partial A_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Bp)) \quad \textbf{(1.2.4)}$$

$$\frac{1}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 Ap\,p_1 d_1}\left(Ap^3 c_2^2 d_1 p_1 + 2\,d_1\left(\left(\left(Bp + \frac{P_{Total}}{2}\right)p_1 + \frac{p_2 P_{Total}}{2}\right)d_2 + p_1\right)c_2 Ap^2 + \left(\left(Bp^2 d_2^2 p_1 + \left(\left(2\,Bp + P_{Total}\right)p_1 + p_2 P_{Total}\right)d_2 + p_1\right)d_1 + c_2 d_2 p_2 P_{Total}\right)Ap + d_2 P_{Total} p_2\right)$$

Now since $\frac{\partial ACon}{\partial Bp}$ and $\frac{\partial KCon}{\partial Bp}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Bp)) = \frac{P_{Total}\left(\left(Ap\left(\left(p_2 - p_1\right)k_1 - k_2 p_1\right)c_2 + p_2 k_1\right)K c_1 - c_2\,Ap\,k_2 p_1\right)d_2}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 K k_1 c_1 p_1}$$

$$simplify(diff(KCon, Bp)) = -\frac{c_2\,Ap\,P_{Total} k_2 d_2}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^2 k_1}$$

the differentiated expressions further simplify to

$$1 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

$$0 = \frac{\partial Bp}{\partial A_{Total}}$$

$$0 = \frac{\partial K}{\partial A_{Total}}$$

However we now notice a contradiction, $0 = \frac{\partial K}{\partial A_{Total}}$ has to be true, while $1 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$. This is not possible since $\frac{\partial ACon}{\partial K}$ has a non-zero denominator and thus the product of $\frac{\partial ACon}{\partial K}$ and $\frac{\partial K}{\partial A_{Total}}$ cannot be 1, while the latter is necessarily zero.

**Thus a biphasic response in Ap is not possible with total substrate concentration $A_{Total}$.**

## 3. Presence of substrate-biphasic dose response in Bp with total substrate concentration ($B_{Total}$)

In this subsection we show the presence of substrate-biphasic in the modified form of the second tier substrate (Bp) with total substrate concentration ($B_{Total}$). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

*ACon*

$$\frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K\, k_1\, c_1} + Ap + \frac{c_2\, Ap\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1} + \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1} + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} - A_{Total} \quad \textbf{(1.3.1)}$$

*BCon*

$$\frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) Ap\, p_1\, d_1} + \frac{d_2\, Bp\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1} + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} + Bp - B_{Total} \quad \textbf{(1.3.2)}$$

*KCon*

$$K + \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1} - K_{Total} \quad \textbf{(1.3.3)}$$

If we are to differentiate the expressions by total substrate amount $B_{Total}$ we would have the following expressions

$$\frac{dACon}{dB_{Total}} = 0 = \frac{\partial ACon}{\partial B_{Total}} + \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$

$$\frac{dBCon}{dB_{Total}} = 0 = \frac{\partial BCon}{\partial B_{Total}} + \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}}$$

$$\frac{dKCon}{dB_{Total}} = 0 = \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}}$$

Now in order to show the presence of the behavior, we assume that the biphasic dose response in Bp

with increasing substrate dose exists.

Thus, assuming that a biphasic response exists, there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial B_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$
$$1 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}$$
$$0 = \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + 1 \cdot \frac{\partial K}{\partial B_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial K_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Bp)) = \frac{P_{Total}\left(K\left(\left(\left(p_2 - p_1\right)k_1 - k_2 p_1\right)Ap\, c_2 + p_2 k_1\right)c_1 - c_2 Ap\, k_2 p_1\right)d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K k_1 c_1 p_1}$$

$$simplify(diff(BCon, Bp)) = \frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 Ap\, p_1 d_1}\left(Ap^3 c_2{}^2 d_1 p_1 + 2 c_2 d_1\left(\left(\left(Bp + \frac{P_{Total}}{2}\right)p_1 + \frac{p_2 P_{Total}}{2}\right)d_2 + p_1\right)Ap^2 + \left(\left(Bp^2 d_2{}^2 p_1 + \left(\left(2\, Bp + P_{Total}\right)p_1 + p_2 P_{Total}\right)d_2 + p_1\right)d_1 + c_2 d_2 p_2 P_{Total}\right)Ap + d_2 P_{Total} p_2\right)$$

$$simplify(diff(KCon, Bp)) = -\frac{c_2 Ap\, P_{Total} k_2 d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1}$$

$$0 = \frac{\partial Ap}{\partial B_{Total}} \cdot \left(\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K} \cdot \frac{\partial KCon}{\partial Ap}\right)$$
$$1 = \frac{\partial BCon}{\partial Ap} \cdot \frac{BAp}{\partial B_{Total}}$$

$$T := simplify(numer(diff(ACon, Ap) - diff(ACon, K) \cdot diff(KCon, Ap)))$$

$$T := -c_1 K^2 \left( \left( -Ap^2 c_2^{\,2} - \left(Bp\, d_2 + 1\right) \left(P_{Total} + 2\, Ap\right) c_2 - \left(Bp\, d_2 + 1\right)^2 \right) p_1 + d_2\, Bp\, P_{Total}\, p_2\, c_2 \right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1^{\,2} + K\, P_{Total}\, c_2\, k_2\, p_1 \left(Bp\, d_2 + 1\right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) \left(K\, c_1 + 1\right) k_1 + c_2^{\,2}\, Ap\, P_{Total}^{\,2}\, k_2^{\,2} \left(Bp\, d_2 + 1\right) p_1 \tag{1.3.4}$$

$$collect(T, K)$$

$$\left( -c_1 \left( \left( -Ap^2 c_2^{\,2} - \left(Bp\, d_2 + 1\right) \left(P_{Total} + 2\, Ap\right) c_2 - \left(Bp\, d_2 + 1\right)^2 \right) p_1 + d_2\, Bp\, P_{Total}\, p_2\, c_2 \right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1^{\,2} + P_{Total}\, c_2\, k_2\, p_1 \left(Bp\, d_2 + 1\right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) c_1\, k_1 \right) K^2 + P_{Total}\, c_2\, k_2\, p_1 \left(Bp\, d_2 + 1\right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1\, K + c_2^{\,2}\, Ap\, P_{Total}^{\,2}\, k_2^{\,2} \left(Bp\, d_2 + 1\right) p_1 \tag{1.3.5}$$

$$simplify\left( -c_1 \left( \left( -Ap^2 c_2^{\,2} - \left(Bp\, d_2 + 1\right) \left(P_{Total} + 2\, Ap\right) c_2 - \left(Bp\, d_2 + 1\right)^2 \right) p_1 + d_2\, Bp\, P_{Total}\, p_2\, c_2 \right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1^{\,2} + P_{Total}\, c_2\, k_2\, p_1 \left(Bp\, d_2 + 1\right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) c_1\, k_1 \right)$$

$$c_1 \left( \left( Ap^2 c_2^{\,2} k_1 + 2 \left( \left( Ap + \frac{P_{Total}}{2} \right) k_1 + \frac{P_{Total}\, k_2}{2} \right) \left(Bp\, d_2 + 1\right) c_2 + k_1 \left(Bp\, d_2 + 1\right)^2 \right) p_1 - P_{Total}\, Bp\, c_2\, d_2\, k_1\, p_2 \right) \left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1 \tag{1.3.6}$$

$$collect\left( \left( Ap^2 c_2^{\,2} k_1 + 2 \left( \left( Ap + \frac{P_{Total}}{2} \right) k_1 + \frac{P_{Total}\, k_2}{2} \right) \left(Bp\, d_2 + 1\right) c_2 + k_1 \left(Bp\, d_2 + 1\right)^2 \right) p_1 - P_{Total}\, Bp\, c_2\, d_2\, k_1\, p_2, P_{Total} \right)$$

$$\left( \left(k_1 + k_2\right) \left(Bp\, d_2 + 1\right) c_2\, p_1 - Bp\, c_2\, d_2\, k_1\, p_2 \right) P_{Total} + \left( Ap^2 c_2^{\,2} k_1 + 2\, k_1\, Ap \left(Bp\, d_2 + 1\right) c_2 + k_1 \left(Bp\, d_2 + 1\right)^2 \right) p_1 \tag{1.3.7}$$

$$simplify\left( collect\left( \left(k_1 + k_2\right) \left(Bp\, d_2 + 1\right) c_2\, p_1 - Bp\, c_2\, d_2\, k_1\, p_2, Bp \right) \right)$$

$$\left( \left(k_1 + k_2\right) \left(Bp\, d_2 + 1\right) p_1 - Bp\, d_2\, k_1\, p_2 \right) c_2 \tag{1.3.8}$$

$$Condition := \left(p_1 - p_2\right) k_1 + k_2\, p_1$$

$$Condition := \left(p_1 - p_2\right) k_1 + k_2\, p_1 \tag{1.3.9}$$

**Thus for a substrate biphasic respose to exist in Bp with increasing BTotal, the above condition would need to be satisfied. This is also a necessary condition to obtain the behavior for some total amounts.**

## Different phosphatase model

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_2 \cdot ApP - k_{b1} \cdot A \cdot K + k_{ub1} \cdot AK :$$
$$dAp := k_1 \cdot AK + k_{ub2} \cdot ApP - k_{b2} \cdot Ap \cdot P1 - p_{b1} \cdot B \cdot Ap + \left(p_{ub1} + p_1\right) \cdot BAp :$$

$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApP := k_{b2} \cdot Ap \cdot P1 - \left(k_{ub2} + k_2\right) \cdot ApP :$$

$$dB := -p_{b1} \cdot B \cdot Ap + p_{ub1} \cdot BAp + p_2 \cdot BpP :$$
$$dBp := p_1 \cdot BAp - p_{b2} \cdot Bp \cdot P2 + p_{ub2} \cdot BpP :$$

$$dBAp := p_{b1} \cdot B \cdot Ap - \left(p_{ub1} + p_1\right) \cdot BAp :$$
$$dBpP := p_{b2} \cdot Bp \cdot P2 - \left(p_{ub2} + p_2\right) \cdot BpP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dP1 := -k_{b2} \cdot Ap \cdot P1 + \left(k_{ub2} + k_2\right) \cdot ApP :$$
$$dP2 := -p_{b2} \cdot Bp \cdot P2 + \left(p_{ub2} + p_2\right) \cdot BpP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + ApP + AK + BAp - A_{Total} :$$
$$KCon := K + AK - K_{Total} :$$
$$P1Con := P1 + ApP - P1_{Total} :$$
$$P2Con := P2 + BpP - P2_{Total} :$$
$$BCon := B + BpP + BAp + Bp - B_{Total} :$$

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of Ap and Bp. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$

We similarly solve for the other variables using the same command.

$$assign(solve(\{dApP, dBAp, dBpP\}, \{ApP, BAp, BpP\}))$$
$$assign(solve(\{dA, dB\}, \{A, B\}))$$

We introduce the following parameters ($c_1$, $c_2$, $d_1$, and $d_2$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) : k_{b2} := c_2 \cdot (k_2 + k_{ub2}) :$
$p_{b1} := d_1 \cdot (p_1 + p_{ub1}) : p_{b2} := d_2 \cdot (p_2 + p_{ub2}) :$

Once this is done, we again solve for the steady state of the phosphatase using the conservation expression for the enzyme (PCon).

$$P1 := solve(P1Con, P1) = \frac{P1_{Total}}{Ap\, c_2 + 1}$$

$$P2 := solve(P2Con, P2) = \frac{P2_{Total}}{Bp\, d_2 + 1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration Ap and Bp

$$A = \frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, K\, k_1\, c_1}$$

$$B = \frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, Ap\, p_1\, d_1}$$

$$AK = \frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, k_1}$$

$$ApP = \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1}$$

$$BAp = \frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, p_1}$$

$$BpP = \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2 + 1}$$

Note that when Ap and Bp are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of Ap and Bp. We now have three expressions, ACon, BCon and KCon - the conservation of the substrates and kinase, whose solution for the variables define the steady state of the system.

## 1. Absence of enzyme-biphasic dose response in the modified forms (Ap and Bp) with total kinase concentration

In this subsection we show the absence of enzyme biphasic in either of the modified forms of the substrates with total kinase amounts. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

$ACon$

$$\frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, K\, k_1\, c_1} + Ap + \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1} + \frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, k_1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, p_1} \qquad \textbf{(2.1.1)}$$

$$-A_{Total}$$

*BCon*

$$\frac{d_2\,Bp\,P2_{Total}\,p_2}{\left(Bp\,d_2+1\right)Ap\,p_1\,d_1}+\frac{d_2\,Bp\,P2_{Total}}{Bp\,d_2+1}+\frac{d_2\,Bp\,P2_{Total}\,p_2}{\left(Bp\,d_2+1\right)p_1}+Bp-B_{Total} \tag{2.1.2}$$

*KCon*

$$K+\frac{c_2\,Ap\,P1_{Total}\,k_2}{\left(Ap\,c_2+1\right)k_1}-K_{Total} \tag{2.1.3}$$

If we are to differentiate the expressions by total kinase amounts we would have the following expressions

$$\frac{dACon}{dK_{Total}}=0=\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$\frac{dBCon}{dK_{Total}}=0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}}=0=\frac{\partial KCon}{\partial K_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

***Absence of biphasic response in Ap with total kinase concentration***

We first begin by showing the absence of a biphasic response in the concentration of Ap as $K_{Total}$ changes. We show this with a proof by contradiction.

If we assume that a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Ap}{\partial K_{Total}}=0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0=\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$0=\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

$$1=\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial K}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Ap}, \frac{\partial BCon}{\partial Ap}, \frac{\partial KCon}{\partial Ap}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial K_{Total}}$ can be zero.

$simplify(diff(ACon, Ap)) =$

$$\frac{\left(Ap^2\, {c_2}^2\, k_1 + \left(\left(2\,Ap + PI_{Total}\right) k_1 + k_2\, PI_{Total}\right) c_2 + k_1\right) K\, c_1 + c_2\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right)^2 K\, k_1\, c_1}$$

$$simplify(diff(BCon, Ap)) = -\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) Ap^2\, p_1\, d_1}$$

$$simplify(diff(KCon, Ap)) = \frac{c_2\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right)^2 k_1}$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Bp}$ is never zero (as shown below), that $\frac{\partial Bp}{\partial K_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Bp)) = \frac{\left(Bp^2\, {d_2}^2\, p_1 + \left(\left(2\,Bp + P2_{Total}\right) p_1 + p_2\, P2_{Total}\right) d_2 + p_1\right) Ap\, d_1 + d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 Ap\, p_1\, d_1}$$

Now since $\frac{\partial ACon}{\partial Bp}$ and $\frac{\partial KCon}{\partial Bp}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Bp)) = \frac{d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 p_1}$$

$$simplify(diff(KCon, Bp)) = 0$$

the differentiated expressions further simplify to

$$0 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

$$0 = \frac{\partial Bp}{\partial K_{Total}}$$

$$1 = \frac{\partial K}{\partial K_{Total}}$$

This implies that $\frac{\partial ACon}{\partial K} = 0$ is a necessary condition for the biphasic behavior to exist. However we can also see that $\frac{\partial ACon}{\partial K} = 0$ is not possible for any feasible steady state of the system (see below)

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) K^2\, k_1\, c_1}$$

**Thus a biphasic response in Ap cannot be possible with total kinase concentration ($K_{Total}$).**

***Absence of biphasic response in Bp with total kinase concentration***

In a similar manner we can show that a biphasic response in Bp can also not exist with total kinase.

We proceed similarly, with a proof by contradiction, starting with the assumption that there exists a biphasic response in Bp with total kinase.

$$\frac{\partial Bp}{\partial K_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$
$$0 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$
$$1 = \frac{\partial K}{\partial K_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial K_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Bp)) = \frac{d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 p_1}$$

$$simplify(diff(BCon, Bp)) = \frac{\left(Bp^2\, {d_2}^2\, p_1 + \left(\left(2\, Bp + P2_{Total}\right) p_1 + p_2\, P2_{Total}\right) d_2 + p_1\right) Ap\, d_1 + d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 Ap\, p_1\, d_1}$$

$$simplify(diff(KCon, Bp)) = 0$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Ap}$ is never zero (as shown below), that $\frac{\partial Ap}{\partial K_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Ap)) = -\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) Ap^2\, p_1\, d_1}$$

Now since $\frac{\partial ACon}{\partial Ap}$ and $\frac{\partial KCon}{\partial Ap}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Ap)) = \frac{\left(Ap^2\, {c_2}^2\, k_1 + \left(\left(2\, Ap + PI_{Total}\right) k_1 + k_2\, PI_{Total}\right) c_2 + k_1\right) K\, c_1 + c_2\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right)^2 K\, k_1\, c_1}$$

$$simplify(diff(KCon, Ap)) = \frac{c_2\, PI_{Total}\, k_2}{\left(c_2\, Ap + 1\right)^2 k_1}$$

the differentiated expressions further simplify to

$$0 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial K_{Total}}$$

$$0 = \frac{\partial Ap}{\partial K_{Total}}$$

$$1 = \frac{\partial K}{\partial K_{Total}}$$

However we can also see that $\frac{\partial ACon}{\partial K} = 0$ is not possible for any feasible steady state of the system.

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, P1_{Total}\, k_2}{(c_2\, Ap + 1)\, K^2\, k_1\, c_1}$$

**Thus a biphasic response in Bp cannot be possible with total kinase concentration ($K_{Total}$).**

## 2. Absence of substrate biphasic dose response in Ap with total substrate concentration ($A_{Total}$)

In this subsection we show the absence of substrate-biphasic in the modified form of the first tier substrate (Ap) with total substrate concentration ($A_{Total}$). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

*ACon*

$$\frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, K\, k_1\, c_1} + Ap + \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1} + \frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, k_1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, p_1} - A_{Total} \tag{2.2.1}$$

*BCon*

$$\frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, Ap\, p_1\, d_1} + \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2 + 1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{(Bp\, d_2 + 1)\, p_1} + Bp - B_{Total} \tag{2.2.2}$$

*KCon*

$$K + \frac{c_2\, Ap\, P1_{Total}\, k_2}{(Ap\, c_2 + 1)\, k_1} - K_{Total} \tag{2.2.3}$$

If we are to differentiate the expressions by total substrate amount $A_{Total}$ we would have the following expressions

$$\frac{dACon}{dA_{Total}} = 0 = \frac{\partial ACon}{\partial A_{Total}} + \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}} + \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

$$\frac{dBCon}{dA_{Total}} = 0 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}} + \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}}$$

$$\frac{dKCon}{dA_{Total}} = 0 = \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial A_{Total}} + \frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}}$$

We now show the absence of a biphasic response in the concentration of Ap with $A_{Total}$ using a proof by contradiction.

Assuming that a biphasic response exists, there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Ap}{\partial A_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$1 = \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

$$0 = \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}}$$

$$0 = \frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}} + \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Ap}, \frac{\partial BCon}{\partial Ap}, \frac{\partial KCon}{\partial Ap}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial A_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Ap)) = \frac{\left(Ap^2\, c_2{}^2\, k_1 + \left(\left(2\, Ap + PI_{Total}\right) k_1 + k_2\, PI_{Total}\right) c_2 + k_1\right) K\, c_1 + c_2\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right)^2 K\, k_1\, c_1}$$

$$simplify(diff(BCon, Ap)) = -\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) Ap^2\, p_1\, d_1}$$

$$simplify(diff(KCon, Ap)) = \frac{c_2\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right)^2 k_1}$$

Now we can make the following inference that since $\frac{\partial BCon}{\partial Bp}$ is never zero (as shown below), that $\frac{\partial Bp}{\partial A_{Total}}$ has to be necessarily zero (at the biphasic peak).

$$simplify(diff(BCon, Bp)) \quad \textbf{(2.2.4)}$$

$$\frac{\left(Bp^2\, d_2{}^2\, p_1 + \left(\left(2\, Bp + P2_{Total}\right) p_1 + p_2\, P2_{Total}\right) d_2 + p_1\right) Ap\, d_1 + d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 Ap\, p_1\, d_1}$$

Now since $\frac{\partial ACon}{\partial Bp}$ and $\frac{\partial KCon}{\partial Bp}$ have non-zero denominators (as shown below)

$$simplify(diff(ACon, Bp)) = \frac{d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 p_1}$$

$$simplify(diff(KCon, Bp)) = 0$$

the differentiated expressions further simplify to

$$1 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$

$$0 = \frac{\partial Bp}{\partial A_{Total}}$$

$$0 = \frac{\partial K}{\partial A_{Total}}$$

However we now notice a contradiction, $0 = \frac{\partial K}{\partial A_{Total}}$ has to be true, while $1 = \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$. This is not possible since $\frac{\partial ACon}{\partial K}$ has a non-zero denominator and thus the product of $\frac{\partial ACon}{\partial K}$ and $\frac{\partial K}{\partial A_{Total}}$ cannot be 1, while the latter is necessarily zero.

**Thus a biphasic response in Ap is not possible with total substrate concentration $A_{Total}$.**

## 3. Absence of substrate biphasic dose response in Bp with total substrate concentration ($B_{Total}$)

In this subsection we show the absence of substrate-biphasic in the modified form of the second tier substrate (Bp) with total substrate concentration ($B_{Total}$). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

*ACon*

$$\frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) K k_1\, c_1} + Ap + \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1} + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) k_1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) p_1} - A_{Total} \tag{2.3.1}$$

*BCon*

$$\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) Ap\, p_1\, d_1} + \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2 + 1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) p_1} + Bp - B_{Total} \tag{2.3.2}$$

*KCon*

$$K + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) k_1} - K_{Total} \tag{2.3.3}$$

If we are to differentiate the expressions by total substrate amount $B_{Total}$ we would have the following expressions

$$\frac{dACon}{dB_{Total}} = 0 = \frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial B_{Total}} + \frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial B_{Total}}$$
$$\frac{dBCon}{dB_{Total}} = 0 = \frac{\partial BCon}{\partial B_{Total}} + \frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}} + \frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial B_{Total}}$$
$$\frac{dKCon}{dB_{Total}} = 0 = \frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial B_{Total}} + \frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}}$$

We now show the absence of a biphasic response in the concentration of Bp with $B_{Total}$ using a proof by contradiction.

Assuming that a biphasic response exists, there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial B_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial B_{Total}}$$
$$1 = \frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}}$$
$$0 = \frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial B_{Total}} + 1\cdot\frac{\partial K}{\partial B_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial B_{Total}} = 0$ can be zero.

$$simplify(diff(ACon, Bp)) = \frac{\left(Ap^2 c_2^2 k_1 + \left(\left(2\,Ap + P1_{Total}\right) k_1 + k_2 P1_{Total}\right) c_2 + k_1\right) K c_1 + c_2 P1_{Total} k_2}{\left(Ap\,c_2 + 1\right)^2 K k_1 c_1}$$
$$simplify(diff(BCon, Bp)) = -\frac{d_2 Bp\, P2_{Total} p_2}{\left(Bp\,d_2 + 1\right) Ap^2 p_1 d_1}$$
$$simplify(diff(KCon, Bp)) = \frac{c_2 P1_{Total} k_2}{\left(Ap\,c_2 + 1\right)^2 k_1}$$

Now by solving equation 3 above (obtained from differentiating KCon) for $\frac{\partial K}{\partial B_{Total}}$ and substituting it in the first equation (obtained from differentiating ACon) we get,

$$0 = \frac{\partial Ap}{\partial B_{Total}}\cdot\left(\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K}\cdot\frac{\partial KCon}{\partial Ap}\right)$$

$$1 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}$$

Now from the first expression above we can infer that $\left( \frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K} \cdot \frac{\partial KCon}{\partial Ap} \right)$ should be equal to zero (since otherwise, $\frac{\partial Ap}{\partial B_{Total}}$ must be equal to zero and this would violate the second expression).

However, we can see below that the expression $\left( \frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K} \cdot \frac{\partial KCon}{\partial Ap} \right)$ is never zero (see below).

$$simplify\left(diff\left(ACon, Ap\right) - diff\left(ACon, K\right) \cdot diff\left(KCon, Ap\right)\right)$$

$$\frac{1}{\left(c_2\, Ap + 1\right)^3 K^2\, k_1{}^2\, c_1} \left( Ap^3\, K^2\, c_1\, c_2{}^3\, k_1{}^2 + 3\, Ap \left( K^2\, c_1 \left( Ap + \frac{PI_{Total}}{3} \right) k_1{}^2 + \frac{K\, k_2\, PI_{Total} \left( K\, c_1 + 1 \right) k_1}{3} + \frac{k_2{}^2\, PI_{Total}{}^2}{3} \right) c_2{}^2 + 3\, K\, k_1 \left( K\, c_1 \left( Ap + \frac{PI_{Total}}{3} \right) k_1 + \frac{k_2\, PI_{Total} \left( K\, c_1 + 1 \right)}{3} \right) c_2 + K^2\, k_1{}^2\, c_1 \right) \tag{2.3.4}$$

**Thus contradiction. The conditions can't be satisfied implying that a biphasic response in Ap is not possible with total substrate concentration $B_{Total}$.**

# Coupled covalent modification systems

**[coupled through enzyme sharing]**

In this file, we study the propensity for substrate and enzyme biphasic dose responses in modified substrate form in coupled covalent modification systems. We consider (as mentioned in the main text), a suite of models of two coupled covalent modification systems (A=Ap and B=Bp) where the coupling is achieved through enzyme sharing (one or more enzymes). By systematically studying models with different degrees of coupling (common kinase and phosphatase, common kinase only, common phosphatase only) we show the following key results

1. Common enzymes (kinase and phosphatase) model:
   a. Enzyme biphasic response is impossible in the modified substrate forms of either modification cycles (Ap or Bp) with changing total kinase concentration ($K_{Total}$)

Note: Substrate biphasic response however in the modified form of either substrate is possible along total substrate amount and is shown in figure 9 in the main text.

2. Separate kinase - common phosphatase model:
   a. Substrate biphasic response is impossible in the modified substrate forms (Ap or Bp) with changing total substrate concentrations (total substrate concentrations of A or B respectively)
   b. Enzyme biphasic response is impossible in the modified substrate forms with either total kinase concentrations.

3. Common kinase - separate phosphatase model:
   a. Substrate biphasic response is impossible in the modified substrate forms (Ap or Bp) with changing total substrate concentrations (total substrate concentration of A or B respectively)
   b. Enzyme biphasic response is impossible in the modified form of either substrate with total kinase concentrations.

These results are summarized in the following tabular column.

**Table 1: Substrate and Enzyme biphasic dose responses in the coupled covalent modification system**

| System | Substrate Biphasic | Enzyme Biphasic |
|---|---|---|
| Common Kinase Common Phosphatase | Yes (see figure N-1 in main text) | Not possible |
| Separate Kinase Common Phosphatase | Not possible | Not possible |
| Common Kinase Separate Phosphatase | Not possible | Not possible |

## Common enzyme model: Presence of substrate biphasic and absence of enzyme biphasic response

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_2 \cdot ApP - k_{b1} \cdot A \cdot K + k_{ub1} \cdot AK :$$
$$dAp := k_1 \cdot AK + k_{ub2} \cdot ApP - k_{b2} \cdot Ap \cdot P :$$
$$dAK := k_{b1} \cdot A \cdot K - \left(k_{ub1} + k_1\right) \cdot AK :$$
$$dApP := k_{b2} \cdot Ap \cdot P - \left(k_{ub2} + k_2\right) \cdot ApP :$$

$$dB := p_2 \cdot BpP - p_{b1} \cdot B \cdot K + p_{ub1} \cdot BK :$$
$$dBp := p_1 \cdot BK + p_{ub2} \cdot BpP - p_{b2} \cdot Bp \cdot P :$$
$$dBK := p_{b1} \cdot B \cdot K - \left(p_{ub1} + p_1\right) \cdot BK :$$
$$dBpP := p_{b2} \cdot Bp \cdot P - \left(p_{ub2} + p_2\right) \cdot BpP :$$

$$dK := -k_{b1} \cdot A \cdot K + \left(k_{ub1} + k_1\right) \cdot AK - p_{b1} \cdot B \cdot K + \left(p_{ub1} + p_1\right) \cdot BK :$$
$$dP := -k_{b2} \cdot Ap \cdot P + \left(k_{ub2} + k_2\right) \cdot ApP - p_{b2} \cdot Bp \cdot P + \left(p_{ub2} + p_2\right) \cdot BpP :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + ApP + AK - A_{Total} :$$
$$BCon := B + Bp + BpP + BK - B_{Total} :$$
$$KCon := K + AK + BK - K_{Total} :$$
$$PCon := P + ApP + BpP - P_{Total} :$$

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of Ap and Bp. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$

We similarly solve for the other variables using the same command.

$$assign(solve(\{dBK, dApP, dBpP\}, \{BK, ApP, BpP\}))$$

$assign(solve(\{dA, dB\}, \{A, B\}))$

We introduce the following parameters ($c_1$, $c_2$, $d_1$, and $d_2$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) : k_{b2} := c_2 \cdot (k_2 + k_{ub2}) :$$
$$p_{b1} := d_1 \cdot (p_1 + p_{ub1}) : p_{b2} := d_2 \cdot (p_2 + p_{ub2}) :$$

Once this is done, we again solve for the steady state of the phosphatase, using the conservation expression for the enzyme (PCon).

$$P := solve(PCon, P) = \frac{P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration Ap and Bp

$$A = \frac{c_2\, Ap\, P_{Total}\, k_2}{(Ap\, c_2 + Bp\, d_2 + 1)\, K k_1\, c_1}$$
$$B = \frac{d_2\, Bp\, P_{Total}\, p_2}{(Ap\, c_2 + Bp\, d_2 + 1)\, K p_1\, d_1}$$
$$AK = \frac{c_2\, Ap\, P_{Total}\, k_2}{(Ap\, c_2 + Bp\, d_2 + 1)\, k_1}$$
$$BK = \frac{d_2\, Bp\, P_{Total}\, p_2}{(Ap\, c_2 + Bp\, d_2 + 1)\, p_1}$$
$$ApP = \frac{c_2\, Ap\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$
$$BpP = \frac{d_2\, Bp\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1}$$

Note that when Ap and Bp are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of Ap and Bp. We now have three expressions, ACon, BCon and KCon - the conservation of the substrates and kinase, whose solution for the variables define the steady state of the system.

## Enzyme biphasic

In this subsection we show the absence of enzyme biphasic in either of the modified forms of the substrates with total kinase amounts. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

Note: Since the kinase and the phosphatase are shared by the two covalent modification cycles, the individual covalent cycles (A=Ap, B=Bp) are equivalent. Thus, a proof presenting the absence of enzyme biphasic in one covalent modification cycle is tantamount to proving the absence in both modification cycle.

$ACon$

$$\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,k_1\,c_1}+Ap+\frac{c_2\,Ap\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}-A_{Total} \tag{1.1.1}$$

*BCon*

$$\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,p_1\,d_1}+Bp+\frac{d_2\,Bp\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-B_{Total} \tag{1.1.2}$$

*KCon*

$$K+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-K_{Total} \tag{1.1.3}$$

If we are to differentiate the expressions by total kinase amounts we would have the following expressions

$$\frac{dACon}{dK_{Total}}=0=\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$
$$\frac{dBCon}{dK_{Total}}=0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}+\frac{\partial BCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$
$$\frac{dKCon}{dK_{Total}}=0=\frac{\partial KCon}{\partial K_{Total}}+\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

Now in order to show the absence of biphasic response in Bp with total kinase concentration, we proceed with a proof by contradiction.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial K_{Total}}=0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0=\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$
$$0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}+\frac{\partial BCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$
$$1=\frac{\partial K}{\partial K_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}$, $\frac{\partial BCon}{\partial Bp}$, $\frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial K_{Total}}$ can be zero.

$$simplify(diff(ACon, Bp)) = -\frac{P_{Total}\, Ap\, d_2\, c_2\, \left(K\left(k_1 + k_2\right) c_1 + k_2\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K\, k_1\, c_1}$$

$$simplify(diff(BCon, Bp)) =$$

$$\frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K\, p_1\, d_1}\left(K\left(Bp^2\, p_1\, {d_2}^2 + 2\left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2\, P_{Total}}{2}\right)\left(Ap\, c_2 + 1\right) d_2 + p_1\left(Ap\, c_2 + 1\right)^2\right) d_1 + d_2\, p_2\, P_{Total}\left(Ap\, c_2 + 1\right)\right)$$

$$simplify(diff(KCon, Bp)) = \frac{P_{Total}\left(Ap\left(p_2\, k_1 - k_2\, p_1\right) c_2 + p_2\, k_1\right) d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1\, p_1}$$

Now solving the final expression obtained above after differentiation for $\frac{\partial K}{\partial K_{Total}}$ and resubstituting it in the first expression (as shown below) we get,

$$\frac{\partial K}{\partial K_{Total}} = 1 - \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} \cdot \left(1 - \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}\right)$$

This substitution is valid, provided $\frac{\partial ACon}{\partial K}$ is non-zero. This is shown below to be the case for any feasible steady state concentration. Thus, simplifying this further, we get

$$\frac{\partial ACon}{\partial K} \qquad \textbf{(1.1.4)}$$

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K^2\, k_1\, c_1}$$

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} - \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} \cdot \frac{\partial ACon}{\partial K}$$

However we know that $\frac{\partial KCon}{\partial K} = 1$. Thus, the expression simplifies to

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K} - \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

which simplifies further to result in

$$\frac{\partial ACon}{\partial K} = 0$$

However this expression (which can be computed as a function of the steady state concentrations - see below) cannot be zero for any feasible steady state.

$$simplify(diff(ACon, K)) =$$

$$-\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K^2\,k_1\,c_1}$$

**We have a contradiction. Thus enzyme biphasic cannot exist in Bp with total kinase concentration ($K_{Total}$).**

## Substrate biphasic

In this subsection we show the presnce of substrate biphasic in the modified form of a substrate form with the respecitve total amount of substrate as dose. We will further show how this substrate biphasic is guaranteed to exist in exactly one of the covalent modification cycles irrespective of the kinetic regime of the (de)modifications. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

Note: Since the kinase and the phosphatase are shared by the two covalent modification cycles, the individual covalent cycles (A=Ap, B=Bp) are equivalent. Thus, a proof presenting the presence of substrate biphasic in one covalent modification cycle is tantamount to proving the presence in the other modification cycle. We thus proceed by focussing on the presence of biphasic dose response with $A_{Total}$ in this sub-section first.

*ACon*

$$\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,k_1\,c_1}+Ap+\frac{c_2\,Ap\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}-A_{Total} \quad \textbf{(1.2.1)}$$

*KCon*

$$K+\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2+Bp\,d_2+1\right)k_1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-K_{Total} \quad \textbf{(1.2.2)}$$

*BCon*

$$\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)K\,p_1\,d_1}+Bp+\frac{d_2\,Bp\,P_{Total}}{Ap\,c_2+Bp\,d_2+1}+\frac{d_2\,Bp\,P_{Total}\,p_2}{\left(Ap\,c_2+Bp\,d_2+1\right)p_1}-B_{Total} \quad \textbf{(1.2.3)}$$

If we are to differentiate the expressions by total substrate amount ($A_{Total}$) we would have the following expressions

$$\frac{dACon}{dA_{Total}}=0=\frac{\partial ACon}{\partial A_{Total}}+\frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial ACon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}+\frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

$$\frac{dBCon}{dA_{Total}}=0=\frac{\partial BCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}+\frac{\partial BCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}$$

$$\frac{dKCon}{dA_{Total}}=0=\frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial A_{Total}}+\frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial A_{Total}}+\frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial A_{Total}}$$

Now in order to show the presence of biphasic response in Ap with total substrate concentration, we proceed with by checking if the necessary conditions for the behavior to exist can be satisfied.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Ap}{\partial A_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$1 = \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$
$$0 = \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}} + \frac{\partial BCon}{\partial K} \cdot \frac{\partial K}{\partial A_{Total}}$$
$$0 = 1 \cdot \frac{\partial K}{\partial A_{Total}} + \frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Ap}, \frac{\partial BCon}{\partial Ap}, \frac{\partial KCon}{\partial Ap}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Ap}{\partial A_{Total}}$ can be zero.

$$simplify(diff(ACon, Ap)) = -\frac{P_{Total}\, Ap\, d_2\, c_2 \left(K \left(k_1 + k_2\right) c_1 + k_2\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K\, k_1\, c_1}$$

$$simplify(diff(BCon, Ap)) =$$
$$\frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K\, p_1\, d_1} \left(K \left(Bp^2\, p_1\, {d_2}^2 + 2 \left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2\, P_{Total}}{2}\right) \left(Ap\, c_2 + 1\right) d_2 + p_1 \left(Ap\, c_2 + 1\right)^2\right) d_1 + d_2\, p_2\, P_{Total} \left(Ap\, c_2 + 1\right)\right)$$

$$simplify(diff(KCon, Ap)) = \frac{P_{Total} \left(Ap \left(p_2\, k_1 - k_2\, p_1\right) c_2 + p_2\, k_1\right) d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1\, p_1}$$

Now solving the final expression obtained above after differentiation for $\frac{\partial K}{\partial K_{Total}}$ and resubstituting it in the other expressions (as shown below) we get,

$$1 = \frac{\partial Bp}{\partial A_{Total}} \cdot \left(\frac{\partial ACon}{\partial Bp} - \frac{\partial ACon}{\partial K} \cdot \frac{\partial KCon}{\partial Bp}\right)$$
$$0 = \frac{\partial Bp}{\partial A_{Total}} \cdot \left(\frac{\partial BCon}{\partial Bp} - \frac{\partial BCon}{\partial K} \cdot \frac{\partial KCon}{\partial Bp}\right)$$
$$\frac{\partial K}{\partial A_{Total}} = -\frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial A_{Total}}$$

From the first expression, it is easy to discern that $\frac{\partial Bp}{\partial A_{Total}}$ cannot be equal to zero. Thus this implies that the the expression contained in the brackets in the second equation has to equal zero for a biphasic dose response to exist. We thus isolate this expression as shown below,

$$simplify(diff(BCon, Bp) - diff(BCon, K)\cdot diff(KCon, Bp))$$

$$\frac{1}{(Ap\,c_2 + Bp\,d_2 + 1)^3\,K^2\,p_1^{\,2}\,d_1\,k_1}\Big(\Big(\big(Bp^2\,d_2^{\,2} + (Ap\,c_2 + 1)\,(P_{Total} + 2\,Bp)\big)\,d_2 + (Ap\,c_2 + 1)^2\Big)\,(Ap\,c_2 + Bp\,d_2 + 1)\,K^2\,d_1\,p_1^{\,2} + K\,p_2\,P_{Total}\,d_2\,(Ap\,c_2 + 1)\,(Ap\,c_2 + Bp\,d_2 + 1)\,(K\,d_1 + 1)\,p_1 + p_2^{\,2}\,P_{Total}^{\,2}\,Bp\,d_2^{\,2}\,(Ap\,c_2 + 1)\Big)\,k_1 - Ap\,Bp\,c_2\,d_2^{\,2}\,k_2\,p_1\,p_2\,P_{Total}^{\,2}\Big) \tag{1.2.4}$$

Collecting only the numerator, and writing it as a polynomial in $P_{Total}$, we get

$$collect\Big(\Big(\big(Bp^2\,d_2^{\,2} + (Ap\,c_2 + 1)\,(P_{Total} + 2\,Bp)\big)\,d_2 + (Ap\,c_2 + 1)^2\Big)\,(Ap\,c_2 + Bp\,d_2 + 1)\,K^2\,d_1\,p_1^{\,2} + K\,p_2\,P_{Total}\,d_2\,(Ap\,c_2 + 1)\,(Ap\,c_2 + Bp\,d_2 + 1)\,(K\,d_1 + 1)\,p_1 + p_2^{\,2}\,P_{Total}^{\,2}\,Bp\,d_2^{\,2}\,(Ap\,c_2 + 1)\Big)\,k_1 - Ap\,Bp\,c_2\,d_2^{\,2}\,k_2\,p_1\,p_2\,P_{Total}^{\,2}, P_{Total}\Big)$$

$$\Big(p_2^{\,2}\,Bp\,d_2^{\,2}\,(Ap\,c_2 + 1)\,k_1 - Ap\,Bp\,c_2\,d_2^{\,2}\,k_2\,p_1\,p_2\Big)\,P_{Total}^{\,2} + \Big((Ap\,c_2 + 1)\,d_2\,(Ap\,c_2 + Bp\,d_2 + 1)\,K^2\,d_1\,p_1^{\,2} + K\,p_2\,d_2\,(Ap\,c_2 + 1)\,(Ap\,c_2 + Bp\,d_2 + 1)\,(K\,d_1 + 1)\,p_1\Big)\,k_1\,P_{Total} + \Big(Bp^2\,d_2^{\,2} + 2\,(Ap\,c_2 + 1)\,Bp\,d_2 + (Ap\,c_2 + 1)^2\Big)\,(Ap\,c_2 + Bp\,d_2 + 1)\,K^2\,d_1\,p_1^{\,2}\,k_1 \tag{1.2.5}$$

We can thus observe that for a feasible steady state of the system, only the coeffecient of the second (leading) exponent of PTotal, can be negative. The other coeffecients are strictly positive. We now isolate this coeffecient below,

$$collect\Big(simplify\Big(Ap\,c_2^{\,2}\,k_2^{\,2}\,(Bp\,d_2 + 1)\,p_1 - Ap\,Bp\,c_2^{\,2}\,d_2\,k_1\,k_2\,p_2\Big), \{Bp, Ap, d_2\}\Big)$$

$$\big(-(k_1\,p_2 - p_1\,k_2)\,k_2\,c_2^{\,2}\,d_2\,Bp + c_2^{\,2}\,k_2^{\,2}\,p_1\big)\,Ap \tag{1.2.6}$$

This implies that, depending on the sign of $k_1\,p_2 - k_2\,p_1$, the coeffecient can be negative or strictly positive (in which case the equation, 1.5, can never be zero for a feasible steady state ruling out substrate biphasic dose response). If $k_1\,p_2 - p_1\,k_2 > 0$, then for a sufficiently large Bp concentration, the coeffecient is negative (irrespective of the value of Ap), guaranteeing that there exists a solution for equation 1.5 (featuring any value of Ap and K at some $P_{Total}$).

Thus for some finite amount of $K_{Total}$ and $B_{Total}$, it is possible to obtain substrate biphasic response in Ap for changing $A_{Total}$, if $k_1\,p_2 - p_1\,k_2 > 0$.

However if

$k_1 p_2 - p_1 k_2 < 0$, substrate biphasic in Ap with ATotal is strictly not possible.

Here by leveraging the fact that the two covalent cycles are essentially equivalent (A=Ap with the kinetic nomenclature $k_i$, and B=Bp with the kinetic nomenclature $p_i$), we can see that if $k_1 p_2 - p_1 k_2 < 0$, then the condition for substrate biphasic in Bp with $B_{Total}$ is trivially satisfied again indicating that there exists some total finite amount of $K_{Total}$ and $A_{Total}$, where substrate biphasic in Bp with $B_{Total}$ is guaranteed to exist.

Thus to conclude - this proves that irrespective of the kinetic regime, either Ap or Bp is exclusively guaranteed to (for some finite amount of substrate and enzyme) present with biphasic dose repsonse with the respective substrates. Specifically, if $k_1 p_2 - p_1 k_2 > 0$ then Ap is capable of substrate biphasic while Bp is not, and when $k_1 p_2 - p_1 k_2 < 0$ then Bp is capable of substrate biphasic while Ap is not.

## Separate kinase common phosphatase model: Absence of substrate and enzyme biphasic dose response

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :
```

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

```
dA := k2·ApP − kb1·A·K1 + kub1·AK1 :
dAp := k1·AK1 + kub2·ApP − kb2·Ap·P :

dAK1 := kb1·A·K1 − (kub1 + k1)·AK1 :
dApP := kb2·Ap·P − (kub2 + k2)·ApP :

dB := p2·BpP − pb1·B·K2 + pub1·BK2 :
dBp := p1·BK2 + pub2·BpP − pb2·Bp·P :

dBK2 := pb1·B·K2 − (pub1 + p1)·BK2 :
dBpP := pb2·Bp·P − (pub2 + p2)·BpP :

dK1 := −kb1·A·K1 + (kub1 + k1)·AK1 :
dK2 := −pb1·B·K2 + (pub1 + p1)·BK2 :
dP := −pb1·B·K2 + (pub1 + p1)·BK2 − pb2·Bp·P + (pub2 + p2)·BpP :
```

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon, PCon, K1Con and K2Con for the substrates and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + ApP + AK1 - A_{Total}:$$
$$BCon := B + Bp + BpP + BK2 - B_{Total}:$$
$$K1Con := K1 + AK1 - K1_{Total}:$$
$$K2Con := K2 + BK2 - K2_{Total}:$$
$$PCon := P + ApP + BpP - P_{Total}:$$

We introduce the following parameters ($c_1$, $c_2$, $d_1$, and $d_2$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) : k_{b2} := c_2 \cdot (k_2 + k_{ub2}) :$$
$$p_{b1} := d_1 \cdot (p_1 + p_{ub1}) : p_{b2} := d_2 \cdot (p_2 + p_{ub2}) :$$

Once this is done, we again solve for the steady state of the phosphatase, using the conservation expression for the enzyme (PCon).

$$AK1 := solve(dAK1, AK1) = K1\ A\ c_1$$

We similarly solve for the other variables using the same command.

$$assign(solve(\{dBK2, dApP, dBpP\}, \{BK2, ApP, BpP\}))$$
$$assign(solve(\{dA, dB\}, \{A, B\}))$$

Once this is done, we again solve for the steady state of the phosphatase, using the conservation expression for the enzyme (PCon).

$$P := simplify(solve((PCon), P))$$

$$P := \frac{P_{Total}}{Ap\ c_2 + Bp\ d_2 + 1} \tag{2.1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration Ap and Bp

$$A = \frac{c_2\ Ap\ P_{Total}\ k_2}{(Ap\ c_2 + Bp\ d_2 + 1)\ K1\ c_1\ k_1}$$
$$B = \frac{d_2\ Bp\ P_{Total}\ p_2}{(Ap\ c_2 + Bp\ d_2 + 1)\ K2\ d_1\ p_1}$$
$$AK1 = \frac{c_2\ Ap\ P_{Total}\ k_2}{(Ap\ c_2 + Bp\ d_2 + 1)\ k_1}$$
$$BK2 = \frac{d_2\ Bp\ P_{Total}\ p_2}{(Ap\ c_2 + Bp\ d_2 + 1)\ p_1}$$
$$ApP = \frac{c_2\ Ap\ P_{Total}}{Ap\ c_2 + Bp\ d_2 + 1}$$
$$BpP = \frac{d_2\ Bp\ P_{Total}}{Ap\ c_2 + Bp\ d_2 + 1}$$

## Substrate biphasic

In this subsection we show the absence of substrate biphasic in either of the modified forms of the substrates with total substrate amounts. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, K1Con, K2Con - shown below) would define the steady state of the system.

Note: Since the two covalent modification cycles are virtually identical to one and other with a shared phosphatase and two unique kinases acting on them, a proof presenting the absence of substrate biphasic in one covalent modification cycle is tantamount to proving the absence in both modification cycle.

*ACon*

$$\frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K1\, c_1\, k_1} + Ap + \frac{c_2\, Ap\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1} + \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1} - A_{Total} \tag{2.1.1}$$

*BCon*

$$\frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K2\, d_1\, p_1} + Bp + \frac{d_2\, Bp\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1} + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} - B_{Total} \tag{2.1.2}$$

*K1Con*

$$K1 + \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1} - K1_{Total} \tag{2.1.3}$$

*K2Con*

$$K2 + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} - K2_{Total} \tag{2.1.4}$$

Now differentiating each of these expressions with respect to $B_{Total}$, we get the following

$$\frac{dACon}{dB_{Total}} = 0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial B_{Total}}$$

$$\frac{dBCon}{dB_{Total}} = 0 = \frac{\partial BCon}{\partial B_{Total}} + \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}} + \frac{\partial BCon}{\partial K2} \cdot \frac{\partial K2}{\partial B_{Total}}$$

$$\frac{dK1Con}{dB_{Total}} = 0 = \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial B_{Total}} + \frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial K1Con}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}}$$

$$\frac{dK2Con}{dB_{Total}} = 0 = \frac{\partial K2Con}{\partial K2} \cdot \frac{\partial K2}{\partial B_{Total}} + \frac{\partial K2Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial K2Con}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}}$$

Now in order to show the absence of biphasic response in Bp with total substrate concentration ($B_{Total}$), we proceed with a proof by contradiction. Note that the same procedure of which will be valid to show the absence of substrate biphasic response in modified form Ap with $A_{Total}$.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial B_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial B_{Total}}$$
$$1 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial BCon}{\partial K2} \cdot \frac{\partial K2}{\partial B_{Total}}$$
$$0 = \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial B_{Total}} + \frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}$$
$$0 = \frac{\partial K2Con}{\partial K2} \cdot \frac{\partial K2}{\partial B_{Total}} + \frac{\partial K2Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial K1Con}{\partial Bp}, \frac{\partial K2Con}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial B_{Total}}$ can be zero.

$$simplify(diff(ACon, Bp)) = -\frac{P_{Total}\, d_2\, c_2 \left(K1 \left(k_1 + k_2\right) c_1 + k_2\right) Ap}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K1\, c_1\, k_1}$$

$$simplify(diff(BCon, Bp)) = \frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K2\, d_1\, p_1} \left(K2 \left(Bp^2\, p_1\, d_2^{\,2} + 2 \left(Ap\, c_2 + 1\right) \left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2\, P_{Total}}{2}\right) d_2 + p_1 \left(Ap\, c_2 + 1\right)^2\right) d_1 + d_2\, p_2\, P_{Total} \left(Ap\, c_2 + 1\right)\right)$$

$$simplify(diff(K1Con, Bp)) = -\frac{c_2\, Ap\, P_{Total}\, k_2\, d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1}$$

$$simplify(diff(K2Con, Bp)) = \frac{d_2\, p_2\, P_{Total} \left(Ap\, c_2 + 1\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 p_1}$$

Now solving the expression obtained by differentiating K1Con above for $\frac{\partial K1}{\partial B_{Total}}$ and resubstituting it in the first expression (as shown below) we get,

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \left(-\frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}\right)$$

which further simplifies to

$$0 = \frac{\partial Ap}{\partial B_{Total}} \cdot \left(\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1Con}{\partial Ap}\right)$$

***Note***

This simplification is valid only if $\frac{\partial K1}{\partial B_{Total}}$ is non-zero at the biphasic peak. This can be easily showed by the following argument.

Suppose, $\frac{\partial K1}{\partial B_{Total}}$ is zero at such a point. Then, by analyzing the other expressions, we get $\frac{\partial Ap}{\partial B_{Total}}$ to be zero. Thus by extension would imply that $\frac{\partial K2}{\partial B_{Total}}$ is zero, and the resulting expression upon differentiating BCon, can never be satisfied. Thus at the biphasic peak (should it exist) $\frac{\partial K1}{\partial B_{Total}}$ is non-zero, allowing us to proceed with our substitution.

This thus leads to the following conclusion, either $\frac{\partial Ap}{\partial B_{Total}}$ is zero, or $\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1Con}{\partial Ap}$ is zero. The former cannot be possible (see note above).

However the latter can also not be possible for a feasible steady state of the system (see below).

$$simplify(diff(ACon, Ap) - diff(ACon, K1) \cdot diff(K1Con, Ap))$$

$$\frac{1}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)^3 K1^2\,c_1\,k_1^2}\left(c_1\,K1^2\left(Ap^2\,c_2^2 + 2\left(Ap + \frac{P_{Total}}{2}\right)\left(Bp\,d_2 + 1\right)c_2 + \left(Bp\,d_2 + 1\right)^2\right)\left(Ap\,c_2 + Bp\,d_2 + 1\right)k_1^2 + K1\,c_2\,k_2\,P_{Total}\left(K1\,c_1 + 1\right)\left(Bp\,d_2 + 1\right)\left(Ap\,c_2 + Bp\,d_2 + 1\right)k_1 + c_2^2\,Ap\,P_{Total}^2\,k_2^2\left(Bp\,d_2 + 1\right)\right) \quad \textbf{(2.1.5)}$$

Since all the parameters and variables involved in the expression are positive, the expression is non-zero always.

We have a contradiction.

Thus, our assumption of the presence of a substrate biphasic in Bp with $B_{Total}$ is wrong.

**Hence substrate biphasic cannot exist in the model.**

## Enzyme biphasic

In this subsection we show the absence of enzyme biphasic in either of the modified forms of the substrates with total kinase concentration (between the respective substrate enzyme pair). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, K1Con, K2Con - shown below) would define the steady state of the system.

Note: Since the two covalent modification cycles are virtually identical to one and other with a shared phosphatase and two unique kinases acting on them, a proof presenting the absence of enzyme biphasic in one covalent modification cycle is tantamount to proving the absence in both modification cycle.

$$ACon$$

$$\frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)K1\,c_1\,k_1} + Ap + \frac{c_2\,Ap\,P_{Total}}{Ap\,c_2 + Bp\,d_2 + 1} + \frac{c_2\,Ap\,P_{Total}\,k_2}{\left(Ap\,c_2 + Bp\,d_2 + 1\right)k_1} \quad \textbf{(2.2.1)}$$

$$-A_{Total}$$

*BCon*

$$\frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) K2\, d_1\, p_1} + Bp + \frac{d_2\, Bp\, P_{Total}}{Ap\, c_2 + Bp\, d_2 + 1} + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} \quad \textbf{(2.2.2)}$$

$$-B_{Total}$$

*K1Con*

$$K1 + \frac{c_2\, Ap\, P_{Total}\, k_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) k_1} - K1_{Total} \quad \textbf{(2.2.3)}$$

*K2Con*

$$K2 + \frac{d_2\, Bp\, P_{Total}\, p_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right) p_1} - K2_{Total} \quad \textbf{(2.2.4)}$$

Now differentiating each of these expressions with respect to $K1_{Total}$, we get the following

$$\frac{dACon}{dK2_{Total}} = 0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial ACon}{\partial Bp} \cdot \frac{\partial Bp}{\partial K2_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial K2_{Total}}$$
$$\frac{dBCon}{dK2_{Total}} = 0 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial K2_{Total}} + \frac{\partial BCon}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}}$$
$$\frac{dK1Con}{dK2_{Total}} = 0 = \frac{\partial K1Con}{\partial K1} \cdot \frac{\partial K1}{\partial K2_{Total}} + \frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial K1Con}{\partial Bp} \cdot \frac{\partial Bp}{\partial K2_{Total}}$$
$$\frac{dK2Con}{dK2_{Total}} = 0 = \frac{\partial K2Con}{\partial K2_{Total}} + \frac{\partial K2Con}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}} + \frac{\partial K2Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial K2Con}{\partial Bp} \cdot \frac{\partial Bp}{\partial K2_{Total}}$$

Now in order to show the absence of biphasic response in Bp with total kinase concentration ($K2_{Total}$), we proceed with a proof by contradiction. Note that the same procedure of which will be valid to show the absence of substrate biphasic response in modified form Ap with $K1_{Total}$.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial K2_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1}{\partial K2_{Total}}$$
$$0 = \frac{\partial BCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial BCon}{\partial K2} \cdot \frac{\partial K2}{\partial K2_{Total}}$$
$$0 = \frac{\partial K1}{\partial K2_{Total}} + \frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}}$$
$$1 = \frac{\partial K2}{\partial K2_{Total}} + \frac{\partial K2Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ACon}{\partial Bp}, \frac{\partial BCon}{\partial Bp}, \frac{\partial K1Con}{\partial Bp}, \frac{\partial K2Con}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial K2_{Total}}$ can be zero.

$$simplify(diff(ACon, Bp)) = -\frac{P_{Total}\, d_2\, c_2 \left(K1 \left(k_1 + k_2\right) c_1 + k_2\right) Ap}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K1\, c_1\, k_1}$$

$$simplify(diff(BCon, Bp)) = \frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 K2\, d_1\, p_1} \left(K2 \left(Bp^2\, p_1\, d_2^{\,2} + 2 \left(Ap\, c_2 + 1\right) \left(\left(Bp + \frac{P_{Total}}{2}\right) p_1 + \frac{p_2\, P_{Total}}{2}\right) d_2 + p_1 \left(Ap\, c_2 + 1\right)^2\right) d_1 + d_2\, p_2\, P_{Total} \left(Ap\, c_2 + 1\right)\right)$$

$$simplify(diff(K1Con, Bp)) = -\frac{c_2\, Ap\, P_{Total}\, k_2\, d_2}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 k_1}$$

$$simplify(diff(K2Con, Bp)) = \frac{d_2\, p_2\, P_{Total} \left(Ap\, c_2 + 1\right)}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^2 p_1}$$

Now solving the expression obtained by differentiating K1Con above for $\frac{\partial K1}{\partial K2_{Total}}$ and resubstituting it in the first expression (as shown below) we get,

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}} + \frac{\partial ACon}{\partial K1} \cdot \left(-\frac{\partial K1Con}{\partial Ap} \cdot \frac{\partial Ap}{\partial K2_{Total}}\right)$$

which further simplifies to

$$0 = \frac{\partial Ap}{\partial K2_{Total}} \cdot \left(\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1Con}{\partial Ap}\right)$$

***Note***

This simplification is valid only if $\frac{\partial K1}{\partial K2_{Total}}$ is non-zero at the biphasic peak. This can be easily showed by the following argument.

Suppose, $\frac{\partial K1}{\partial K2_{Total}}$ is zero at such a point. Then, by analyzing the other expressions, we get $\frac{\partial Ap}{\partial K2_{Total}}$ to be zero. Thus by extension would imply that $\frac{\partial K2}{\partial K2_{Total}}$ is zero, and the resulting expression upon differentiating K2Con, can never be satisfied. Thus at the biphasic peak (should it exist) $\frac{\partial K1}{\partial K2_{Total}}$ is non-zero, allowing us to proceed with our substitution.

This thus leads to the following conclusion, either

$\frac{\partial Ap}{\partial K2_{Total}}$ is zero, or

$\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K1} \cdot \frac{\partial K1Con}{\partial Ap}$ is zero. The former cannot be possible (see note above).

However the latter can also not be possible for a feasible steady state of the system (see below).

$$simplify(diff(ACon, Ap) - diff(ACon, K1) \cdot diff(K1Con, Ap))$$

$$\frac{1}{\left(Ap\, c_2 + Bp\, d_2 + 1\right)^3 KI^2\, c_1\, k_1^2} \left( c_1\, KI^2 \left( Ap^2\, c_2^2 + 2 \left( Ap + \frac{P_{Total}}{2} \right) \left( Bp\, d_2 + 1 \right) c_2 \right. \right. \quad \textbf{(2.2.5)}$$

$$\left. + \left( Bp\, d_2 + 1 \right)^2 \right) \left( Ap\, c_2 + Bp\, d_2 + 1 \right) k_1^2 + KI\, c_2\, k_2\, P_{Total} \left( KI\, c_1 + 1 \right) \left( Bp\, d_2 \right.$$

$$\left. \left. + 1 \right) \left( Ap\, c_2 + Bp\, d_2 + 1 \right) k_1 + c_2^2\, Ap\, P_{Total}^2\, k_2^2 \left( Bp\, d_2 + 1 \right) \right)$$

Since all the parameters and variables involved in the expression are positive, the expression is non-zero always.

We have a contradiction.

Thus, our assumption of the presence of an enzyme biphasic in Bp with $K2_{Total}$ is wrong.

**Hence enzyme biphasic cannot exist in the model.**

## Common kinase separate phosphatase model: Absence of substrate and enzyme biphasic dose response

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_{b1} \cdot A \cdot K + k_{ub1} \cdot AK + k_2 \cdot ApP1 :$$
$$dAp := -k_{b2} \cdot Ap \cdot P1 + k_{ub2} \cdot ApP1 + k_1 \cdot AK :$$
$$dB := -p_{b1} \cdot B \cdot K + p_{ub1} \cdot BK + p_2 \cdot BpP2 :$$
$$dBp := -p_{b2} \cdot Bp \cdot P2 + p_{ub2} \cdot BpP2 + p_1 \cdot BK :$$

$$dAK := k_{b1} \cdot A \cdot K - \left( k_{ub1} + k_1 \right) \cdot AK :$$
$$dApP1 := k_{b2} \cdot Ap \cdot P1 - \left( k_{ub2} + k_2 \right) \cdot ApP1 :$$
$$dBK := p_{b1} \cdot B \cdot K - \left( p_{ub1} + p_1 \right) \cdot BK :$$
$$dBpP2 := p_{b2} \cdot Bp \cdot P2 - \left( p_{ub2} + p_2 \right) \cdot BpP2 :$$

$$dK := -k_{b1} \cdot A \cdot K + \left( k_{ub1} + k_1 \right) \cdot AK - p_{b1} \cdot B \cdot K + \left( p_{ub1} + p_1 \right) \cdot BK :$$

$dP1 := -k_{b2} \cdot Ap \cdot P1 + \left(k_{ub2} + k_2\right) \cdot ApP1 :$
$dP2 := -p_{b2} \cdot Bp \cdot P2 + \left(p_{ub2} + p_2\right) \cdot BpP2 :$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, BCon, KCon, P1Con and P2Con for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + ApP1 + AK - A_{Total} :$
$BCon := B + Bp + BpP2 + BK - B_{Total} :$
$KCon := K + AK + BK - K_{Total} :$
$P1Con := P1 + ApP1 - P1_{Total} :$
$P2Con := P2 + BpP2 - P2_{Total} :$

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of Ap and Bp. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$AK := solve(dAK, AK) = \frac{k_{b1}\, A\, K}{k_{ub1} + k_1}$$

We similarly solve for the other variables using the same command.

$assign(solve(\{dBK, dApP1, dBpP2\}, \{BK, ApP1, BpP2\}))$
$assign(solve(\{dA, dB\}, \{A, B\}))$

We introduce the following parameters ($c_1$, $c_2$, $d_1$, and $d_2$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) :$
$p_{b1} := d_1 \cdot \left(p_1 + p_{ub1}\right) : p_{b2} := d_2 \cdot \left(p_2 + p_{ub2}\right) :$

Once this is done, we again solve for the steady state of the phosphatases, using the respective conservation expressions (P1Con and P2Con).

$$P1 := solve(P1Con, P1) = \frac{P1_{Total}}{Ap\, c_2 + 1}$$

$$P2 := solve(P2Con, P2) = \frac{P2_{Total}}{Bp\, d_2 + 1}$$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration Ap and Bp

$$A = \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) K\, k_1\, c_1}$$

$$B = \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) K\, p_1\, d_1}$$

$$AK = \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) k_1}$$

$$BK = \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) p_1}$$

$$ApP1 = \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1}$$

$$BpP2 = \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2 + 1}$$

Note that when Ap and Bp are positive, steady state concentrations of the other variable concentrations are positive as well. Thus we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of Ap and Bp. We now have three expressions, ACon, BCon and KCon - the conservation of the substrates and kinase, whose solution for the variables define the steady state of the system.

## Substrate biphasic

In this subsection we show the absence of substrate biphasic in either of the modified forms of the substrates with their respective total substrate amounts. As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

Note: Since the two covalent modification cycles are virtually identical to one and other with a shared kinase and two unique phosphatases acting on them, a proof presenting the absence of substrate biphasic in one covalent modification cycle is tantamount to proving the absence in both modification cycle.

*ACon*

$$\frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) K\, k_1\, c_1} + Ap + \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2 + 1} + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) k_1} - A_{Total} \tag{3.1.1}$$

*BCon*

$$\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) K\, p_1\, d_1} + Bp + \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2 + 1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) p_1} - B_{Total} \tag{3.1.2}$$

*KCon*

$$K + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) k_1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) p_1} - K_{Total} \tag{3.1.3}$$

If we are to differentiate the expressions by total substrate amount ($B_{Total}$) we would have the following expressions

$$\frac{dACon}{dB_{Total}} = 0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$

$$\frac{dBCon}{dB_{Total}} = 0 = \frac{\partial BCon}{\partial B_{Total}} + \frac{\partial BCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}} + \frac{\partial BCon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$

$$\frac{dKCon}{dB_{Total}} = 0 = \frac{\partial KCon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial KCon}{\partial Bp} \cdot \frac{\partial Bp}{\partial B_{Total}}$$

Now in order to show the absence of biphasic response in Bp with total substrate concentration, we proceed with a proof by contradiction.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial B_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$

$$1 = \frac{\partial BCon}{\partial K} \cdot \frac{\partial K}{\partial B_{Total}}$$

$$0 = \frac{\partial K}{\partial B_{Total}} + \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial Bp}{\partial K_{Total}}$ can be zero.

$$simplify(diff(BCon, Bp)) = \frac{\left(Bp^2\, d_2{}^2\, p_1 + \left(\left(2\, Bp + P2_{Total}\right) p_1 + p_2\, P2_{Total}\right) d_2 + p_1\right) K\, d_1 + d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 K\, p_1\, d_1}$$

$$simplify(diff(KCon, Bp)) = \frac{d_2\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right)^2 p_1}$$

We also know that $\frac{\partial BCon}{\partial K}$ is non-zero (See below). Thus we can assert that $\frac{\partial K}{\partial B_{Total}}$ is also non-zero.

$$simplify(diff(BCon, K)) = -\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2 + 1\right) K^2\, p_1\, d_1}$$

This allows us to solve for it using the expression obtained from differentiating KCon and substituting in the expression belonging to ACon.

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}} + \frac{\partial ACon}{\partial K} \cdot \left(-\frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial B_{Total}}\right)$$

This expression can be further simplified as shown below

$$0 = \frac{\partial Ap}{\partial B_{Total}} \cdot \left( \frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K} \frac{\partial KCon}{\partial Ap} \right)$$

Now, in order for this to be true, either the $\frac{\partial Ap}{\partial B_{Total}}$ is zero (which it cannot be - see note below), or $\frac{\partial ACon}{\partial Ap} - \frac{\partial ACon}{\partial K} \frac{\partial KCon}{\partial Ap}$ is zero. We show below that latter can also not be zero for any feasible steady state concentration.

***Note***

Suppose $\frac{\partial Ap}{\partial B_{Total}}$ is zero. Then, this implies that $\frac{\partial ACon}{\partial K}$ must be zero (this is since, $\frac{\partial K}{\partial B_{Total}}$ is non-zero). However, $\frac{\partial ACon}{\partial K}$ is non-zero as shown below.

$$simplify(diff(ACon, K)) = -\frac{c_2\, Ap\, PI_{Total}\, k_2}{\left(Ap\, c_2 + 1\right) K^2\, k_1\, c_1}$$

$$simplify(diff(ACon, Ap) - diff(ACon, K) \cdot diff(KCon, Ap)) =$$
$$\frac{1}{\left(c_2\, Ap + 1\right)^3 K^2\, k_1^{\,2}\, c_1} \left( Ap^3\, K^2\, c_1\, c_2^{\,3}\, k_1^{\,2} + 3 \left( K^2\, c_1 \left( Ap + \frac{PI_{Total}}{3} \right) k_1^{\,2} + \frac{K\, k_2\, PI_{Total} \left(K\, c_1 + 1\right) k_1}{3} + \frac{k_2^{\,2}\, PI_{Total}^{\,2}}{3} \right) Ap\, c_2^{\,2} + 3 \left( K\, c_1 \left( Ap + \frac{PI_{Total}}{3} \right) k_1 + \frac{k_2\, PI_{Total} \left(K\, c_1 + 1\right)}{3} \right) k_1\, K\, c_2 + K^2\, k_1^{\,2}\, c_1 \right)$$

Since all the parameters and variables involved in the expression are positive, the expression is non-zero always.

We have a contradiction.

Thus, our assumption of the presence of a substrate biphasic in Bp with $B_{Total}$ is wrong.

**Hence substrate biphasic cannot exist in the model.**

## Enzyme biphasic

In this subsection we show the absence of enzyme biphasic in either of the modified forms of the substrates with total kinase concentration (between the respective substrate enzyme pair). As mentioned earlier, feasible solutions to the three coupled expressions (ACon, BCon, KCon - shown below) would define the steady state of the system.

Note: Since the two covalent modification cycles are virtually identical to one and other with a shared kinase and two unique phosphatases acting on them, a proof presenting the absence of enzyme biphasic in one covalent modification cycle is tantamount to proving the absence in both modification cycle.

*ACon*

$$\frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2+1\right) K k_1 c_1} + Ap + \frac{c_2\, Ap\, P1_{Total}}{Ap\, c_2+1} + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2+1\right) k_1} - A_{Total} \tag{3.2.1}$$

*BCon*

$$\frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2+1\right) K p_1 d_1} + Bp + \frac{d_2\, Bp\, P2_{Total}}{Bp\, d_2+1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2+1\right) p_1} - B_{Total} \tag{3.2.2}$$

*KCon*

$$K + \frac{c_2\, Ap\, P1_{Total}\, k_2}{\left(Ap\, c_2+1\right) k_1} + \frac{d_2\, Bp\, P2_{Total}\, p_2}{\left(Bp\, d_2+1\right) p_1} - K_{Total} \tag{3.2.3}$$

Now differentiating each of these expressions with respect to $K1_{Total}$, we get the following

$$\frac{dACon}{dK_{Total}} = 0 = \frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$\frac{dBCon}{dK_{Total}} = 0 = \frac{\partial BCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}} + \frac{\partial BCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}} + \frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}} + \frac{\partial KCon}{\partial Bp}\cdot\frac{\partial Bp}{\partial K_{Total}}$$

Now in order to show the absence of biphasic response in Bp with total kinase concentration ($K_{Total}$), we proceed with a proof by contradiction. Note that the same procedure of which will be valid to show the absence of substrate biphasic response in modified form Ap with $K_{Total}$.

If we assume that such a biphasic response exists, then there should exist a steady state of the system where the following should be satisfied (at the biphasic peak)

$$\frac{\partial Bp}{\partial K_{Total}} = 0$$

Thus, the above expressions obtained after differentiation can be simplified as follows

$$0 = \frac{\partial ACon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}} + \frac{\partial ACon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$0 = \frac{\partial BCon}{\partial K}\cdot\frac{\partial K}{\partial K_{Total}}$$

$$1 = \frac{\partial K}{\partial K_{Total}} + \frac{\partial KCon}{\partial Ap}\cdot\frac{\partial Ap}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial BCon}{\partial Bp}, \frac{\partial KCon}{\partial Bp}$ have a non-zero denominator

(as shown below) and thus the products involving $\frac{\partial Bp}{\partial K_{Total}}$ can be zero.

$$simplify(diff(BCon, Bp)) = \frac{K\left(Bp^2 d_2^2 p_1 + \left(\left(2\,Bp + P2_{Total}\right) p_1 + p_2 P2_{Total}\right) d_2 + p_1\right) d_1 + d_2 P2_{Total} p_2}{\left(Bp\,d_2 + 1\right)^2 K p_1 d_1}$$

$$simplify(diff(KCon, Bp)) = \frac{d_2 P2_{Total} p_2}{\left(Bp\,d_2 + 1\right)^2 p_1}$$

Now since $\frac{\partial BCon}{\partial K}$ is non-zero (see below), $\frac{\partial K}{\partial K_{Total}}$ must be zero.

$$simplify(diff(BCon, K)) = -\frac{d_2 Bp\, P2_{Total} p_2}{\left(Bp\,d_2 + 1\right) K^2 p_1 d_1}$$

Again, since $\frac{\partial ACon}{\partial K}$ has a non-zero denominator (see below), the products involving $\frac{\partial K}{\partial K_{Total}}$ can be zero.

$$simplify(diff(BCon, K)) = \frac{K\left(Bp^2 d_2^2 p_1 + \left(\left(2\,Bp + P2_{Total}\right) p_1 + p_2 P2_{Total}\right) d_2 + p_1\right) d_1 + d_2 P2_{Total} p_2}{\left(Bp\,d_2 + 1\right)^2 K p_1 d_1}$$

This results in the following expressions after reduction for the differentiated expressions

$$0 = \frac{\partial ACon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

$$1 = \frac{\partial KCon}{\partial Ap} \cdot \frac{\partial Ap}{\partial K_{Total}}$$

However, we have a contradiction here. Both $\frac{\partial ACon}{\partial Ap}$ and $\frac{\partial KCon}{\partial Ap}$ are non-zero (see below)

$$simplify(diff(ACon, Ap)) = \frac{K\left(Ap^2 c_2^2 k_1 + \left(\left(2\,Ap + P1_{Total}\right) k_1 + k_2 P1_{Total}\right) c_2 + k_1\right) c_1 + c_2 P1_{Total} k_2}{\left(Ap\,c_2 + 1\right)^2 K k_1 c_1}$$

$$simplify(diff(KCon, Ap)) = \frac{c_2 P1_{Total} k_2}{\left(Ap\,c_2 + 1\right)^2 k_1}$$

However this simultaneously implies that $\frac{\partial Ap}{\partial K_{Total}}$ must be zero and non-zero to satisfy the two remaining expressions from earlier.
We have a contradiction.

Thus, our assumption of the presence of an enzyme biphasic in Bp with $K_{Total}$ is wrong.

**Hence enzyme biphasic cannot exist in the model.**

# Biphasic interactions within common network motifs

The following document includes models N1-N4. As mentioned in the main text each of these models involve a basic 3-node motif representing, a. positive feedback (N1), b. negative feedback (N2), c. incoherent feedforward (N3), and d. coherent feedforward (N4). Each node consists of an active form and an inactive form shuttling between each other. In each of these models, the basic characteristic of the motif is represented following which the role of biphasic interactions in the pathway are analyzed by considering the interaction of the pathway node (responsible for the feedback/feedforward) to be biphasic. All models are constructed using simple mass action kinetics and the biphasic interaction is modelled using the following expression $f(x) = b_1 \cdot x e^{-b_2 \cdot x}$ where a and b are parameters. This ensures a biphasic response is acheieved at $x$ value of $\frac{1}{b_2}$ and a corresponding maximum $f(x)$ of $\left( \frac{b_1}{b_2 \cdot e} \right)$.

## Proof of absence of multistability in negative feedback motif (Open System)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

In this sub-section we show that the basic negative feedback motif (Open system) used in figure S3, is incapable of exhibiting multistability. To recap; figure S3 shows how the introduction of a simple biphasic response in the feedback interaction can allow the system to present multistability.

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero. The model is described here again (the code for the same can be found in the MatCont package under the name N5__Bi_NFB). In this instance, for the open system, Switch = 0.

$$dR := k_0 \cdot S - k_1 \cdot R \cdot Ap \cdot (1 - Switch) - Switch \cdot R \cdot Ap \cdot b_1 \cdot \exp\left(-b_2 \cdot Ap\right) - k_2 \cdot R :$$

$$dAp := \frac{k_3 \cdot A \cdot R}{K_3 + A} - \frac{k_4 \cdot Ap}{K_4 + Ap} :$$

$$dA := -\frac{k_3 \cdot A \cdot R}{K_3 + A} + \frac{k_4 \cdot Ap}{K_4 + Ap} :$$

$$Switch := 0 :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A_{Total} - A - Ap :$

Now we begin solving the system of equations to obtain correlations between the steady state concentrations of the variables. Primarily to obtain steady state correlations of variables as a function of concentrations of A and R. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$R := solve(dA, R) = \frac{k_4\, Ap\, (K_3 + A)}{(K_4 + Ap)\, k_3\, A}$$

$assign(solve(\{ACon\}, \{A\})) :$

Now having solved for the steady state of the system in terms of Ap, the only two equation that remains (which defines the steady state of the system) is the differential equation dR

$simplify(numer(-dR))$

$$k_1 k_4 Ap^3 + \left(\left(-K_3 k_1 - k_1 A_{Total} + k_2\right) k_4 - k_0 S k_3\right) Ap^2 + \left(-k_2 \left(A_{Total} + K_3\right) k_4 + S k_0 k_3 \left(A_{Total} - K_4\right)\right) Ap + K_4 S k_0 k_3 A_{Total} \quad \textbf{(1.1)}$$

The feasible solutions of the above equation for Ap defines the steady state of the system. We can also see that the equation is a third degree polynomial in Ap, which can atmost accept three feasible solutions. Note that the feasible solution for Ap should lie between 0 and $A_{Total}$.

We can also observe that the coeffecient of the third exponent and the constant term are both positive. This implies that one root of the cubic polynomial is necessarily negative. Similarly since the polynomial evaluated at Ap = $A_{Total}$ is negative (see below), there exists exactly one solution.

$simplify\left(eval\left(T, Ap = A_{Total}\right)\right)$

$$-K_3 k_4 A_{Total} \left(k_1 A_{Total} + k_2\right) \quad \textbf{(1.2)}$$

## Choice of parameters for biphasic in interaction

In this sub-section we show our the rational used to choose parameter values when comparing the behavior of a network motif with and without biphasic in interaction. The main aim, as stated in the models and methods is to maintain parity in the strength of the interaction modelled, whether using a bilinear response (modelled using the term: $f(x) = k \cdot x$) indicating general motif response, or using a biphasic response modelled using the following term $f(x) = b_1 \cdot x \cdot \exp\left(-b_2 \cdot x\right)$

$restart : with(LinearAlgebra) : with(Student[LinearAlgebra]) : with(VectorCalculus) :$

In order to ensure parity in the strength of the interaction. We first assume that the system is operating over the range (for values of x) 0 to a finite value $X_{Total}$. In such a case, we wish to choose parameters $b_1$ and $b_2$ for the biphasic interaction term, such that the averaged strength over 0 to $X_{Total}$ is equal to that arising out of the bilinear (basal models') interaction. This can be written as shown below

$$\int_0^{X_{Total}} f_{Biphasic}(x)\,dx = \int_0^{X_{Total}} f_{basal}(x)\,dx$$

Substituiting the respective terms, this becomes

$$\int_0^{X_{Total}} b_1 \cdot x \cdot \exp(-b_2 \cdot x)\,dx = \int_0^{X_{Total}} k \cdot x\,dx :$$

Evaluating this expression we get

$$-\frac{b_1\left(e^{-b_2 X_{Total}} b_2 X_{Total} + e^{-b_2 X_{Total}} - 1\right)}{b_2^{\,2}} = \frac{k X_{Total}^{\,2}}{2}$$

Now we assign b2 the value of $b_2 = \frac{2^{0.5}}{X_{Total}}$ . With this, the above expression becomes

$$0.2065321413\, b_1 X_{Total}^{\,2} = \frac{k X_{Total}^{\,2}}{2}$$

Thus, further simplification of this implies that b1 should take the value as shown below

$$solve\left(0.2065321413\, b_1 X_{Total}^{\,2} = \frac{k X_{Total}^{\,2}}{2}, b_1\right)$$

$$2.420930693\, k \qquad \textbf{(2.1)}$$

Thus, in order to have parity in the averaged strength of both the biphasic interaction and the basal interaction over the range of operation (0 to $X_{Total}$). We approximate this and take the value of $b_1$ and $b_2$ to be as shown below.

$$\boldsymbol{b_1 := 2.5 \cdot k} \text{ and } \boldsymbol{b_2 := \frac{2^{0.5}}{X_{Total}}}$$

**Note:** In instances such as the upstream signal regulation not having an explicit total amount (unlike when the total amount, i.e. the range of activation, is evident when the regulation comes from a conserved substrate) the XTotal value is assumed to be taken arbitrarily such that it represents approximately half the range over which the bifurcation is carried out.

# Biphasic interactions within integral feedback control motif

In this section we provide the following proof pertaining to the effect of biphasic interactions on integral feedback control motif responose.
1. The presence of a biphasic interaction can induce a saturation effect that can result in a reduced range of signal regulation for the model.
2. The biphasic interaction allows for in addition to the saturation effect above a loss of stability in one of the branches of steady state.

We first begin by describing the three relevant models,
1. Basic model of integral feedback control motif
2. The integral feedaback control model with biphasic signal regulation
3. Model with biphasic response within the motif interactions

The first two models are provided for completeness - however please only run the third model as appropriate for the proof that follows. Similar to the other worksheets, please run the model, and one of the proofs (not the whole worksheet together)

## Basic model

$$restart : with(Student[LinearAlgebra]) : with(VectorCalculus) :$$

$$dA := kaa * S - kda * A * M :$$
$$dP := kap * A - kdp :$$
$$dM := kam * P - kdm * M :$$

## Biphasic signal regulation model

$$restart : with(Student[LinearAlgebra]) : with(VectorCalculus) :$$

$$dA := b1 * S * \exp(-b2 * S) - kda * A * M :$$
$$dP := kap * A - kdp :$$
$$dM := kam * P - kdm * M :$$

## Biphasic response within motif interactions

```
restart : with(Student[LinearAlgebra]) : with(VectorCalculus) :

dA := kaa * S − kda * A * M :
dP := kap * A − kdp :
dM := b1 * P * exp(−b2 * P) − kdm * M :
```

In the subsequent subsections we provide proofs for the following insights.

1. The presence of a biphasic interaction can induce a saturation effect that can result in a reduced range of signal regulation for the model.
2. The biphasic interaction allows for in addition to the saturation effect above a loss of stability in one of the branches of steady state.

## 1. Saturating effect of interaction leads to a reduced range of signal regulation

**Prereq: please run only model 3 before running this section.**

In order to illustrate this proof, we solve for the system at steady state to obtain relationships between the steady state concentrations of A, and P as shown below. We do this by solving the differential equations for A and P at steady state.

$assign(solve(\{dA, dP\}, \{A, M\}))$

Once we do this and resubstitute, the remaining differential equation for M simplies as follows

$simplify(dM)$

$$\frac{b1\ P\ e^{-b2\ P}\ kda\ kdp - kdm\ kaa\ S\ kap}{kda\ kdp} \tag{4.1}$$

Solving this equation for the signal value (at steady state) provides the following relationship between any upstream signal and the steady state concentration of P.

$solve(dM, S)$

$$\frac{b1\ P\ e^{-b2\ P}\ kda\ kdp}{kdm\ kaa\ kap} \tag{4.2}$$

Now, the expression on the right hand side is the biphasic interaction term (in variable P), multiplied by a constant (combination of the kinetic constants). This maximum can be evaluated as a function of the kinetic parameters as shown below,

$$Max\_value\_of\_S := simplify\left(eval\left(\frac{b1\ P\ e^{-b2\ P}\ kda\ kdp}{kdm\ kaa\ kap}, P = solve\left(diff\left(\frac{b1\ P\ e^{-b2\ P}\ kda\ kdp}{kdm\ kaa\ kap}, P\right), P\right)\right)\right)$$

$$Max\_value\_of\_S := \frac{b1\ e^{-1}\ kda\ kdp}{b2\ kdm\ kaa\ kap} \tag{4.3}$$

Thus, the right hand side has a maximum concentration that it can take as P varies. This implies that any signal value beyond that maximum (as given by the equation) cannot be supported as a steady state by the system.

**Hence we have shown how the saturation effect of interaction leads to a reduced range of signal regulation.** This can also be easily verified with a computation by choosing an input signal beyond the maximum value predicted by the equation above.

## 2. The biphasic interaction induces a loss of stability in one of the branches of steady state (necessarily)

**Prereq: please run only model 3 before running this section.**

In order to show this result pertaining to the stability of the steady states, we first make the point that there exists two steady states for any given choice of signal activation from this motif. This is apparent from the earlier steady states observed in proof 1. These steady state concentration relationships are rewritten below

$$A := \frac{kdp}{kap} : M := \frac{kaa\ S\ kap}{kda\ kdp} :$$

and $$S := \frac{b1\ P\ e^{-b2\ P}\ kda\ kdp}{kdm\ kaa\ kap} :$$

Thus for any given value of S, A is fixed, M takes a specific value, and there are two values of P that will satisfy the last equation. This implies that there are two steady state branches (Defined by the steady states of A, M and P) for every signal value.

Having estabished that - we will show below that one of these branches necessarily is unstable, and in particular looses stabilty at the point of inflection (where the biphasic nature of the interaction begins or kicks in).

In order to ascertain information regarding the stability of the steady states, we calculate the Jacobian and the Characteristic polynomial of the system at any given steady state. This is done as shown below.

$$J := Jacobian([dA, dP, dM], [A, P, M]) :$$

$$C := CharacteristicPolynomial(J, x) :$$

$$T := collect(eval(collect(C, x)), x)$$

$$T := x^3 + (M\ kda + kdm)\ x^2 + M\ kda\ kdm\ x - (P\ b2 - 1)\ A\ b1\ kap\ kda\ e^{-P\ b2} \quad \textbf{(5.1)}$$

From the above expression, for any feasible steady state we can observe that the coeffecient of the leading, second degree and first degree exponent of x are all positive. The roots of the polynomial (solutions) represents the eigen values of the system.

We solve the system to obtain steady state concentrations for A and M (as obtained in proof 1) and we resubstitute it into the expression T.

$assign(solve(\{dA, dM\}, \{M, A\}))$

We now extract the coeffecient of the exponents as shown below,

$C3 := coeff(T, x, 3) : C2 := coeff(T, x, 2) : C1 := coeff(T, x, 1) : C0 := coeff(T, x, 0) :$

Now at a given steady state, for the system to be stable, all eigen values of the steady state should be negative and real. In this polynomial, this will get determined depending on the sign of the constant term.

However we can see that this term can be negative (giving rise to a unstable steady state).

$C0$

$$-\frac{(P\,b2-1)\,S\,kaa\,kdm\,kap}{P} \tag{5.2}$$

This constant term is negative when $P\,b2-1$ is positive for a given steady state.

Examining the biphasic in interaction term as shown in the model, we can see that it is precisely negative when the biphasic effect kick's in within the interaction, or rather there is a decreasing function value for increasing input value of P. This can be seen below by checking the sign of the gradient of the function as P changes.

$simplify(diff(b1 * P * \exp(-b2 * P), P))$

$$-b1\,e^{-P\,b2}\,(P\,b2-1) \tag{5.3}$$

As we can see from this gradient, when $P\,b2-1$ is positive, the gradient is negative. The system originally as we showed initially has two steady state solutions, one corresponding to the initial phase of the interaction, and the other belonging to the waning / latter phase of the interaction.

**Hence proved Thow the biphasic interaction induces a loss of stability in one of the branches of steady state (necessarily).**

These two insights show how the waning phase of the biphasic interaction necessarily is unstable. Thus even though the biphasic behavior introduces an additional steady state and non-linearity in to the system.. it only diminishes the range over which signal regualtion is capable and further, the steady state introduced is unstable.

# Extracellular Regulated Kinase model (ERK)

## Exploring the features and requirements of obtaining biphasic dose response in pYpTErk with total substrate and enzyme amounts

In this file we analytically study the presence of enzyme and substrate biphasic dose response in the doubly phosphorylated Erk (pYpTErk). In particular we establish the following key results regarding them,

**1. Enzyme biphasic response in pYpTErk** (biphasic behavior in the dose response curve of pYpTErk as $Mek_{Total}$ ($K_{Total}$) changes): Presence of enzyme biphasic response is shown for a choice of parameters. Further, we show how certain analytical expressions (involving kinetic constants and total phosphatase amount) guarentees the presence of enzyme biphasic for some total concentration of Erk and Mek.

**2. Substrate biphasic response** (biphasic behavrior in the dose response curve of pYpTErk as $Erk_{Total}$ changes): We show how for any choice of biochemistry kinetics, the Erk model is guarenteed to present with substrate biphasic response in pYpTErk (at some total concentration of Erk and Mek).

We note that the key signature of biphasic behavior in the dose response curve of the system is the presence of a steady state of the system, where the following conditions are satisfied.

$$\frac{dpYpTErk}{dMek_{Total}} = 0 \text{ (for enzyme biphasic in pYpTErk)}$$

$$\frac{dpYpTErk}{dErk_{Total}} = 0 \text{ (for substrate biphasic in pYpTErk)}$$

Please note that we'll be using $K_{Total}$ interchangably for $Mek_{Total}$ (and K for Mek) for the remainder of this script for brevity.

**The Erk model:** We first describe the mathematical description of the Erk model (adapted from Bluthgen et al., 2018)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text. Here dERk represents d[Erk]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dErk := -k_{b1}\cdot Erk\cdot K + k_{ub1}\cdot C1 + p_4\cdot DT1 + k_4\cdot DY1 :$$
$$dC1 := k_{b1}\cdot Erk\cdot K - (k_{ub1} + k_1)\cdot C1 :$$
$$dpYErk := k_1\cdot C1 - k_{b4}\cdot pYErk\cdot P + k_{ub4}\cdot DY1 - k_{b2}\cdot pYErk\cdot K + k_{ub2}\cdot CY2 :$$
$$dDY1 := k_{b4}\cdot pYErk\cdot P - (k_{ub4} + k_4)\cdot DY1 :$$
$$dCY2 := k_{b2}\cdot pYErk\cdot K - (k_{ub2} + k_2)\cdot CY2 :$$
$$dpYpTErk := k_2\cdot CY2 - p_{b3}\cdot pYpTErk\cdot P + p_{ub3}\cdot D2 + p_2\cdot CT2 :$$
$$dD2 := p_{b3}\cdot pYpTErk\cdot P - (p_{ub3} + p_3)\cdot D2 :$$
$$dCT2 := p_{b2}\cdot pTErk\cdot K - (p_{ub2} + p_2)\cdot CT2 :$$
$$dpTErk := -p_{b2}\cdot pTErk\cdot K + p_{ub2}\cdot CT2 + p_3\cdot D2 - p_{b4}\cdot pTErk\cdot P + p_{ub4}\cdot DT1 :$$
$$dDT1 := p_{b4}\cdot pTErk\cdot P - (p_{ub4} + p_4)\cdot DT1 :$$

$$dK := -k_{b1}\cdot Erk\cdot K + (k_{ub1} + k_1)\cdot C1 - k_{b2}\cdot pYErk\cdot K + (k_{ub2} + k_2)\cdot CY2 - p_{b2}\cdot pTErk\cdot K + (p_{ub2} + p_2)\cdot CT2 :$$
$$dP := -k_{b4}\cdot pYErk\cdot P + (k_{ub4} + k_4)\cdot DY1 - p_{b3}\cdot pYpTErk\cdot P + (p_{ub3} + p_3)\cdot D2 - p_{b4}\cdot pTErk\cdot P + (p_{ub4} + p_4)\cdot DT1 :$$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, P1Con, P2Con and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$KCon := K_{Total} - K - C1 - CY2 - CT2 :$$
$$PCon := P_{Total} - P - DY1 - D2 - DT1 :$$
$$ErkCon := Erk_{Total} - Erk - pYErk - pTErk - pYpTErk - C1 - CY2 - CT2 - DY1 - D2 - DT1 :$$

We now solve the system described at steady state to obtain expression linking the steady state concentrations of the various species. Here we use the Maple command *solve*

$$assign(solve(\{dC1, dCY2, dCT2, dDY1, dD2, dDT1\}, \{C1, CY2, CT2, DY1, D2, DT1\}))$$

Simultaneously we introduce the following parameters ($c_1$, $c_2$, $c_4$, $d_2$, $d_3$, and $d_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$k_{b1} := c_1\cdot(k_{ub1} + k_1) : k_{b2} := c_2\cdot(k_{ub2} + k_2) : k_{b4} := c_4\cdot(k_{ub4} + k_4) :$$
$$p_{b2} := d_2\cdot(p_{ub2} + p_2) : p_{b3} := d_3\cdot(p_{ub3} + p_3) : p_{b4} := d_4\cdot(p_{ub4} + p_4) :$$

Once this is done, we again solve for the steady states of the unmodified and partially modified Erks (in terms of the other variables and parameters).

$$assign(solve(\{dErk, dpYErk, dpTErk\}, \{Erk, pYErk, pTErk\}))$$

Further we now introduce a ratio, $\epsilon$ = K/P (defined as the ratio of the free enzymes), and solve for the steady state of the free phosphatase (using the conservation expression PCon).

$K := \text{epsilon} \cdot P$ :

$P := simplify(solve(PCon, P)) =$

$$\left(P_{Total}\,\epsilon\,c_2\,k_2\left(d_2\,\epsilon\,p_2+d_4\,p_4\right)\right)\Big/\left(c_2\,d_2\,k_2\,p_2\left(d_3\,pYpTERk+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)pYpTERk+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,pYpTERk\,p_3\,p_4\right)$$

Note: For the sake of brevity of the expressions further, we rename pYpTErk as App.

$pYpTERk := App$ :

This systematic solving and renaming results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration App (pYpTErk) and $\epsilon$.

$$simplify(Erk) = \frac{\left(\epsilon\,c_2\,k_2+c_4\,k_4\right)d_4\,p_4\,d_3\,App\,p_3}{k_2\,c_2\,\epsilon^2\,k_1\,c_1\left(d_2\,\epsilon\,p_2+d_4\,p_4\right)}$$

$$simplify(pYErk) = \frac{d_4\,p_4\,d_3\,App\,p_3}{\epsilon\,c_2\,k_2\left(d_2\,\epsilon\,p_2+d_4\,p_4\right)}$$

$$simplify(pTErk) = \frac{d_3\,App\,p_3}{d_2\,\epsilon\,p_2+d_4\,p_4}$$

$$simplify(C1) = \frac{P_{Total}\left(\epsilon\,c_2\,k_2+c_4\,k_4\right)d_4\,p_4\,d_3\,App\,p_3}{k_1\left(c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4\right)}$$

$$simplify(CY2) = \frac{c_2\,P_{Total}\,d_4\,p_4\,d_3\,App\,p_3\,\epsilon}{c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4}$$

$$simplify(CT2) = \frac{d_2\,P_{Total}\,\epsilon^2\,c_2\,k_2\,App\,d_3\,p_3}{c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4}$$

$$simplify(DY1) = \frac{c_4\,P_{Total}\,d_4\,p_4\,d_3\,App\,p_3}{c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4}$$

$$simplify(D2) = \frac{d_3\,App\,P_{Total}\,\epsilon\,c_2\,k_2\left(\epsilon\,d_2\,p_2+d_4\,p_4\right)}{c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4}$$

$$simplify(DY1) = \frac{c_4\,P_{Total}\,d_4\,p_4\,d_3\,App\,p_3}{c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+\left(d_3\left(p_3+p_4\right)App+p_4\right)k_2\,d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4}$$

Note that when the steady state concentrations of App and $\epsilon$ are positive, the steady state concentrations of the other variable concentrations are positive as well. Thus, so far, we have solved the system of equations at steady state to arrive at expressions linking the steady state concentrations of the variables with that of App and $\epsilon$. We now have two expressions, ErkCon and KCon - the conservation of the Erk and Mek, whose

solution for the variables define the steady state of the full system.

## Substrate Biphasic

In this subsection, we analytically show the presence of substrate biphasic (in the full system) for any choice of underlying kinetics. As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system.

$$\frac{\partial App}{\partial Erk_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ErkCon = 0 (see below) whose solutions to the variables App and $\epsilon$ define the steady state of the system.

If we differentiate both these with respect to the total substrate concentration in the system, we get

$$\frac{dErkCon}{dErk_{Total}} = 0 = \frac{\partial ErkCon}{\partial Erk_{Total}} + \frac{\partial ErkCon}{\partial App} \cdot \frac{\partial App}{\partial Erk_{Total}} + \frac{\partial ErkCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial Erk_{Total}}$$
$$\frac{dKCon}{dErk_{Total}} = 0 = \frac{\partial KCon}{\partial App} \cdot \frac{\partial App}{\partial Erk_{Total}} + \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial Erk_{Total}}$$

Now in order to show the presence of a substrate biphasic response and study its features, we begin by first imposing the necessary features that the system must satisfy for the behavior to exist. We begin with the basic tenet that for the behavior there should exist a steady state of the system where $\frac{\partial App}{\partial Erk_{Total}} = 0$ is satisfied.
At this point then, the above expressions simply as follows

$$1 = \frac{\partial ErkCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial Erk_{Total}}$$
$$0 = \frac{\partial KCon}{\partial \epsilon} \cdot \frac{\partial \epsilon}{\partial Erk_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ErkCon}{\partial Erk}, \frac{\partial KCon}{\partial Erk}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial Erk_{Total}}$ can be zero.

$$simplify(diff(ErkCon, App)) =$$
$$\left(-d_2^3 \left(App^2 p_2 d_3^2 + \left(p_3 P_{Total} + 2\left(App + \frac{P_{Total}}{2}\right) p_2\right) d_3 + p_2\right) k_1 c_2^3 p_2^2 k_2^3 c_1 \epsilon^7 \right.$$
$$- d_2^2 c_2^3 \left(\left(App^2 p_2 p_3 d_3^3 k_1 + 2 k_1 App \left(App\left(p_3 + \frac{3 p_4}{2}\right) d_4 + p_3\right) p_2 d_3^2\right.\right.$$
$$\left. + \left(\left(\left(P_{Total}\left(p_2 + 2 k_1\right) p_3 + 6 k_1\left(App + \frac{P_{Total}}{2}\right) p_2\right) p_4 + 2 k_1\left(App + \frac{P_{Total}}{2}\right) p_3 p_2\right) d_4\right.\right.$$

$$
\begin{aligned}
&+ p_2 p_3 k_1 \Big) d_3 + 3\, p_2 p_4 d_4 k_1 \Big) k_2 + p_2 p_3 p_4 P_{Total} d_3 d_4 k_1 \Big) p_2 k_2^2 c_1 \epsilon^6 \\
&- 2\, d_2 c_2^2 d_4 \Bigg( \Bigg( \Bigg( \frac{\left( \left( 2 k_1 c_1 + p_2 d_2 \right) p_4 + 2 p_3 c_1 k_1 \right) App^2 p_3 p_2 d_3^3}{2} \\
&+ \frac{1}{2} \Big( App\, p_2 \left( App\, c_1 k_1 \left( p_3 + 3 p_4 \right) \left( p_3 + p_4 \right) d_4 + 2 \left( \left( 2 k_1 c_1 + p_2 d_2 \right) p_4 \right. \right. \right. \\
&\left. \left. \left. + p_3 c_1 k_1 \right) p_3 \right) d_3^2 \Big) + 2\, p_4 \Bigg( \Bigg( \Bigg( \frac{P_{Total} \left( p_2 + \frac{k_1}{2} \right) p_3}{2} + \frac{3 k_1 \left( App + \frac{P_{Total}}{2} \right) p_2}{2} \Bigg) p_4 \\
&+ k_1 \left( App + \frac{P_{Total}}{2} \right) p_3 p_2 \Bigg) c_1 d_4 + \frac{p_2 p_3 \left( 2 k_1 c_1 + p_2 d_2 \right)}{4} \Bigg) d_3 + \frac{3 p_2 p_4^2 c_1 d_4 k_1}{2} \Bigg) k_2 \\
&+ p_2 p_3 p_4^2 P_{Total} c_1 d_3 d_4 k_1 \Bigg) c_2 + d_2 p_4 \Bigg( \frac{App^2 d_3^2 k_1}{2} + App\, k_1 \left( App\, c_4 + 1 \right) d_3 + App\, c_4 k_1 \\
&+ \left( \frac{1}{2} + \frac{P_{Total} c_4}{2} \right) k_1 + \frac{k_4 P_{Total} c_4}{2} \Bigg) p_3 d_3 p_2^2 c_1 \Bigg) k_2^2 \epsilon^5 - c_2^2 d_4 \Bigg( 2 \Bigg( \Bigg( \Bigg( \Bigg( p_2 d_2 \\
&+ \frac{k_1 c_1}{2} \Bigg) p_4 + \frac{p_3 c_1 k_1}{2} \Bigg) App^2 p_3 \left( p_3 + p_4 \right) d_3^3 \\
&+ \frac{p_4 App \left( App\, c_1 k_1 \left( p_3 + p_4 \right)^2 d_4 + 2 \left( \left( k_1 c_1 + 2 p_2 d_2 \right) p_4 + p_3 \left( k_1 c_1 + p_2 d_2 \right) \right) p_3 \right) d_3^2}{2} \\
&+ p_4^2 \Bigg( \Bigg( \Bigg( \frac{p_3 P_{Total}}{2} + k_1 \left( App + \frac{P_{Total}}{2} \right) \Bigg) p_4 + p_3 k_1 \left( App + \frac{P_{Total}}{2} \right) \Bigg) c_1 d_4 + p_3 \Bigg( p_2 d_2 \\
&+ \frac{k_1 c_1}{2} \Bigg) \Bigg) d_3 + \frac{p_4^3 c_1 d_4 k_1}{2} \Bigg) k_2 + \frac{p_3 p_4^3 P_{Total} c_1 d_3 d_4 k_1}{2} \Bigg) d_4 c_2 + d_2 p_4 \Bigg( \Big( 2 k_1 c_1 \left( p_3 \right. \\
&\left. + p_4 \right) d_4 + c_4 \left( 2 p_3 c_1 k_1 + k_4 p_2 d_2 \right) \Big) App^2 d_3^2 + 2\, App \left( k_1 c_1 \left( p_3 + 2 p_4 \right) \left( App\, c_4 + 1 \right) d_4 \right. \\
&\left. + c_4 \left( p_3 c_1 k_1 + k_4 p_2 d_2 \right) \right) d_3 + 4 p_4 \Bigg( App\, c_4 k_1 + \left( \frac{1}{2} + \frac{P_{Total} c_4}{2} \right) k_1 + \frac{k_4 P_{Total} c_4}{2} \Bigg) c_1 d_4 \\
&+ k_4 p_2 c_4 d_2 \Bigg) p_3 d_3 p_2 \Bigg) k_2^2 \epsilon^4 - 2\, p_4 c_2 p_3 d_3 \Bigg( \frac{\left( d_3 \left( p_3 + p_4 \right) App + p_4 \right)^2 d_4 k_2^2 c_2^2}{2} \\
&+ \Bigg( \Bigg( \frac{k_1 c_1 \left( p_3 + p_4 \right)^2 d_4}{2} + \Big( \left( p_3 \left( k_1 c_1 + p_2 d_2 \right) + k_4 p_2 d_2 \right) p_4 + p_3 \left( p_3 c_1 k_1 \right. \\
&\left. + k_4 p_2 d_2 \right) \Big) c_4 \Bigg) App^2 d_3^2 + App \left( p_4 c_1 k_1 \left( p_3 + p_4 \right) \left( App\, c_4 + 1 \right) d_4 + \Big( \Big( p_3 \left( k_1 c_1 \right. \right. \\
&\left. \left. + p_2 d_2 \right) + 2 k_4 p_2 d_2 \Big) p_4 + k_4 p_2 p_3 d_2 \Big) c_4 \right) d_3 + p_4 \Bigg( p_4 \Bigg( App\, c_4 k_1 + \left( \frac{1}{2} + \frac{P_{Total} c_4}{2} \right) k_1 \\
&+ \frac{k_4 P_{Total} c_4}{2} \Bigg) c_1 d_4 + k_4 p_2 c_4 d_2 \Bigg) \Bigg) k_2 c_2 \\
&+ \frac{App\, p_2 p_3 p_4 c_1 c_4 d_2 d_3 k_1 \left( App\, c_4 + 2\, App\, d_3 + 2 \right)}{2} \Bigg) d_4^2 k_2 \epsilon^3
\end{aligned}
$$

$$-2\,p_4\,c_2\,p_3\,d_3\,d_4^{\,2}\,k_2\,c_4\left(\frac{1}{2}\left(\left(d_3\left(p_3+p_4\right)App+p_4\right)\left(\left(\left(k_4+2\,p_3\right)p_4+k_4\,p_3\right)App\,d_3+k_4\,p_4\right)d_4\,k_2\,c_2\right)+p_4\,App\,p_3\,d_3\left(\left(k_1\,c_1\left(p_3+p_4\right)d_4+c_4\left(k_4\,p_2\,d_2+\frac{p_3\,c_1\,k_1}{2}\right)\right)App\,d_3+\frac{p_4\,c_1\,k_1\left(App\,c_4+2\right)d_4}{2}+k_4\,p_2\,c_4\,d_2\right)\right)\epsilon^2-2\,p_4^{\,2}\left(\left(\left(\left(k_4+\frac{p_3}{2}\right)p_4+k_4\,p_3\right)App\,d_3+k_4\,p_4\right)k_2\,c_2+\frac{App\,p_3\,p_4\,c_1\,d_3\,k_1}{2}\right)App\,p_3^{\,2}\,d_3^{\,2}\,d_4^{\,3}\,c_4^{\,2}\,\epsilon-App^2\,c_4^{\,3}\,d_3^{\,3}\,d_4^{\,3}\,k_4\,p_3^{\,3}\,p_4^{\,3}\Bigg)\Bigg/\left(k_1\,c_2\left(c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+d_4\left(d_3\left(p_3+p_4\right)App+p_4\right)c_2\,k_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4\right)^2\epsilon^2\,k_2\,c_1\left(d_2\,\epsilon\,p_2+d_4\,p_4\right)\right)$$

*simplify*(*diff*(*KCon*, *App*)) =

$$\left(d_3\left(p_2\,\epsilon^3\,c_2\,d_2\,k_1\,k_2-\left(\left(-p_3-p_4\right)d_4+p_3\,d_2\right)k_2\,k_1\,c_2\,\epsilon^2-\left(\left(k_1+k_2\right)c_2-k_1\,c_4\right)p_4\,d_4\,p_3\,\epsilon-p_3\,p_4\,c_4\,d_4\,k_4\right)\epsilon\,k_2\,P_{Total}\left(d_2\,\epsilon\,p_2+d_4\,p_4\right)c_2\right)\Big/\left(k_1\left(c_2\,d_2\,k_2\,p_2\left(App\,d_3+1\right)\epsilon^2+k_2\left(\left(App\,d_3+1\right)p_4+App\,d_3\,p_3\right)d_4\,c_2\,\epsilon+c_4\,d_3\,d_4\,App\,p_3\,p_4\right)^2\right)$$

Observing, the second expression above (from differentiation of the total kinase concentration) we can see that either $\frac{\partial KCon}{\partial \epsilon}$ or $\frac{\partial \epsilon}{\partial Erk_{Total}}$ must be equal to zero. However, $\frac{\partial \epsilon}{\partial Erk_{Total}}$ cannot be zero as, if it was indeed zero, then since the denominator of $\frac{\partial ErkCon}{\partial \epsilon}$ is non-zero, there would be a contradiction with the first expression (from differentiation of the total Erk concentration).

Thus, $\frac{\partial KCon}{\partial \epsilon}$ must be equal to zero. This expression is evaluated below and stored as T.

$$T := simplify\left(\frac{numer(simplify((diff(KCon, \text{epsilon}))))}{P_{Total}\cdot c_2}\right)$$

$$T := \epsilon^2\,c_2\left(-\left(d_3\left(p_3+p_4\right)App+p_4\right)k_1\,p_4\,d_4^{\,2}+d_2\left(p_3\,App^2\left(\left(p_2-k_1\right)p_4-p_3\,k_1\right)d_3^{\,2}+App\left(\left(-2\,k_1\,p_2\,\epsilon+p_3\left(p_2-k_1\right)\right)p_4-2\,k_1\,p_2\,\epsilon\,p_3\right)d_3-2\,p_2\,p_4\,\epsilon\,k_1\right)d_4-p_2^{\,2}\,\epsilon^2\,d_2^{\,2}\,k_1\left(d_3\,App+1\right)\right)k_2^{\,2}+p_3\,App\,d_3\,p_4\left(-\left(App\left(\left(p_3-k_4\right)p_4-k_4\,p_3\right)d_3+2\left(\epsilon\,k_1-\frac{k_4}{2}\right)p_4\right)c_4\,d_4+\epsilon\,d_2\left(App\left(p_2\,\epsilon\,c_2\,k_1+2\,c_4\left(-p_3\,k_1+k_4\,p_2\right)\right)d_3+\left(k_1\left(c_2-3\,c_4\right)\epsilon+2\,c_4\,k_4\right)p_2\right)\right)d_4\,k_2-App^2\,c_4\,d_3^{\,2}\,d_4^{\,2}\,k_1\,p_3^{\,2}\,p_4^{\,2} \qquad \textbf{(1.1)}$$

Thus as long as the above equation (expression T) is satsifed, there exists a steady state at some $K_{Total}$ and $Erk_{Total}$ corresponding to the conservation equations, where there exists a biphasic peak (behavior) in the dose response.

Now in order to show that the expression T is indeed satisfied for every choice of underlying kinetics, we rewrite $App = m \cdot \epsilon$

$App := m \cdot \text{epsilon}$ :

This simplifies the expression T as shown below.

$simplify(T)$

$$\left( -m\, p_2^{\,2}\, \epsilon^4\, c_2\, d_2^{\,2}\, d_3\, k_1\, k_2^{\,2} + c_2 \left( \left( d_3 \left( \left( \left( p_2 - k_1 \right) p_4 - p_3\, k_1 \right) p_3\, m\, d_3 - 2\, p_2\, k_1 \left( p_3 + p_4 \right) \right) m\, d_4 - p_2^{\,2}\, d_2\, k_1 \right) k_2 + m^2\, p_2\, p_3\, p_4\, d_3^{\,2}\, d_4\, k_1 \right) d_2\, k_2\, \epsilon^3 + 2\, p_4 \left( \frac{c_2 \left( -m\, d_3\, k_1 \left( p_3 + p_4 \right) d_4 + \left( m\, p_3 \left( p_2 - k_1 \right) d_3 - 2\, p_2\, k_1 \right) d_2 \right) k_2}{2} + \left( m\, c_4 \left( -p_3\, k_1 + k_4\, p_2 \right) d_3 + \frac{p_2\, k_1 \left( c_2 - 3\, c_4 \right)}{2} \right) d_3\, p_3\, m\, d_2 \right) d_4\, k_2\, \epsilon^2 - \left( p_4\, c_2\, d_4\, k_1\, k_2^{\,2} + d_3\, c_4\, p_3\, m \left( \left( \left( \left( p_3 - k_4 \right) p_4 - p_3\, k_4 \right) m\, d_3 + 2\, p_4\, k_1 \right) d_4 - 2\, p_2\, d_2\, k_4 \right) k_2 + m^2\, p_3^{\,2}\, p_4\, c_4\, d_3^{\,2}\, d_4\, k_1 \right) p_4\, d_4\, \epsilon + m\, p_3\, p_4^{\,2}\, c_4\, d_3\, d_4^{\,2}\, k_2\, k_4 \right) \epsilon \qquad \textbf{(1.2)}$$

We can observe that the expression is now a fifth order polynomial in $\epsilon$. One of the roots factors out resulting in a simplified quartic polynomial in $\epsilon$, with the constant term always positive and the leading coeffecient always having a negative sign. This indicates that for any choice of value for m (which appears in all the coeffecients) there exists a postive solution in $\epsilon$ for this quartic expression.

This indicates that for every choice of underlying kinetics and total phosphatase concentration, expression T = 0 can be satisfied (guarenteeing the existence of biphasic response in the dose response for some total concentrations of ErkTotal and KTotal which can be arrived at by evaluating the respective conservation expressions).

**Thus we have shown how substrate biphasic response is guarenteed to exist in pYpTErk, for some total concentration of Erk and Kinase in the system, for any choice of underlying kinetics.**

Note: We illustrate the validity of this argument by providing an example of a prediction of the presence of a biphasic response for an arbitraty value of 1 for m (i.e. App = $\epsilon$) and some abitrary choice of kinetics (all values are equal to 1), validated by computational bifurcation analysis.

$m := 1$ :

$k_1 := 1 : k_2 := 1 : k_4 := 1 : p_2 := 1 : p_3 := 1 : p_4 := 1$ :

$k_{b1} := 1 : k_{b2} := 1 : k_{b4} := 1 : p_{b2} := 1 : p_{b3} := 1 : p_{b4} := 1 : k_{ub1} := 1 : k_{ub2} := 1 : k_{ub4} := 1 : p_{b2} := 1 : p_{b3} := 1 : p_{b4} := 1 : p_{ub2} := 1 : p_{ub3} := 1 : p_{ub4} := 1$ :

$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} : d_3 := \frac{p_{b3}}{p_3 + p_{ub3}} : d_2 := \frac{p_{b2}}{p_2 + p_{ub2}} : d_4 := \frac{p_{b4}}{p_4 + p_{ub4}}$ :

$P_{Total} := 1$ :

$\epsilon := evalf(solve(T, useassumptions) \text{ assuming } \epsilon > 0) = 0.236067977$

$App = 0.236067977$

$ErkCon = Erk_{Total} - 6.236067987$

$$KCon = K_{Total} - 0.5278640449$$

**The computational evidence of the substrate biphasic dose response behavior predicted above is shown in supplementary figure 1A.**

## Enzyme biphasic

In this subsection, we analytically show the presence of enzyme biphasic (in the full system) for some choice of kinetic parameter values. We further show the existence of an analytical condition involving kinetic constants that will guarantee the presense of biphasic dose response in pYpTErk with $Mek_{Total}$ ($K_{Total}$). As noted earlier, the biphasic behavior is characterized by the following condition being satisfied for some steady state of the system.

$$\frac{\partial App}{\partial K_{Total}} = 0$$

We now have two remaining conservations, KCon = 0 & ErkCon = 0 (see below) whose solutions to the variables App and $\epsilon$ define the steady state of the system.

If we differentiate these both with respect to the total enzyme (kinase/Mek) concentration in the system, we get

$$\frac{dErkCon}{dK_{Total}} = 0 = \frac{\partial ErkCon}{\partial App}\cdot\frac{\partial App}{\partial K_{Total}} + \frac{\partial ErkCon}{\partial \epsilon}\cdot\frac{\partial \epsilon}{\partial K_{Total}}$$
$$\frac{dKCon}{dK_{Total}} = 0 = \frac{\partial KCon}{\partial K_{Total}} + \frac{\partial KCon}{\partial App}\cdot\frac{\partial App}{\partial K_{Total}} + \frac{\partial KCon}{\partial \epsilon}\cdot\frac{\partial \epsilon}{\partial K_{Total}}$$

Now in order to show the presence of a substrate biphasic response and study its features, we begin with exploring the necessary features that the system must satisfy for the behavior to exist. We begin with the basic tenet that for the behavior there should exist a steady state of the system where $\frac{\partial App}{\partial K_{Total}} = 0$ is satisfied.
At this point then, the above expressions simply as follows

$$0 = \frac{\partial ErkCon}{\partial \epsilon}\cdot\frac{\partial \epsilon}{\partial K_{Total}}$$
$$1 = \frac{\partial KCon}{\partial \epsilon}\cdot\frac{\partial \epsilon}{\partial K_{Total}}$$

***Note***

This simplification is possible since the functions $\frac{\partial ErkCon}{\partial Erk}, \frac{\partial KCon}{\partial Erk}$ have a non-zero denominator (as shown below) and thus the products involving $\frac{\partial App}{\partial Erk_{Total}}$ can be zero.

$$simplify(diff(ErkCon, App)) =$$
$$\left(-\left(App^2 p_2 d_3^2 + \left(p_3 P_{Total} + 2\left(App + \frac{P_{Total}}{2}\right)p_2\right)d_3 + p_2\right)c_2^3 k_2^3 c_1 p_2^2 d_2^3 k_1 \epsilon^7\right.$$

$$
\begin{aligned}
&-c_2^{\,3}\,k_2^{\,2}\,c_1\,p_2\,d_2^{\,2}\left(\left(App^2\,p_2\,p_3\,d_3^{\,3}\,k_1+2\,App\,p_2\,k_1\left(App\left(p_3+\frac{3\,p_4}{2}\right)d_4+p_3\right)d_3^{\,2}\right.\right.\\
&+\left(\left(\left(P_{Total}\left(p_2+2\,k_1\right)p_3+6\left(App+\frac{P_{Total}}{2}\right)p_2\,k_1\right)p_4+2\left(App+\frac{P_{Total}}{2}\right)p_3\,p_2\,k_1\right)d_4\right.\\
&\left.\left.+p_2\,p_3\,k_1\right)d_3+3\,p_2\,p_4\,d_4\,k_1\right)k_2+p_2\,p_3\,p_4\,P_{Total}\,d_3\,d_4\,k_1\Bigg)\epsilon^6\\
&-2\left(\left(\left(\frac{p_3\,App^2\left(\left(2\,k_1\,c_1+p_2\,d_2\right)p_4+2\,p_3\,c_1\,k_1\right)p_2\,d_3^{\,3}}{2}\right.\right.\right.\\
&+\frac{1}{2}\left(App\left(App\,c_1\,k_1\left(p_3+3\,p_4\right)\left(p_3+p_4\right)d_4+2\,p_3\left(\left(2\,k_1\,c_1+p_2\,d_2\right)p_4\right.\right.\right.\right.\\
&\left.\left.\left.\left.+p_3\,c_1\,k_1\right)\right)p_2\,d_3^{\,2}\right)+2\,p_4\left(\left(\left(\left(\frac{\left(p_2+\frac{k_1}{2}\right)P_{Total}\,p_3}{2}+\frac{3\left(App+\frac{P_{Total}}{2}\right)p_2\,k_1}{2}\right)p_4\right.\right.\right.\\
&\left.\left.\left.+\left(App+\frac{P_{Total}}{2}\right)p_3\,p_2\,k_1\right)c_1\,d_4+\frac{p_2\,p_3\left(2\,k_1\,c_1+p_2\,d_2\right)}{4}\right)d_3+\frac{3\,p_2\,p_4^{\,2}\,c_1\,d_4\,k_1}{2}\right)k_2\\
&\left.+p_2\,p_3\,p_4^{\,2}\,P_{Total}\,c_1\,d_3\,d_4\,k_1\right)c_2+p_3\left(\frac{App^2\,d_3^{\,2}\,k_1}{2}+App\,k_1\left(App\,c_4+1\right)d_3+App\,c_4\,k_1\right.\\
&\left.+\left(\frac{1}{2}+\frac{P_{Total}\,c_4}{2}\right)k_1+\frac{k_4\,P_{Total}\,c_4}{2}\right)d_3\,c_1\,p_2^{\,2}\,d_2\,p_4\Bigg)c_2^{\,2}\,d_4\,k_2^{\,2}\,d_2\,\epsilon^5-c_2^{\,2}\,d_4\left(2\,d_4\left(\left(\left(p_3\right.\right.\right.\\
&\left.\left.+p_4\right)p_3\,App^2\left(\left(p_2\,d_2+\frac{k_1\,c_1}{2}\right)p_4+\frac{p_3\,c_1\,k_1}{2}\right)d_3^{\,3}\right.\\
&\left.+\frac{App\left(App\,c_1\,k_1\left(p_3+p_4\right)^2\,d_4+2\,p_3\left(\left(k_1\,c_1+2\,p_2\,d_2\right)p_4+p_3\left(k_1\,c_1+p_2\,d_2\right)\right)\right)p_4\,d_3^{\,2}}{2}\right.\\
&+\left(c_1\left(\left(\frac{p_3\,P_{Total}}{2}+k_1\left(App+\frac{P_{Total}}{2}\right)\right)p_4+p_3\,k_1\left(App+\frac{P_{Total}}{2}\right)\right)d_4+p_3\left(p_2\,d_2\right.\right.\\
&\left.\left.\left.+\frac{k_1\,c_1}{2}\right)\right)p_4^{\,2}\,d_3+\frac{p_4^{\,3}\,c_1\,d_4\,k_1}{2}\right)k_2+\frac{p_3\,p_4^{\,3}\,P_{Total}\,c_1\,d_3\,d_4\,k_1}{2}\Bigg)c_2+p_3\left(App^2\left(2\,k_1\,c_1\left(p_3\right.\right.\right.\\
&\left.\left.\left.+p_4\right)d_4+c_4\left(2\,p_3\,c_1\,k_1+k_4\,p_2\,d_2\right)\right)d_3^{\,2}+2\left(k_1\,c_1\left(p_3+2\,p_4\right)\left(App\,c_4+1\right)d_4\right.\right.\\
&\left.\left.+c_4\left(p_3\,c_1\,k_1+k_4\,p_2\,d_2\right)\right)App\,d_3+4\left(App\,c_4\,k_1+\left(\frac{1}{2}+\frac{P_{Total}\,c_4}{2}\right)k_1\right.\right.\\
&\left.\left.\left.+\frac{k_4\,P_{Total}\,c_4}{2}\right)c_1\,p_4\,d_4+k_4\,p_2\,c_4\,d_2\right)d_3\,p_2\,d_2\,p_4\right)k_2^{\,2}\,\epsilon^4\\
&-2\,p_3\,c_2\,d_4^{\,2}\,d_3\left(\frac{\left(d_3\left(p_3+p_4\right)App+p_4\right)^2\,d_4\,k_2^{\,2}\,c_2^{\,2}}{2}+\left(App^2\left(\frac{k_1\,c_1\left(p_3+p_4\right)^2\,d_4}{2}\right.\right.\right.\\
&\left.\left.\left.+c_4\left(\left(p_3\left(k_1\,c_1+p_2\,d_2\right)+k_4\,p_2\,d_2\right)p_4+p_3\left(p_3\,c_1\,k_1+k_4\,p_2\,d_2\right)\right)\right)d_3^{\,2}\right.\right.
\end{aligned}
$$

$$+ App\left(p_4 c_1 k_1 \left(p_3 + p_4\right)\left(App\, c_4 + 1\right) d_4 + c_4\left(\left(p_3\left(k_1 c_1 + p_2 d_2\right) + 2 k_4 p_2 d_2\right) p_4 + k_4 p_2 p_3 d_2\right)\right) d_3 + \left(\left(App\, c_4 k_1 + \left(\frac{1}{2} + \frac{P_{Total}\, c_4}{2}\right) k_1 + \frac{k_4 P_{Total}\, c_4}{2}\right) c_1 p_4 d_4 + k_4 p_2 c_4 d_2\right) p_4\right) k_2 c_2 + \frac{App\, p_2 p_3 p_4 c_1 c_4 d_2 d_3 k_1 \left(App\, c_4 + 2\, App\, d_3 + 2\right)}{2}\Bigg) k_2 p_4 \epsilon^3 - 2\left(\frac{\left(d_3\left(p_3 + p_4\right) App + p_4\right) d_4 k_2 \left(App\left(\left(k_4 + 2 p_3\right) p_4 + k_4 p_3\right) d_3 + k_4 p_4\right) c_2}{2} + p_3 App\, d_3 p_4\left(\left(k_1 c_1\left(p_3 + p_4\right) d_4 + c_4\left(k_4 p_2 d_2 + \frac{p_3 c_1 k_1}{2}\right)\right) App\, d_3 + \frac{p_4 c_1 k_1 \left(App\, c_4 + 2\right) d_4}{2} + k_4 p_2 c_4 d_2\right)\right) p_3 c_2 d_4^2 d_3 c_4 k_2 p_4 \epsilon^2 - 2 p_3^2 App\, d_4^3 d_3^2 c_4^2 p_4^2\left(\left(App\left(\left(k_4 + \frac{p_3}{2}\right) p_4 + k_4 p_3\right) d_3 + k_4 p_4\right) k_2 c_2 + \frac{App\, p_3 p_4 c_1 d_3 k_1}{2}\right) \epsilon - App^2 c_4^3 d_3^3 d_4^3 k_4 p_3^3 p_4^3\Bigg) \Big/ \left(\epsilon^2 c_2 k_2 c_1 k_1\left(d_2 \epsilon p_2 + d_4 p_4\right)\left(c_2 d_2 k_2 p_2\left(App\, d_3 + 1\right)\epsilon^2 + \left(d_3\left(p_3 + p_4\right) App + p_4\right) c_2 k_2 d_4 \epsilon + c_4 d_3 d_4 App\, p_3 p_4\right)^2\right)$$

$$simplify(diff(KCon, App)) =$$

$$\left(P_{Total}\left(p_2 \epsilon^3 c_2 d_2 k_1 k_2 - \left(\left(-p_3 - p_4\right) d_4 + p_3 d_2\right) k_1 c_2 k_2 \epsilon^2 - p_4 d_4\left(\left(k_1 + k_2\right) c_2 - k_1 c_4\right) p_3 \epsilon - p_3 p_4 c_4 d_4 k_4\right)\left(d_2 \epsilon p_2 + d_4 p_4\right) c_2 d_3 k_2 \epsilon\right) \Big/ \left(\left(c_2 d_2 k_2 p_2\left(App\, d_3 + 1\right)\epsilon^2 + \left(\left(App\, d_3 + 1\right) p_4 + App\, d_3 p_3\right) d_4 c_2 k_2 \epsilon + c_4 d_3 d_4 App\, p_3 p_4\right)^2 k_1\right)$$

Observing, the first expression above (from differentiation of the total Erk concentration) we can see that either $\frac{\partial ErkCon}{\partial \epsilon}$ or $\frac{\partial \epsilon}{\partial K_{Total}}$ must be equal to zero. However, $\frac{\partial \epsilon}{\partial K_{Total}}$ cannot be zero as, if it was indeed zero, then since the denominator of $\frac{\partial KCon}{\partial \epsilon}$ is non-zero, there would be a contradiction with the first expression (from differentiation of the total Erk concentration).

Thus, $\frac{\partial ErKCon}{\partial \epsilon}$ must be equal to zero. This expression (a polynomial in App and $\epsilon$) is calculated as shown below and stored as T.

$T := simplify(numer(simplify(diff(ErkCon, \text{epsilon}))))$ :

Now rewriting App as m/$\epsilon$ allows us to simplify T as a ninth degree polynomial in epsilon (whose coeffecients are functions of the variable m).

$App := \frac{m}{\text{epsilon}}$ :

However, this rearrangement allows us to isolate and observe the leading coeffecient and constant term of the polynomial as shown below.

$T := simplify(T \cdot \epsilon^3)$ :

Error, recursive assignment

$C0 := simplify(coeff(T, \text{epsilon}, 0))$

$$= 2\,d_3^{\,3}\,p_3^{\,3}\,App^3\,d_4^{\,4}\,p_4^{\,4}\,c_4^{\,3}\,k_4$$

$C9 := simplify(coeff(T, \text{epsilon}, 0))$

$$C9 := 2\,d_3^{\,3}\,p_3^{\,3}\,App^3\,d_4^{\,4}\,p_4^{\,4}\,c_4^{\,3}\,k_4 \qquad \textbf{(2.1)}$$

We can notice that the sign of C0 is always positive for all choices of kinetic parameters and feasible values of m, while the sign of C9 is decided purely by $P_{Total}$ and the kinetic constants (and not m). We can use this insight to conclude that, should C9 be negative in sign, then there exists necessarily a positive feasible solution of $\epsilon$ for every feasible value m, thereby guarenteeing the existence of some feasible App and $\epsilon$ where the expression is 0 (indicating the presence of biphasic behavior in the dose response of pYpTErk with $K_{Total}$).

i.e. If C9<0 then there is the guarenteed existence of biphasic dose response with kinase for some total concentrations. We further isolate this condition from C9 as shown below (stored as the expression *Condition*),

$$Condition := collect\left(-\left(\left(\left(\left(-d_4\left(p_4+k_1\right)P_{Total}-k_1\right)p_2+p_4\,P_{Total}\,d_4\,k_1\right)k_2-d_4\,k_1\,p_2\,p_4\,P_{Total}\right)\right) < 0, P_{Total}\right)$$

$$= \left(-\left(-d_4\left(p_4+k_1\right)p_2+p_4\,d_4\,k_1\right)k_2+p_2\,p_4\,d_4\,k_1\right)P_{Total}+k_1\,k_2\,p_2<0$$

Now, when *Condition* is negative, the signs of C9 and C0 are negative and positive respectively indicating the presence presence of atleast one positive feasible solution for epsilon for every choice for a value of m (and thus App, and the concentration of the other variables).

**Thus we have shown how enzyme biphasic response is guarenteed to exist in pYpTErk, for some total concentration of Erk and Kinase in the system, when the above condition is satisfied.**

**This is not a necessary condition though, biphasic responses can/might exist when the above condition is not satisfied. However this condition represents ONE possible sufficiency condition for obtaining the biphasic dose response in the substrate.**

Note: We illustrate the validity of this argument by showing the method predicting the presence of a biphasic response for the arbitraty value of 1 for m (i.e. App = $\frac{3}{\epsilon}$) and some abitrary choice of kinetics that satisfies *Condition*.

$m := 3$ :

$k_1 := 1 : k_2 := 1 : k_4 := 1 : p_2 := 0.1 : p_3 := 1 : p_4 := 1$ :

$k_{b1} := 1 : k_{b2} := 1 : k_{b4} := 1 : p_{b2} := 1 : p_{b3} := 1 : p_{b4} := 1 : k_{ub1} := 1 : k_{ub2} := 1 : k_{ub4} := 1 : p_{b2} := 1 :$
$p_{b3} := 1 : p_{b4} := 1 : p_{ub2} := 1 : p_{ub3} := 1 : p_{ub4} := 1$ :

$$c_1 := \frac{k_{b1}}{k_1+k_{ub1}} : c_2 := \frac{k_{b2}}{k_2+k_{ub2}} : c_4 := \frac{k_{b4}}{k_4+k_{ub4}} : d_3 := \frac{p_{b3}}{p_3+p_{ub3}} : d_2 := \frac{p_{b2}}{p_2+p_{ub2}} : d_4 := \frac{p_{b4}}{p_4+p_{ub4}} :$$

$P_{Total} := 1$ :

$Condition = -0.2500000000 < 0$

$\epsilon := evalf(solve(T, useassumptions) \text{ assuming } \epsilon > 0) = 4.551992989$
$App = 0.6590519817$
$ErkCon = Erk_{Total} - 2.771482748$
$KCon =$

$K_{Total} - 4.159218072$

**The computational evidence of the biphasic dose response behavior predicted above is shown in supplementary figure 1B.**

# Network regulation meets substrate modification chemistry
# Supplementary file 1

**Vaidhiswaran Ramesh[1], Thapanar Suwanmajo[2,3], J Krishnan[1,4,*]**

**Affiliations:**
1 - Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, Imperial College London, London, SW7 2AZ, UK
2 - Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
3 - Center of Excellence in Materials Science and Technology, Chiang Mai University, Chiang Mai 50200, Thailand
4 - Institute for Systems and Synthetic Biology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

* - Corresponding author (j.krishnan@ic.ac.uk)

This supplementary file 1 (Maple workbook and its PDF version) provides a detailed description of the mathematical models employed in the manuscript and further details the various analytical proofs referenced. A brief description of the semi-analytical approaches and how they can be brought to bear in such models is also shown.

The general organization of this Maple workbook is as follows. There are a total of 12 Maple worksheets organized across 3 folders (and many subfolders). As seen below, the folders and subfolders are broadly organized based on the type of the network regulation. Under each folder/subfolder, the relevant analysis/proofs for the respective system/model is presented. Each folder is prefixed with a "Read_Me" file that summarizes the results enclosed and provides a brief synthesis.

**Organization of Results**

Read_Me (This file): General explanation on the organization and accessibility of the text [Page 1]

Folder 2: Feedback networks (with results pertinent to multi-stationarity [Page 2-86] and biphasic response behavior [Page 87-105])
Folder 3: Feedforward networks (with results pertinent to multi-stationarity and isolas) [Page 106-122]
Folder 4: Semi-analytical approaches [Page 123-130]

Each worksheet begins with a set of instructions on how to run the document, and is intended to be read and accessed independent of the other documents in the workbook. The worksheets are numbered (as are the folders) and this is used to reference specific results in the main text of the manuscript.

Please take caution to run the Maple worksheets exactly as instructed (e.g., running the sections and subsections relevant to the proof and not the entire document).

# Feedback models (positive and negative feedback mediated by App or Ap)

This folder deals with results pertaining to the ordered DSP network (with common kinase and common phosphatase) with feedback from App (see Fig. 1B(i, ii)) and analogous feedback from Ap, realized through activation/deactivation of the kinase. Detailed mathematical models are presented in the beginning of each file (in section '0.') along with instructions on how to run the file (please ignore if you are reading this on the PDF version).

The organization of the results in this section are as follows. Each folder deals with a specific set of results (summarized below) pertaining to feedback in a class of models. The results correspond to the presence/absence of multi-stationarity in the ordered DSP with feedback regulation (realized in different ways), and the presence and features of biphasic responses in the dose-response curve of the maximally modified substrate form as a function of signal strength.

**Background**
Before exploring and summarizing the various results in this context, we draw attention to related results in literature to contextualize the work that follows. In particular we wish to highlight results from literature exploring the innate capacity of the ordered DSP network to enable multi-stationarity, and biphasic responses in the dose-response of the maximally modified substrate with signal/enzyme.

Due to the work of Wang and Sontag, (On the number of steady states in a multiple futile cycle. Journal of Mathematical Biology, 57(1): 29–52, 2008) and others in literature, we know that the ordered DSP network with common kinase and phosphatase is capable of at most exhibiting three steady states (stable or unstable) - thus permitting multi-stationarity without the need for any explicit feedback regulation. Conradi and Mincheva, (Catalytic constants enable the emergence of bistability in dual phosphorylation. J. R. Soc. Interface 11: 20140158) expands on this work and provides insights into how certain key parameters (catalytic constants) enable such multi-stationarity.

It has similarly been established by Suwanmajo and Krishnan, (Biphasic responses in multisite phosphorylation systems J. R. Soc. Interface.102013074220130742, 2013) that multisite modification is capable of inherently presenting biphasic responses in the dose-response curve of maximally modified substrate as a function of total enzyme concentration.

We now recapitulate a couple of computational insights established in the figures in the main text of this manuscript.

1. Positive feedback (PFB) from App and Negative feedback (NFB) from App were each able to present tristability when the feedback was mediated through a simple covalent modification tier.
2. Analogous feedback from the Ap substrate form (both positive and negative feedback analogues - each representing a mixed feedback in this instance) did not present tristability computationally (to the best of our knowledge).
3. Biphasic behavior is perturbed in a unique manner by feedback; wherein, the computational results suggest that the peak concentration reached by the maximally modified substrate form is unaffected, as is the saturation concentration at infinite signal strength. However we see PFB and NFB perturb the position of the biphasic peak vis-a-vis signal strength moving it towards lower and higher values respectively.

**Analytical results**
The analytical proofs in this folder build on these to present new insights in an effort to delineate the source of multi-stationarity in such networks and the different ways in which PFB and NFB can interact with the modification biochemistry to enable multi-stationarity and perturb biphasic responses.

In particular:

1. We show that when the feedback is from App, both positive feedback and negative feedback is capable of exhibiting increased multi-stationarity (tristability), even in the enzyme limiting regime (Total substrate concentration is much greater than the total enzyme concentrations), so long as the feedback is mediated through a covalent modification tier (Folder 2.2).
2. In the absence of such a tier (when the feedback is directly mediated) we show that the App PFB is

capable of increased multi-stationarity (tristability) both in the full system and when the network operates in the enzyme limiting regime. App NFB on the other hand is incapable of increased tristability when the system operates in the enzyme limiting regime. Analogous negative feedback (through kinase deactivation) from Ap is also shown to be incapable of exhibiting tristability, irrepsective of enzyme regime (Folder 2.3).
3. To further study the effect of different feedbacks - we study reduced networks of DSP without (some/all) key complexes (fully saturated limit) and show the key differences between PFB and NFB from App and Ap (Folder 2.4). In particular these results include multiple instances where the presence of key components of the double site modification system enable positive feedback and negative feedback to act in different ways - indicating how the biochemistry of the multisite modification can alter expected network regulation behavior and capability.

We also present results pertaining to the influence of the different feedback regulation on the biphasic behavior exhibited by the maximally modified substrate form in dose-response curves as a funciton of signal.

4. We present results to understand the systems' capacity to exhibit biphasic behavior and how this is impacted by the feedback regulation. In particular these results show how key features of the biphasic response are retained even as network regulation perturbs other characteristics of the behavior (Folder 2.5)

**Synthesis**
These results show that there are different ingredients of the feedback and key components of the multisite modification that contribute to the emergence of new qualitative systems behavior. In the case of multi-stationarity it is shown that each of these ingredients is integral to establishing multi-stationarity, while for biphasic behavior, we show that network regulation through feedback acts mostly as a lever that perturbs the behavior, with the behavior itself and its significant features being guaranteed by key drivers belonging to the biochemistry of the multisite modification. Thus depending on the behavior of concern, network regulation or multisite modification biochemistry or a combination of both can enable and introduce interesting and nonintuitive information processing characteristics, that is beyond simple interpolations of any isolated single-level behavior.

# App positive feedback (PFB)

This document deals with results pertaining to the ordered DSP model (with common kinase and common phosphatase) with positive feedback from the maximally modified substrate form (App) through activation of the kinase (mediated by a covalent modification tier).

We detail key analytical proofs concerning the behavioral characteristics of the PFB model, drawing strict boundaries between the possibility and impossibility of multi-stationarity. These results are organized as follows,

**0. Model:** The App PFB model mediated by a covalent modification tier.
**1. Solving the system of equations at steady state:** Systematically solving (algebraically) the ODEs at steady state to obtain expressions relating the steady state concentrations of the variables in terms of each other.
**2. Presence of increased multi-stationarity (enabling tristability):** Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values
**3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime**: Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values (when the network is operating in the enzyme limiting regime, i.e. total substrate concentration is much greater than total enzyme concentrations).

**Instructions**
Run the first two sections (0. and 1.) before running only the required section, i.e. 2. or 3. (ignore if you are accessing this on the PDF).

# 0. App feedback (positive) mediated by a covalent modification tier (PFB)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot S \cdot Ks + k_{PFB} \cdot X \cdot Ks - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks - X \cdot k_{PFB} \cdot Ks + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP :$$

$$dXs := -d_a \cdot Xs \cdot App + d_d \cdot X :$$
$$dX := d_a \cdot Xs \cdot App - d_d \cdot X :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase in the absence of signal. It can take the value of 0 without affecting any of the model dynamics. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$
$$XCon := X + Xs - X_{Total} :$$

# 1. Solving the system of equations at steady state

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of App and a ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{d_d X}{App\ d_a}$$

We similarly solve for the other variables using the same command.

```
Ks := solve(dKs, Ks) :
AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP :=
      solve(dApP, ApP) :

Sol := solve({dA, dAp}, {A, Ap}) :
assign(Sol)
```

We now introduce a new ratio, $\epsilon = \frac{K}{P}$, and we also introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

```
K := ϵ·P :
k_b1 := c_1·(k_1 + k_ub1) : k_b2 := c_2·(k_2 + k_ub2) : k_b3 := c_3·(k_3 + k_ub3) : k_b4 := c_4·(k_4 + k_ub4) :
```

Once this is done, we again solve for the steady state of the phosphatase and the substrate involved in the covalent modification cycle using the conservation expression for the enzyme (PCon), and the substrate respectively (XCon).

```
X := solve(XCon, X) :
P := solve(simplify(PCon), P) :
```

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration (App) and $\epsilon$.

$$AK = \frac{k_4 c_4 c_3 App\ k_3 P_{Total}}{k_1 \left(App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2\right)}$$

$$ApK = \frac{c_2 c_3 App\ k_3 \epsilon P_{Total}}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$AppP = \frac{c_3 App\ P_{Total} \epsilon k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$ApP = \frac{c_4 c_3 App\ k_3 P_{Total}}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$Ks = \frac{u_d \epsilon^2 P_{Total} k_2 c_2}{\left(App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2\right)\left(S\ u_a + \frac{X_{Total} App\ d_a k_{PFB}}{App\ d_a + d_d} + u_\beta\right)}$$

$$K = \frac{\epsilon^2 P_{Total} k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$P = \frac{P_{Total} \epsilon k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

Note that when App and $\epsilon$ are positive, the other variable concentrations are positive as well. Thus we have partially solved the system of equations to arrive at expressions linking the steady state concentrations of the variables with that of App and $\epsilon$.

## 2. Presence of increased multi-stationarity (enabling

## tristability)

We now have two expressions, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform this simplification.

$T := simplify(numer(ACon))$

$$T := k_1 c_2^2 k_2 c_1 \left(App^2 c_3 k_2 + \left(\left(\left(-A_{Total} + P_{Total}\right) k_2 + P_{Total} k_3\right) c_3 + k_2\right) App - A_{Total} k_2\right) \epsilon^3 + \left(k_1 \left(c_3 + c_4\right) App + \left(1 + \left(-A_{Total} + P_{Total}\right) c_4\right) k_1 + P_{Total} c_4 k_4\right) k_3 App\, c_2 k_2 c_1 c_3 \epsilon^2 + k_3 App \left(c_3 \left(c_1 k_1 k_3 + c_2 k_2 k_4\right) App + c_2 k_2 k_4\right) c_4 c_3 \epsilon + App^2 c_3^2 c_4^2 k_3^2 k_4 \quad (3.1)$$

$G := simplify(numer(KCon))$

$$G := \left(d_a \left(S u_a + k_{PFB} X_{Total} + u_\beta\right) \left(c_2 \left(-k_2 K_{Total} + P_{Total} k_3\right) \epsilon - k_3 K_{Total} c_4\right) c_3 App^2 + \left(P_{Total} c_2 d_a k_2 \left(S u_a + k_{PFB} X_{Total} + u_d + u_\beta\right) \epsilon^2 + \left(d_d \left(-k_2 K_{Total} + P_{Total} k_3\right) \left(S u_a + u_\beta\right) c_3 - K_{Total} d_a k_2 \left(S u_a + k_{PFB} X_{Total} + u_\beta\right)\right) c_2 \epsilon - k_3 K_{Total} c_3 c_4 d_d \left(S u_a + u_\beta\right)\right) App + \epsilon c_2 \left(P_{Total} \left(S u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S u_a + u_\beta\right)\right) d_d k_2\right) k_1 + \left(d_a \left(S u_a + k_{PFB} X_{Total} + u_\beta\right) App + d_d \left(S u_a + u_\beta\right)\right) k_3 c_4 k_4 P_{Total} App\, c_3 \quad (3.2)$$

The following parameter set admits tristability in the model. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking them.

$k_1 := 2.226443427156099 : k_2 := 34.441456743290490 : k_3 := 80.224008796362380 : k_4 := 15.270118706808322 :$
$k_{b1} := 79.008883130076730 : k_{b2} := 2.552189964337821e+03 : k_{b3} := 2.490887667747111e+02 : k_{b4} := 54.631264692401650 :$
$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 : \quad k_{PFB} := 25.862959673799264 :$
$S := 1.607685603083320 : u_a := 4.679399737235169e-4 : u_d := 19.769496729482345 :$
$u_\beta := 0 : d_a := 1 : d_d := 1 :$
$A_{Total} := 29.341031578113412 : K_{Total} := 17.307197261951845 : P_{Total} := 4.934324749201547 : X_{Total} := 1 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

These parameter values result in the expressions T and G taking the form.

$simplify(T) =$
$(1.028487799 \times 10^9\, App^2 - 1.294574855 \times 10^{10}\, App - 9.840219823 \times 10^9)\, \epsilon^3 + (2.137257745 \times 10^8\, App^2 + 1.087254646 \times 10^9\, App)\, \epsilon^2 + (1.070249873 \times 10^8\, App^2 + 3.128599508 \times 10^7\, App)\, \epsilon + 1.042057402 \times 10^7\, App^2$
$simplify(G) =$
$(1.243378476 \times 10^6\, App + 538684.4977)\, \epsilon^2 + (-2.546308422 \times 10^6\, App^2 - 2.471867688 \times 10^6\, App - 71.89730664)\, \epsilon + 786558.7762\, App^2 + 22.87871322\, App$

Now we use the inbuilt *Maple* solver *solve* to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ :
$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)]$ :
$Array(data)$

$$\begin{bmatrix} 1 & App = 9.134930984 \times 10^{-6} & \epsilon = 0.0001724258844 \\ 2 & App = 0.0001199395341 & \epsilon = 0.0006431460374 \\ 3 & App = 0.1028503992 & \epsilon = 0.03202356481 \\ 4 & App = 2.400532536 & \epsilon = 0.2289151319 \\ 5 & App = 13.17482346 & \epsilon = 27.75568492 \end{bmatrix} \quad \textbf{(3.3)}$$

As we can see we thus have 5 positive solutions for App, $\epsilon$ and by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have shown increased multi-stationarity in the model, involving five positive steady states (enabling tristability).**

## 3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime

In this subsection of the analysis we wish to demonstrate the presence of increased multi-stationarity (enabling tristability) in the limit where the concentrations of enzymes are much less than the the concentration of the substrates in the system. In such a scenario the conservation expression for the substrate concentration can be suitably modified to ignore the complexes involving the enzymes. This expression (ACon_Red = 0) is given below.

$ACon\_Red := A_{Total} - A - Ap - App$ :

Now the steady state of the system is defined by the expressions ACon_Red and KCon which are expressions involving App and $\epsilon$. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplification.

$T := simplify(numer(ACon\_Red))$

$$T := -c_1 c_2 k_1 k_2 (App - A_{Total}) \epsilon^2 - c_3 App k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App k_3 \quad \textbf{(4.1)}$$

$G := simplify(numer(KCon))$

$$\begin{aligned} G := {} & ((c_2 (-k_2 K_{Total} + k_3 P_{Total}) \epsilon - k_3 K_{Total} c_4) (S u_a + k_{PFB} X_{Total} + u_\beta) c_3 d_a App^2 \\ & + (P_{Total} c_2 d_a k_2 (S u_a + k_{PFB} X_{Total} + u_d + u_\beta) \epsilon^2 + (d_d (-k_2 K_{Total} + k_3 P_{Total}) (S u_a \\ & + u_\beta) c_3 - K_{Total} d_a k_2 (S u_a + k_{PFB} X_{Total} + u_\beta)) c_2 \epsilon - k_3 K_{Total} c_3 c_4 d_d (S u_a + u_\beta)) App \\ & + d_d \epsilon c_2 k_2 (P_{Total} (S u_a + u_d + u_\beta) \epsilon - K_{Total} (S u_a + u_\beta))) k_1 + (d_a (S u_a + k_{PFB} X_{Total} \\ & + u_\beta) App + d_d (S u_a + u_\beta)) c_4 k_4 c_3 P_{Total} App k_3 \end{aligned} \quad \textbf{(4.2)}$$

The following parameter set admits increased multi-stationarity (enabling tristability) in the model. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking the variables.

$k_1 := 1.496019776669208 : k_2 := 24.894304930673254 : k_3 := 69.949359115788620 : k_4 := 11.899049706995182 :$
$k_{b1} := 0.240968878931967 : k_{b2} := 50.609447250163065 : k_{b3} := 1.067396494946383e+02 : k_{b4} := 30.082694888232470 :$
$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 :$
$A_{Total} := 1.370106929866000e+02 : K_{Total} := 6.428145913885926 : P_{Total} := 1.881093427076993 : X_{Total} := 1 :$
$S := 1.607685603083320e-04 : k_{PFB} := 22.650066364381720 :$
$u_a := 0.378358419530540 : u_d := 2.198199412258635 : u_\beta := 0 :$
$d_a := 1 : d_d := 1 :$
$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$

These parameter values result in the expressions T and G taking the form.

$simplify(T) = (-7.027123898\ App + 962.7911151)\ \epsilon^2 - 15.19887017\ App\ \epsilon - 2920.332734\ App$
$simplify(G) =$
$(3402.296735\ App + 300.9914413)\ \epsilon^2 + (-2833.917283\ App^2 - 10597.93762\ App - 0.02846131274)\ \epsilon + 70968.50234\ App^2 + 0.1905897413\ App$

Now we use the inbuilt *Maple* solver *solve* to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ :
$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)] :$
$Array(data)$

$$\begin{bmatrix} 1 & App = 2.953072699 \times 10^{-9} & \epsilon = 0.00009464274056 \\ 2 & App = 0.005254219372 & \epsilon = 0.1262860076 \\ 3 & App = 0.02468970125 & \epsilon = 0.2738775809 \\ 4 & App = 112.7858430 & \epsilon = 49.30917649 \\ 5 & App = 121.1812579 & \epsilon = 65.28758768 \end{bmatrix} \quad \textbf{(4.3)}$$

As we can see we have 5 positive solutions for App, $\epsilon$ and thus by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have the increased multi-stationarity involving five steady states (enabling tristability) even when the system operates in the enzyme limiting regime (where the total enzyme concentrations are much lower than the total substrate concentration).**

# App negative feedback (NFB)

This document deals with the results pertaining to the ordered DSP model (with common kinase and common phosphatase) with negative feedback from the maximally modified substrate form (App) through deactivation of the kinase mediated by a covalent modification tier.

We detail the following key analytical proofs concerning the behavioral characteristics of the NFB model, drawing strict boundaries between the possibility and impossibility of multi-stationarity. These results are organized as follows,

**0. Model:** The App NFB model mediated by a covalent modification tier.
**1. Solving the system of equations at steady state:** Systematically solving (algebraically) the ODEs at steady state to obtain expressions relating the steady state concentrations of the variables in terms of each other.
**2. Presence of increased multi-stationarity (enabling tristability):** Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values
**3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime**: Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values (when the network is operating in the enzyme limiting regime, i.e. total substrate is much greater than total enzyme).

**Instructions**
Run the first two sections (0. and 1.) before running only the required section, i.e. 2. or 3. (ignore if you are accessing this on the PDF).

# 0. App feedback (negative) mediated by a covalent modification tier (PFB)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot S \cdot Ks - k_{NFB} \cdot X \cdot K - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks + X \cdot k_{NFB} \cdot K + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP :$$

$$dXs := -d_a \cdot Xs \cdot App + d_d \cdot X :$$
$$dX := d_a \cdot Xs \cdot App - d_d \cdot X :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase in the absence of signal. It can take the value of 0 without affecting any of the model dynamics. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$
$$XCon := X + Xs - X_{Total} :$$

# 1. Solving the system of equations at steady state

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of App and a ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{d_d X}{App\ d_a}$$

We similarly solve for the other variables using the same command as shown.

```
Ks := solve(dKs, Ks) :
X := solve(XCon, X) :
AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP :=
    solve(dApP, ApP) :

Sol := solve( {dA, dAp}, {A, Ap}) :
assign(Sol)
```

We now introduce a new ratio, $\epsilon = \frac{K}{P}$, and we also introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

```
K := ϵ·P :
k_b1 := c_1·(k_1 + k_ub1) : k_b2 := c_2·(k_2 + k_ub2) : k_b3 := c_3·(k_3 + k_ub3) : k_b4 := c_4·(k_4 + k_ub4) :
```

Once this is done, we again solve for the steady state concentration of the free phosphatase using the conservation expression for the enzyme (PCon),

```
P := solve(simplify(PCon), P) :
```

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration (App) and $\epsilon$.

$$AK = \frac{k_4 c_4 c_3 App\ k_3 P_{Total}}{k_1 \left(App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2\right)}$$

$$ApK = \frac{c_2 c_3 App\ k_3 \epsilon P_{Total}}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$AppP = \frac{c_3 App\ P_{Total} \epsilon k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$ApP = \frac{c_4 c_3 App\ k_3 P_{Total}}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$Ks = \frac{\epsilon^2 P_{Total} k_2 c_2 \left(\frac{X_{Total} App\ d_a k_{NFB}}{App\ d_a + d_d} + u_d\right)}{\left(App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2\right)\left(S\ u_a + u_\beta\right)}$$

$$K = \frac{\epsilon^2 P_{Total} k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

$$P = \quad = \frac{P_{Total} \epsilon k_2 c_2}{App\ c_2 c_3 \epsilon k_2 + c_4 c_3 App\ k_3 + \epsilon k_2 c_2}$$

Note that when App and $\epsilon$ are positive, the other variable concentrations are positive as well. Thus we have partially solved the system of equations to arrive at expressions linking the steady state concentrations of the variables with that of App and $\epsilon$.

## 2. Presence of increased multi-stationarity (enabling tristability)

We now have two expressions, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform this simplification.

$$T := simplify(numer(ACon))$$

$$T := k_1 \left(App^2 c_3 k_2 + \left(\left(\left(-A_{Total} + P_{Total}\right) k_2 + P_{Total} k_3\right) c_3 + k_2\right) App - A_{Total} k_2\right) c_1 k_2 c_2^2 \epsilon^3 + c_3 \left(k_1 \left(c_3 + c_4\right) App + \left(1 + \left(-A_{Total} + P_{Total}\right) c_4\right) k_1 + P_{Total} c_4 k_4\right) c_1 k_2 App\, k_3 c_2 \epsilon^2 + c_3 \left(c_3 \left(c_1 k_1 k_3 + c_2 k_2 k_4\right) App + c_2 k_2 k_4\right) c_4 App\, k_3 \epsilon + App^2 c_3^2 c_4^2 k_3^2 k_4 \quad (3.1)$$

$$G := simplify(numer(KCon))$$

$$G := \left(\left(c_2 \left(-k_2 K_{Total} + P_{Total} k_3\right) \epsilon - k_3 K_{Total} c_4\right) c_3 d_a \left(S u_a + u_\beta\right) App^2 + \left(P_{Total} c_2 d_a k_2 \left(S u_a + k_{NFB} X_{Total} + u_d + u_\beta\right) \epsilon^2 + \left(S u_a + u_\beta\right) \left(d_d \left(-k_2 K_{Total} + P_{Total} k_3\right) c_3 - K_{Total} d_a k_2\right) c_2 \epsilon - k_3 K_{Total} c_3 c_4 d_d \left(S u_a + u_\beta\right)\right) App + k_2 \left(P_{Total} \left(S u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S u_a + u_\beta\right)\right) d_d \epsilon c_2\right) k_1 + k_4 c_4 c_3 App\, k_3 P_{Total} \left(App\, d_a + d_d\right) \left(S u_a + u_\beta\right) \quad (3.2)$$

The following parameter set admits tristability in the model. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking the variables.

$$k_1 := 1 : k_2 := 1 : k_3 := 0.9 : k_4 := 100 :$$
$$k_{b1} := 1 : k_{b2} := 1 : k_{b3} := 200 : k_{b4} := 3 :$$
$$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 :$$
$$u_a := 1 : u_d := 4 : u_\beta := 0.01 :$$
$$d_a := 2 : d_d := 1 :$$
$$S := 0.067 : k_{NFB} := 2 :$$
$$A_{Total} := 40 : K_{Total} := 20 : P_{Total} := 1 : X_{Total} := 10 :$$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

These parameter values result in the expressions T and G taking the form.

$$simplify(T) = \left(13.15789474\, App^2 - 501.1907895\, App - 5.\right) \epsilon^3 + \left(2493.778285\, App^2 + 66.59718604\, App\right) \epsilon^2 + \left(14943.63841\, App^2 + 140.6982804\, App\right) \epsilon + 791.8402438\, App^2$$

$$simplify(G) = \left(24.077\, App + 2.038500000\right) \epsilon^2 + \left(-154.8105264\, App^2 - 78.94526315\, App - 0.7700000000\right) \epsilon + 34.66805627\, App^2 + 17.33402814\, App =$$

Now we use the inbuilt *Maple* solver *solve* to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions) \text{ assuming } \epsilon > 0, App > 0 :$$
$$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)] :$$
$$Array(data)$$

$$\begin{bmatrix} 1 & App = 0.006643206660 & \epsilon = 0.4817332947 \\ 2 & App = 0.08183096913 & \epsilon = 1.838201268 \\ 3 & App = 0.3039385674 & \epsilon = 3.945827691 \\ 4 & App = 0.8840197992 & \epsilon = 7.984809681 \\ 5 & App = 6.405197361 & \epsilon = 43.66319243 \end{bmatrix} \quad \textbf{(3.3)}$$

As we can see we thus have 5 positive solutions for App, $\epsilon$ and thus by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have shown increased multi-stationarity in the model, involving five positive steady states (enabling tristability).**

## 3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime

In this section of the analysis we wish to demonstrate the presence of increased multi-stationarity (enabling tristability) in the limit where the concentrations of enzymes are much less than the concentration of the substrates in the system. In such a scenario the conservation expression for the substrate concentration can be suitably modified to ignore the complexes involving the enzymes. This expression (ACon_Red = 0) is given below.

$ACon\_Red := A_{Total} - A - Ap - App$ :

Now the steady state of the system is defined by the expressions ACon_Red and KCon involving App and $\epsilon$. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplification.

$T := simplify(numer(ACon\_Red))$

$$T := -c_1 c_2 k_1 k_2 \left(App - A_{Total}\right) \epsilon^2 - c_3 App\, k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App\, k_3 \quad \textbf{(4.1)}$$

$G := simplify(numer(KCon))$

$$\begin{aligned} G := {} & \left(\left(S u_a + u_\beta\right) d_a c_3 \left(c_2 \left(-k_2 K_{Total} + k_3 P_{Total}\right) \epsilon - k_3 K_{Total} c_4\right) App^2 + \left(P_{Total} c_2 d_a k_2 \left(S u_a\right.\right.\right. \\ & + k_{NFB} X_{Total} + u_d + u_\beta\left.\left.\right) \epsilon^2 + \left(d_d \left(-k_2 K_{Total} + k_3 P_{Total}\right) c_3 - K_{Total} d_a k_2\right) \left(S u_a + u_\beta\right) c_2 \epsilon \\ & - k_3 K_{Total} c_3 c_4 d_d \left(S u_a + u_\beta\right)\left.\right) App + d_d \epsilon k_2 \left(P_{Total} \left(S u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S u_a\right.\right. \\ & \left.\left. + u_\beta\right)\right) c_2\left.\right) k_1 + k_4 c_4 c_3 App\, k_3 P_{Total} \left(App\, d_a + d_d\right) \left(S u_a + u_\beta\right) \end{aligned} \quad \textbf{(4.2)}$$

The following parameter set admits tristability in the model. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking the variables.

$k_1 := 4.164912079945205$ : $k_2 := 0.410512609395469$ : $k_3 := 0.229071749963111$ : $k_4 := 76.606951634495970$ :
$k_{b1} := 0.693683329319052$ : $k_{b2} := 2.155108632165148$ : $k_{b3} := 1.356531391212892e+02$ : $k_{b4} := 4.415668113427349$ :
$k_{ub1} := 0.418828504079182$ : $k_{ub2} := 0.428862554615646$ : $k_{ub3} := 0.298624002337422$ : $k_{ub4} := 0.585710088258362$ :
$u_a := 0.138455242394071$ : $u_d := 7.050392004917759$ : $u_\beta := 0$ :
$d_a := 0.844606655897927$ : $d_d := 0.754817479130625$ :

$S := 12.0506182578860 15 : k_{NFB} := 33.111593992179540 :$
$A_{Total} := 1.703694160053648e+02 : K_{Total} := 6.865336220784366 : P_{Total} := 4.146324379634737 :$
$X_{Total} := 13.877955348813080 :$
$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$

These parameter values result in the expressions T and G taking the form.

$simplify(T) = (-0.6643342110\ App + 113.1822316)\ \epsilon^2 - 37.11631467\ App\ \epsilon - 258.0514906\ App$
$simplify(G) =$
$(7198.330075\ App + 119.7872189)\ \epsilon^2 + (-7238.219977\ App^2 - 6511.202826\ App$
$- 37.95490283)\ \epsilon + 1372.066011\ App^2 + 1226.203228\ App$

Now we use the inbuilt *Maple* solver *solve* to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ :
$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)]$ :
$Array(data)$

$$\begin{bmatrix} 1 & App = 0.003456278215 & \epsilon = 0.08933988374 \\ 2 & App = 0.02204808056 & \epsilon = 0.2278664329 \\ 3 & App = 0.4658271790 & \epsilon = 1.111407290 \\ 4 & App = 1.149399580 & \epsilon = 1.825100948 \\ 5 & App = 111.7550123 & \epsilon = 113.0721667 \end{bmatrix} \quad \textbf{(4.3)}$$

As we can see we thus have 5 positive solutions for App, $\epsilon$ and thus by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have shown increased multi-stationarity in the model, involving five positive steady states (enabling tristability) even when the system operates in the limiting enzyme regime (where the total enzyme concentrations are much lower than the total substrate concentration).**

# App positive feedback (PFB) - Direct

This document deals with results pertaining to the ordered DSP network (with common kinase and common phosphatase) with positive feedback from the maximally modified substrate form (App) by direct activation of the kinase.

In this document we detail the following key analytical proofs concerning the behvaioral characteristics of the PFB model, drawing strict boundaries between the possibility and impossibility of multi-stationarity. These results are organized as follows,

**0. Model:** The App PFB model mediated by direct activation of the kinase by the substrate form.
**1. Solving the system of equations at steady state:** Systematically solving (algebraically) the ODEs at steady state to obtain expressions relating the steady state concentrations of the variables in terms of each other.
**2. Presence of increased multi-stationarity (enabling tristability):** Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values.
**3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime**: Tristability is established in the system by working with the reduced steady state equations of the model and is demonstrated for a choice of parameter values (when the network is operating in the enzyme limiting regime, i.e. total substrate concentration is much greater than total enzyme concentration).

**Instructions**
Run the first two sections (0. and 1.) before running only the required section, i.e. 2. or 3. (ignore if you are accessing this on the PDF).

# 0. App feedback (positive) by activation of the kinase by the substrate form (PFB)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modelled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot S \cdot Ks + k_{PFB} \cdot App \cdot Ks - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks - App \cdot k_{PFB} \cdot Ks + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase in the absence of signal. It can take the value of 0 without affecting any of the model dynamics. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$

# 1. Solving the system of equations at steady state

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of App and a ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{App\ k_{PFB} + S\ u_a + u_\beta}$$

We similarly solve for the other variables using the same command as shown.

$AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP := solve(dApP, ApP) :$

$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$
$A := eval(A, Sol) : Ap := eval(Ap, Sol) :$

We now introduce a new ratio, $\epsilon$ = K/P, and also rewrite the kinetic constants with the new nomenclature (for $k_{b1}$, $k_{b2}$, $k_{b3}$, and $k_{b4}$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$K := \epsilon \cdot P :$
$k_{b1} := c_1 \cdot (k_1 + k_{ub1}) : k_{b2} := c_2 \cdot (k_2 + k_{ub2}) : k_{b3} := c_3 \cdot (k_3 + k_{ub3}) : k_{b4} := c_4 \cdot (k_4 + k_{ub4}) :$

Once this is done, we again solve for the steady state of phosphatase using the conservation expression for the enzyme (PCon).

$P := solve(simplify(PCon), P) :$

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration (App) and $\epsilon$.

$$AK = \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 (App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2)}$$

$$ApK = \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2}$$

$$AppP = \frac{c_3 App P_{Total} \epsilon k_2 c_2}{App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2}$$

$$ApP = \frac{c_4 c_3 App k_3 P_{Total}}{App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2}$$

$$Ks = \frac{u_d \epsilon^2 P_{Total} k_2 c_2}{(App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2) (App k_{PFB} + S u_a + u_\beta)}$$

$$K = \frac{\epsilon^2 P_{Total} k_2 c_2}{App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2}$$

$$P = \quad = \frac{P_{Total} \epsilon k_2 c_2}{App c_2 c_3 \epsilon k_2 + c_4 c_3 App k_3 + \epsilon k_2 c_2}$$

Thus we have partially solved the system of equations to arrive at expressions linking the steady state concentrations of the variables with that of App and $\epsilon$

## 2. Presence of increased multi-stationarity (enabling tristability)

We now have two expressions, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplificiation.

$T := simplify(numer(ACon))$
$$T := \left(App^2 c_3 k_2 + \left(\left(-A_{Total} + P_{Total}\right) k_2 + P_{Total} k_3\right) c_3 + k_2\right) App - A_{Total} k_2\right) k_1 c_2^2 k_2 c_1 \epsilon^3 \quad \textbf{(3.1)}$$

$$+ App\, k_3\, c_2\, k_2\, c_1 \left(k_1 \left(c_3 + c_4\right) App + \left(1 + \left(-A_{Total} + P_{Total}\right) c_4\right) k_1 + P_{Total}\, c_4\, k_4\right) c_3\, \epsilon^2$$
$$+ App\, k_3\, c_4 \left(c_3 \left(c_1\, k_1\, k_3 + c_2\, k_2\, k_4\right) App + c_2\, k_2\, k_4\right) c_3\, \epsilon + App^2\, c_3^2\, c_4^2\, k_3^2\, k_4$$

$G := simplify(numer(KCon))$

$$G := \left(-k_{PFB} \left(c_2 \left(k_2\, K_{Total} - P_{Total}\, k_3\right) \epsilon + K_{Total}\, c_4\, k_3\right) c_3\, App^2 + \left(k_{PFB}\, P_{Total}\, c_2\, k_2\, \epsilon^2 \right. \right. \quad \textbf{(3.2)}$$
$$\left. - c_2 \left(\left(k_2\, K_{Total} - P_{Total}\, k_3\right) \left(S\, u_a + u_\beta\right) c_3 + k_{PFB}\, K_{Total}\, k_2\right) \epsilon - K_{Total}\, c_3\, c_4\, k_3 \left(S\, u_a + u_\beta\right)\right)$$
$$App + c_2\, k_2 \left(P_{Total} \left(S\, u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S\, u_a + u_\beta\right)\right) \epsilon\Big)\, k_1$$
$$+ k_4\, c_4\, c_3\, App\, k_3\, P_{Total} \left(App\, k_{PFB} + S\, u_a + u_\beta\right)$$

The following data point shows a scenario with tristability in the system. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking the variables.

$k_1 := 2.226443427156099 : k_2 := 34.441456743290490 : k_3 := 80.224008796362380 : k_4 := 15.270118706808322 :$
$k_{b1} := 79.008883130076730 : k_{b2} := 2.552189964337821e+03 : k_{b3} := 2.490887667747111e+02 : k_{b4} := 54.631264692401650 :$
$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 :$
$A_{Total} := 29.341031578113412 : K_{Total} := 17.307197261951845 : P_{Total} := 4.934324749201547 :$

$S := 1.607685603083320e-04 : u_a := 4.679399737235169e-04 : u_d := 19.769496729482345 : k_{PFB} := 25.862959673799264 :$

$u_\beta := 0 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

These parameter values result in the expressions T and G taking the form.

$simplify(T) =$
$(1.028487799 \times 10^9\, App^2 - 1.294574855 \times 10^{10}\, App - 9.840219823 \times 10^9)\, \epsilon^3 + (2.137257745 \times 10^8\, App^2 + 1.087254646 \times 10^9\, App)\, \epsilon^2 + (1.070249873 \times 10^8\, App^2 + 3.128599508 \times 10^7\, App)\, \epsilon + 1.042057402 \times 10^7\, App^2$
$=$
$simplify(G) =$
$(704693.9782\, App + 538664.0018)\, \epsilon^2 + (-2.546234358 \times 10^6\, App^2 - 2.471721733 \times 10^6\, App - 0.007189730664)\, \epsilon + 786535.8969\, App^2 + 0.002287871322\, App$
$=$

Now we use the inbuilt *Maple* solver solve to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ :
$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)]$ :
$Array(data)$

$$\begin{bmatrix} 1 & App = 5.603494211 \times 10^{-14} & \epsilon = 1.334757760 \times 10^{-8} \\ 2 & App = 0.0001954834874 & \epsilon = 0.0008297305948 \\ 3 & App = 0.1059530789 & \epsilon = 0.03264859832 \\ 4 & App = 2.304615646 & \epsilon = 0.2229433927 \\ 5 & App = 13.23075949 & \epsilon = 48.22134836 \end{bmatrix} \quad \textbf{(3.3)}$$

As we can see we thus have 5 positive solutions for App, $\epsilon$ and thus by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have shown increased multi-stationarity in the model, involving five positive steady states (enabling tristability).**

# 3. Presence of increased multi-stationarity (enabling tristability) - Enzyme limiting regime

Again similar to the analysis earlier - we now have two expressions, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system.

However in this subsection of the analysis we wish to demonstrate the presence of increased multi-stationarity (enabling tristability) in the limit where the total concentrations of enzymes are much less than the total concentration of the substrates in the system. In such a scenario the conservation expression for the substrate concentration can be suitably modified to ignore the complexes involving the enzymes. This expression (ACon_Red = 0) is given below.

$ACon\_Red := A_{Total} - A - Ap - App$ :

Now thus the steady state of the system is defined by the expressions ACon_Red and KCon involving App and $\epsilon$. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplificiation.

$T := simplify(numer(ACon\_Red))$

$$T := -c_1 c_2 k_1 k_2 \left(App - A_{Total}\right) \epsilon^2 - c_3 App\, k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App\, k_3 \quad \textbf{(4.1)}$$

$G := simplify(numer(KCon))$

$$\begin{aligned} G := {} & \left(-c_3 \left(c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right) \epsilon + K_{Total} c_4 k_3\right) k_{PFB} App^2 + \left(k_{PFB} P_{Total} c_2 k_2 \epsilon^2\right.\right. \\ & - c_2 \left(\left(k_2 K_{Total} - k_3 P_{Total}\right) \left(S u_a + u_\beta\right) c_3 + k_{PFB} K_{Total} k_2\right) \epsilon - K_{Total} c_3 c_4 k_3 \left(S u_a + u_\beta\right)\Big) \\ & App + \left(P_{Total} \left(S u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S u_a + u_\beta\right)\right) \epsilon k_2 c_2\Big) k_1 \\ & + k_4 c_4 c_3 App\, k_3 P_{Total} \left(App\, k_{PFB} + S u_a + u_\beta\right) \end{aligned} \quad \textbf{(4.2)}$$

The following data point shows a scenario with tristability in the system. Note that positive steady state concentration solutions of {App, $\epsilon$} solutions of expressions T and G (from above) will necessarily correspond to positive concentrations of the other variables at steady state as we have seen earlier with the expressions linking the variables.

$k_1 := 1.496019776669208$ : $k_2 := 24.894304930673254$ : $k_3 := 69.949359115788620$ : $k_4 := 11.899049706995182$ :
$k_{b1} := 0.240968878931967$ : $k_{b2} := 50.609447250163065$ : $k_{b3} := 1.067396494946383e{+}02$ : $k_{b4} := 30.082694888232470$ :
$k_{ub1} := 1$ : $k_{ub2} := 1$ : $k_{ub3} := 1$ : $k_{ub4} := 1$ :

$k_{PFB} := 22.650066364381720 : S := 1.607685603083320\text{e-}04 : u_a := 0.378358419530540 : u_d := 2.198199412258635 :$
$A_{Total} := 1.370106929866000\text{e}+02 : K_{Total} := 6.428145913885926 : P_{Total} := 1.881093427076993 :$
$u_\beta := 0 :$

$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

These parameter values result in the expressions T and G taking the form.

$$simplify(T) = (-7.027123898\, App + 962.7911151)\, \epsilon^2 - 15.19887017\, App\, \epsilon - 2920.332734\, App$$
$$simplify(G) = (3101.305293\, App + 300.9914413)\, \epsilon^2 + (-2833.909672\, App^2 - 10597.90916\, App - 0.02846131274)\, \epsilon + 70968.31172\, App^2 + 0.1905897413\, App$$
$$=$$

Now we use the inbuilt *Maple* solver *solve* to solve these coupled polynomial equations. Here we are naturally only interested in the positive solutions of App and $\epsilon$ and thus only isolate them.

$Sol := solve(\{T, G\}, \{App, \epsilon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ :
$data := [seq([i, Sol[i, 1], Sol[i, 2]], i = 1..5)]$ :
$Array(data)$

$$\begin{bmatrix} 1 & App = 2.953072699 \times 10^{-9} & \epsilon = 0.00009464274056 \\ 2 & App = 0.005154346920 & \epsilon = 0.1250795861 \\ 3 & App = 0.02576198263 & \epsilon = 0.2797670435 \\ 4 & App = 105.5761561 & \epsilon = 41.16827231 \\ 5 & App = 126.4729476 & \epsilon = 84.78615270 \end{bmatrix} \quad \textbf{(4.3)}$$

As we can see we thus have 5 positive solutions for App, $\epsilon$ and thus by extension for the other variables that satisfy the system of equations (from the system of ODEs) and the conservation conditions.

**Thus we have shown increased multi-stationarity in the model, involving five positive steady states (enabling tristability) even when we operate in the limiting enzyme regime where the enzyme concentrations are much lower than the substrate concentration.**

# App negative feedback (NFB) - Direct

This document deals with results pertaining to the ordered DSP (with common kinase and common phosphatase) network with negative feedback from the maximally modified substrate form (App) by direct deactivation of the kinase.

In this document we detail the following key analytical proofs concerning the behavioral characteristics of the NFB model, drawing strict boundaries between the possibility and impossibility of multi-stationarity. These results are organized as follows,

**0. Model:** The App NFB model mediated by direct activation of the kinase by the substrate form.
**1. Solving the system of equations at steady state:** Systematically solving (algebraically) the ODEs at steady state to obtain expressions relating the steady state concentrations of the variables in terms of each other.
**2. Absence of increased multi-stationarity (tristability) - Enzyme limiting regime:** We provide an analytical proof for the absence of increased multi-stationarity (tristability) in the network when the system is operating in the enzyme limiting regime.

## 0. App feedback (negative) by deactivation of the kinase by the substrate form (NFB)

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot S \cdot Ks - k_{NFB} \cdot App \cdot K - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks + App \cdot k_{NFB} \cdot K + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase in the absence of signal. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$

## 1. Solving the system of equations at steady state

Now we begin by solving the system of equations to obtain correlations between the steady state concentrations of the variables, primarily to obtain steady state correlations of variables as a function of App and the ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\,(App\,k_{NFB} + u_d)}{S\,u_a + u_\beta}$$

We similarly solve for the other variables using the same command as shown.

$$AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP :=$$

$solve(dApP, ApP)$ :

$Sol := solve(\{dA, dAp\}, \{A, Ap\})$ :
$assign(Sol)$

We now introduce a new ratio, $\epsilon$ = K/P, and also rewrite the kinetic constants with the new nomenclature (for $k_{b1}$, $k_{b2}$, $k_{b3}$, and $k_{b4}$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$K := \epsilon \cdot P$ :
$k_{b1} := c_1 \cdot (k_1 + k_{ub1})$ : $k_{b2} := c_2 \cdot (k_2 + k_{ub2})$ : $k_{b3} := c_3 \cdot (k_3 + k_{ub3})$ : $k_{b4} := c_4 \cdot (k_4 + k_{ub4})$ :

Once this is done, we again solve for the steady state of phosphatase using the conservation expression for the enzyme (PCon).

$P := solve(simplify(PCon), P)$ :

This results in the following steady state correlations between the various variables in terms of the substrate concentration (App) and $\epsilon$.

$$AK = \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 (App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2)}$$

$$ApK = \frac{c_2 c_3 App \epsilon k_3 P_{Total}}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2}$$

$$AppP = \frac{c_3 App P_{Total} \epsilon k_2 c_2}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2}$$

$$ApP = \frac{c_4 c_3 App k_3 P_{Total}}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2}$$

$$Ks = \frac{\epsilon^2 P_{Total} k_2 c_2 (App k_{NFB} + u_d)}{(App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2)(S u_a + u_\beta)}$$

$$K = \frac{\epsilon^2 P_{Total} k_2 c_2}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2}$$

$$P = = \frac{P_{Total} \epsilon k_2 c_2}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2}$$

Thus we have partially solved the system of equations to arrive at correlations between the steady state concentrations of the variables (particularly concentrations of variables in terms of App and $\epsilon$).

## 2. Absence of increased multi-stationarity (tristability) - Enzyme limiting regime

In this subsection of the analysis we wish to demonstrate the absence of increased multi-stationarity (enabling tristability) in the limit where the concentrations of enzymes are much less than the concentration of the substrates in the system. In such a scenario the conservation expression for the substrate concentration can be suitably modified to ignore the complexes involving the enzymes. This expression (ACon_Red = 0) is given below.

$ACon\_Red := A_{Total} - A - Ap - App$ :

Now the steady state of the system is defined by the expressions ACon_Red and KCon involving App

and $\epsilon$. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplification.

$T := simplify(numer(ACon\_Red))$

$$T := -c_1 c_2 k_1 k_2 \left(App - A_{Total}\right) \epsilon^2 - c_3 App\, k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App\, k_3 \tag{3.1}$$

$G := simplify(numer(KCon))$

$$G := \left(P_{Total} c_2 k_2 \left(App\, k_{NFB} + S u_a + u_d + u_\beta\right) \epsilon^2 - c_2 \left(c_3 \left(K_{Total} k_2 - k_3 P_{Total}\right) App + K_{Total} k_2\right) \left(S u_a + u_\beta\right) \epsilon - K_{Total} App\, c_3 c_4 k_3 \left(S u_a + u_\beta\right)\right) k_1 + k_4 c_4 c_3 App\, k_3 P_{Total} \left(S u_a + u_\beta\right) \tag{3.2}$$

Now we can solve T, to obtain an expression linking the steady state concentration of App with the ratio $\epsilon$ as shown below.

$$App := simplify(solve(T, App)) = \frac{A_{Total} \epsilon^2 k_2 c_2 k_1 c_1}{k_1 \epsilon \left(c_2 \epsilon k_2 + c_3 k_3\right) c_1 + k_4 c_4 c_3 k_3}$$

Resubstituting this into G results in the following expression.

$simplify(numer(G))$

$$c_2 \left(P_{Total} c_1 c_2 k_1 k_2 \left(S u_a + k_{NFB} A_{Total} + u_d + u_\beta\right) \epsilon^3 - k_1 c_1 \left(\left(-P_{Total} \left(A_{Total} \left(S u_a + u_\beta\right) c_2 + S u_a + u_d + u_\beta\right) k_3 + A_{Total} K_{Total} c_2 k_2 \left(S u_a + u_\beta\right)\right) c_3 + K_{Total} c_2 k_2 \left(S u_a + u_\beta\right)\right) \epsilon^2 - k_3 c_3 \left(\left(S u_a + u_\beta\right) \left(K_{Total} \left(c_4 A_{Total} + 1\right) k_1 - A_{Total} P_{Total} c_4 k_4\right) c_1 - P_{Total} c_4 k_4 \left(S u_a + u_d + u_\beta\right)\right) \epsilon - K_{Total} c_3 c_4 k_3 k_4 \left(S u_a + u_\beta\right)\right) k_1 k_2 \epsilon \tag{3.3}$$

Note that the solutions of this polynomial represents the steady state concentrations of the system. In order to have increased multi-stationarity (enabling tristability), we require five distinct solutions for the system. However this being a third order polynomial in $\epsilon$, five such distinct steady states are impossible.

**Thus we have shown that App NFB is incapable of generating increased multi-stationarity in the model, involving five positive steady states, in the enzyme limiting regime.**

# Ap negative feedback (NFB) - Direct

This document deals with results pertaining to the ordered DSP network (with common kinase and common phosphatase) with 'negative' feedback from the partially modified substrate form (Ap) by direct deactivation of the kinase.

In this document we detail the following key analytical proofs concerning the behavioral characteristics of the NFB model, drawing strict boundaries between the possibility and impossibility of multi-stationarity. These results are organized as follows,

**0. Model:** The Ap NFB model mediated by direct deactivation of the kinase by the substrate form.
**1. Solving the system of equations at steady state:** Systematically solving (algebraically) the ODEs at steady state to obtain expressions relating the steady state concentrations of the variables in terms of each other.
**2. Absence of increased multi-stationarity (tristability):** We provide an analytical proof for the absence of increased multi-stationarity (tristability) in the network.

# 0. Ap FB (negative) by activation of the kinase by the substrate form (NFB)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot S \cdot Ks - k_{NFB} \cdot Ap \cdot K - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks + Ap \cdot k_{NFB} \cdot K + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase in the absence of signal. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$

# 1. Solving the system of equations at steady state

Now we begin by solving the system of equations to obtain expressions linking the steady state concentrations of the variables, primarily to obtain expressions for the steady state concentrations of variables as a function of concentrations of Ap and a ratio $\epsilon$ (defined below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\,(Ap\,k_{NFB} + u_d)}{S\,u_a + u_\beta}$$

We similarly solve for the other variables using the same command as shown.

$$AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP := solve(dApP, ApP) :$$

$Sol := solve(\{dA, dAp\}, \{A, App\})$ :
$assign(Sol)$

We now introduce a new ratio, $\epsilon$ = K/P, and also rewrite the kinetic constants with the new nomenclature (for $k_{b1}$, $k_{b2}$, $k_{b3}$, and $k_{b4}$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$K := \epsilon \cdot P$ :
$k_{b1} := c_1 \cdot (k_1 + k_{ub1})$ : $k_{b2} := c_2 \cdot (k_2 + k_{ub2})$ : $k_{b3} := c_3 \cdot (k_3 + k_{ub3})$ : $k_{b4} := c_4 \cdot (k_4 + k_{ub4})$ :

Once this is done, we again solve for the steady state of phosphatase using the conservation expression for the enzyme (PCon).

$P := solve(simplify(PCon), P)$ :

This results in the following expressions for the steady state concentrations of the various species in terms of the steady state substrate concentration (Ap) and $\epsilon$.

$$AK = \frac{c_4\, Ap\, k_4\, P_{Total}\, k_3}{k_1\, (Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3)}$$
$$ApK = \frac{c_2\, Ap\, \epsilon\, P_{Total}\, k_3}{Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3}$$
$$AppP = \frac{Ap\, \epsilon\, k_2\, c_2\, P_{Total}}{Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3}$$
$$ApP = \frac{c_4\, Ap\, P_{Total}\, k_3}{Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3}$$
$$Ks = \frac{\epsilon\, P_{Total}\, k_3\, (Ap\, k_{NFB} + u_d)}{(Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3)\, (S\, u_a + u_\beta)}$$
$$K = \frac{\epsilon\, P_{Total}\, k_3}{Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3}$$
$$P = = \frac{P_{Total}\, k_3}{Ap\, \epsilon\, k_2\, c_2 + Ap\, c_4\, k_3 + k_3}$$

Note that when App and $\epsilon$ are positive, the other variable concentrations are positive as well. Thus we have partially solved the system of equations to arrive at expressions linking the steady state concentrations of the variables with that of Ap and $\epsilon$.

## 2. Absence of increased multi-stationarity (tristability)

We now have two conservation equations that are unsolved, namely ACon = 0 and KCon = 0. The solution set {Ap, $\epsilon$} of these two coupled expressions would thus represent the steady state solutions of the system. We note here that the denominator in these expressions are always positive, and thus their solutions can be obtained from the numerators. We rename these numerators (two coupled polynomials) as expressions T and G. Here we use the inbuilt *Maple* command *numer* to perform the simplification.

$T := simplify(numer(KCon))$ :
$G := simplify(numer(ACon))$ :

Now we rearrange the polynomial G as a function in App as shown below. We use the inbuilt *Maple* function *collect* to collect the expression in coefficients of Ap as shown.

$collect(G, Ap)$

$$\left(c_3 c_4 \left(\epsilon k_1 c_1 + c_4 k_4\right) k_3^2 + k_2 \epsilon c_2 \left(c_1 k_1 \left(c_3 + c_4\right) \epsilon + c_3 c_4 k_4\right) k_3 + c_1 c_2^2 \epsilon^3 k_1 k_2^2\right) Ap^2 + \left(c_3 \left(\epsilon^2 P_{Total} c_1 c_2 k_1 - \left(\left(-1 + \left(A_{Total} - P_{Total}\right) c_4\right) k_1 - P_{Total} c_4 k_4\right) c_1 \epsilon + c_4 k_4\right) k_3^2 - k_2 \epsilon^2 c_2 \left(-1 + \left(A_{Total} - P_{Total}\right) c_3\right) c_1 k_1 k_3\right) Ap - c_1 c_3 \epsilon k_1 k_3^2 A_{Total} \tag{3.1}$$

$collect(T, Ap)$

$$\left(\left(\left(\epsilon \left(S c_2 u_a + c_2 u_\beta + k_{NFB}\right) P_{Total} - K_{Total} c_4 \left(S u_a + u_\beta\right)\right) k_3 - \epsilon K_{Total} c_2 k_2 \left(S u_a + u_\beta\right)\right) k_1 + c_4 k_4 P_{Total} k_3 \left(S u_a + u_\beta\right)\right) Ap + \left(\epsilon \left(S u_a + u_d + u_\beta\right) P_{Total} - K_{Total} \left(S u_a + u_\beta\right)\right) k_3 k_1 \tag{3.2}$$

We now solve KCon to obtain an expression for steady state concentration of Ap in terms of the ratio $\epsilon$ (as shown below). Re-substituting this into the expression for ACon would then result in a single polynomial in $\epsilon$, whose feasible solutions (positive solutions of $\epsilon$) represent the various feasible steady states of the full system (as the concentration of the other variables can be obtained from $\epsilon$).

$Ap := solve(T, Ap)$

$$Ap := \left(\left(S \epsilon u_a P_{Total} - S u_a K_{Total} + \epsilon u_d P_{Total} + \epsilon P_{Total} u_\beta - K_{Total} u_\beta\right) k_3 k_1\right) / \left(S c_2 \epsilon k_1 k_2 u_a K_{Total} - S c_2 \epsilon k_1 k_3 u_a P_{Total} + S c_4 k_1 k_3 u_a K_{Total} - S c_4 k_3 k_4 u_a P_{Total} + c_2 \epsilon k_1 k_2 K_{Total} u_\beta - c_2 \epsilon k_1 k_3 P_{Total} u_\beta + c_4 k_1 k_3 K_{Total} u_\beta - c_4 k_3 k_4 P_{Total} u_\beta - \epsilon k_1 k_3 k_{NFB} P_{Total}\right) \tag{3.3}$$

Resubstituting these expressions back into ACon, it simplifies as follows

$simplify(numer(ACon))$

$$k_3 P_{Total} \left(P_{Total} c_1 c_2^2 k_1^2 k_2^2 \left(S u_a + u_d + u_\beta\right)^2 \epsilon^5 - \left(\left(\left(S u_a + u_\beta\right) c_2 + k_{NFB}\right) c_3 P_{Total}^2 k_3^2 - k_2 \left(\left(\left(\left(-S u_a - u_\beta\right) c_2 - k_{NFB}\right) P_{Total} + \left(A_{Total} + K_{Total}\right) \left(S u_a + u_\beta\right) c_2 + S u_a + k_{NFB} A_{Total} + u_d + u_\beta\right) c_3 + \left(-S u_a - u_\beta\right) c_2 + \left(S u_a + u_d + u_\beta\right) c_4 - k_{NFB}\right) P_{Total} k_3 + k_2^2 K_{Total} \left(S u_a + u_\beta\right) \left(1 + \left(A_{Total} - P_{Total}\right) c_3\right) c_2\right) c_1 k_1^2 \left(S u_a + u_d + u_\beta\right) c_2 \epsilon^4 + \left(\left(-\left(\left(\left(S u_a + u_\beta\right) c_2 + k_{NFB}\right) c_4 \left(S u_a + u_d + u_\beta\right) P_{Total} + \left(S u_a + u_\beta\right)^2 \left(A_{Total} - K_{Total}\right) c_2^2 - \left(\left(A_{Total} + K_{Total}\right) \left(S u_a + u_d + u_\beta\right) c_4 - S u_a - 2 k_{NFB} A_{Total} + k_{NFB} K_{Total} - u_d - u_\beta\right) \left(S u_a + u_\beta\right) c_2 - \left(\left(S u_a + u_d + u_\beta\right) c_4 - k_{NFB}\right) \left(S u_a + k_{NFB} A_{Total} + u_d + u_\beta\right)\right) c_3 P_{Total} k_3^2 + k_2 K_{Total} \left(S u_a + u_\beta\right) \left(\left(\left(\left(S u_a + u_\beta\right) c_2 + \left(2 S u_a + 2 u_d + 2 u_\beta\right) c_4 + k_{NFB}\right) P_{Total} + \left(A_{Total} - K_{Total}\right) \left(S u_a + u_\beta\right) c_2 - 2 A_{Total} \left(S u_a + u_d + u_\beta\right) c_4 - S u_a + k_{NFB} A_{Total} - u_d - u_\beta\right) c_3 + \left(S u_a + u_\beta\right) c_2 + \left(-S u_a - u_d - u_\beta\right) c_4 + k_{NFB}\right) c_2 k_3 \tag{3.4}$$

$$
\begin{aligned}
&- K_{Total}^{\,2}\, c_2^{\,2}\, c_3\, k_2^{\,2}\, \left(S\, u_a + u_\beta\right)^2\Big)\, c_1\, k_1 + k_3\, P_{Total}\, c_4\, k_4\, \left(S\, u_a + u_d + u_\beta\right)\, \Big(-2\, c_3\, c_1\, \Big(\left(S\, u_a\right.\\
&\left. + u_\beta\right)\, c_2 + \frac{k_{NFB}}{2}\Big)\, P_{Total}\, k_3 + \Big(\left(-1 + \left(A_{Total} + K_{Total} - P_{Total}\right)\, c_3\right)\, \left(S\, u_a + u_\beta\right)\, c_1\\
&+ c_3\, \left(S\, u_a + u_d + u_\beta\right)\Big)\, k_2\, c_2\Big)\Big)\, k_1\, \epsilon^3 + k_3\, \Big(c_3\, \Big(\left(c_4\, \Big(\left(S\, u_a + u_\beta\right)\, c_2 + \left(S\, u_a + u_d + u_\beta\right)\, c_4\right.\\
&\left. + k_{NFB}\right)\, P_{Total} + \left(S\, u_a + u_\beta\right)\, \left(1 + \left(A_{Total} - K_{Total}\right)\, c_4\right)\, c_2 - \left(c_4\, A_{Total} + 1\right)\, \Big(\left(S\, u_a + u_d\right.\\
&\left. + u_\beta\right)\, c_4 - k_{NFB}\Big)\Big)\, k_3 - 2\, K_{Total}\, c_2\, c_4\, k_2\, \left(S\, u_a + u_\beta\right)\Big)\, K_{Total}\, c_1\, \left(S\, u_a + u_\beta\right)\, k_1^{\,2} + \Big(\\
&-2\, \Big(\Big(\frac{\left(S\, u_a + u_d + u_\beta\right)\, c_4\, P_{Total}}{2} + \left(A_{Total} - K_{Total}\right)\, \left(S\, u_a + u_\beta\right)\, c_2\\
&- \frac{\left(A_{Total} + K_{Total}\right)\, \left(S\, u_a + u_d + u_\beta\right)\, c_4}{2} + \frac{S\, u_a}{2} + k_{NFB}\, A_{Total} - \frac{k_{NFB}\, K_{Total}}{2} + \frac{u_d}{2} + \frac{u_\beta}{2}\Big)\\
&\left(S\, u_a + u_\beta\right)\, c_1 + \frac{\left(S\, u_a + u_d + u_\beta\right)\, \Big(\left(S\, u_a + u_\beta\right)\, c_2 + \left(-S\, u_a - u_d - u_\beta\right)\, c_4 + k_{NFB}\Big)}{2}\Big)\\
&c_3\, P_{Total}\, k_3 + k_2\, \Big(\left(S\, u_a + u_\beta\right)\, \left(1 + \left(A_{Total} - K_{Total} + P_{Total}\right)\, c_3\right)\, c_1 - c_3\, \left(S\, u_a + u_d\right.\\
&\left. + u_\beta\right)\Big)\, K_{Total}\, \left(S\, u_a + u_\beta\right)\, c_2\Big)\, c_4\, k_4\, k_1 - P_{Total}^{\,2}\, c_1\, c_3\, c_4^{\,2}\, k_3\, k_4^{\,2}\, \left(S\, u_a + u_\beta\right)\, \left(S\, u_a + u_d\right.\\
&\left. + u_\beta\right)\Big)\, \epsilon^2 + c_3\, \left(S\, u_a + u_\beta\right)\, k_3^{\,2}\, c_4\, \Big(-K_{Total}^{\,2}\, c_1\, c_4\, \left(S\, u_a + u_\beta\right)\, k_1^{\,2} + \Big(\left(c_4\, P_{Total} + 1 + \left(A_{Total}\right.\right.\right.\\
&\left.\left.\left. - K_{Total}\right)\, c_4\right)\, \left(S\, u_a + u_\beta\right)\, c_1 + \left(S\, u_a + u_\beta\right)\, c_2 + \left(-S\, u_a - u_d - u_\beta\right)\, c_4 + k_{NFB}\right)\, K_{Total}\, k_4\, k_1\\
&- \Big(\left(A_{Total} - K_{Total}\right)\, \left(S\, u_a + u_\beta\right)\, c_1 + S\, u_a + u_d + u_\beta\Big)\, P_{Total}\, c_4\, k_4^{\,2}\Big)\, \epsilon\\
&+ K_{Total}\, c_3\, c_4^{\,2}\, k_3^{\,2}\, k_4^{\,2}\, \left(S\, u_a + u_\beta\right)^2\Big)
\end{aligned}
$$

This is a fifth order polynomial in $\epsilon$. Since we are looking for increased multi-stationarity (enabling tristability), we are looking for five positive solutions for $\epsilon$ as roots for this polynomial. However we can note that the leading coefficient of the polynomial in $\epsilon$ and the constant are always positive (irrespective of parameter value) thus implying by a simple extension of Vieta's formulas for roots of a polynomial, that the system cannot accommodate five positive roots (or that one root is necessarily negative and real).

**Thus we have shown that Ap NFB is incapable of generating increased multi-stationarity in the model involving five positive steady states.**

# Ordered DSP with unsaturated limit of enzyme action (no complexes)
# Impact of feedback on network behavior (Multi-stationarity)

This worksheet explores the effect of feedback regulation on the propensity for multi-stationarity in the ordered DSP (with common kinase and common phosphatase), when the phosphorylation and dephosphorylation are carried out in the unsaturated limit of enzyme action. In order to study this further, we first consider the basic isolated model without any feedback, and show that this system is necessarily incapable of multi-stationarity. Further on we study systematically the effect of each type of feedback on the models' capacity to generate multi-stationarity. This analysis is carried out for both when the feedback is effected directly on the kinase through activation and inactivation, and when the feedback is mediated through a covalent modification tier. The organization of these results (into subsections) and the emerging synthesis is expanded on below.

**0. Model:** The ordered DSP model (all complexes are absent due to unsaturated limit of enzyme action) without any feedback (called model M1)
**1. Absence of multistability in M1:** Analytic proof establishing that M1 model has only one possible feasible steady state (ruling out multi-stationarity).

**2. Feedback from App (Activating kinase - Positive feedback):** We show how the App positive feedback on M1 is able to generate multi-stationarity, (Multi-stationarity is observed both when a covalent modification tier mediates the positive feedback and when the feedback is directly effected).
**3. Feedback from App (Deactivating kinase - Negative feedback):** We prove that App negative feedback on M1 is unable to generate multi-stationarity (both with and without a tier mediating the feedback).
**4. Feedback from Ap (Activating kinase - Positive feedback):** We prove that Ap 'positive' feedback on M1 is unable to generate multi-stationarity (both with and without a tier mediating the feedback).
**5. Feedback from Ap (Deactivating kinase - Negative feedback):** We prove that Ap 'negative' feedback on M1 is unable to generate multi-stationarity (both with and without a tier mediating the feedback).

A summary of these results is tabulated below.

**`able 1: Summary: Capacity of ordered DSP (no complex) with feedback to generate multi-stationarity**

| Model | Tier | No Tier |
|---|---|---|
| Isolated ordered DSP (M1) | **No** | |
| App Positive Feedback | **Yes** | **Yes** |
| App Negative Feedback | **No** | **No** |
| Ap Positive Feedback | **No** | **No** |
| Ap Negative Feedback | **No** | **No** |

These results indicate that while there are similarities between analogous feedback effected by the maximally modified substrate and the partially modified substrate, App feedback registers much more strongly (as positive or negative feedback) and in the case of positive feedback is capable of generating multi-stationarity in the network when the other feedbacks are incapable of it. This inability to generate multi-stationarity is the case even when the feedback is mediated by the additional covalent modification tier (bringing additional nonlinearity to the system). This allows us to conclude that the emerging nature of the feedback at the systems level is a careful confluence of both the biochemistry of the multisite modification system and the nature of the feedback (i.e. positive or negative mediated directly or through a tier).

**Instructions**
Each subsection apart from 0. & 1. is self contained - hence only run the necessary sections modularly (ignore if you are accessing this on the PDF).

## 0. Model (M1)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks + u_d \cdot K :$$
$$dK := u_a \cdot Ks - u_d \cdot K :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model and consequently P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$

## 1. Absence of multi-stationarity in M1

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain the steady state expression of Ap as a function of parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{u_a}$$

We similarly solve for the other variables using the same command.

$$Sol := solve(\{dA, dApp\}, \{A, App\}) :$$
$$assign(Sol)$$

With this we have solved all ODEs at steady state. We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables define the steady state concentrations of the variables and thus the system. We now solve these too to obtain a expression for Ap.

$$K := simplify(solve(KCon, K)) :$$

$Ap = simplify(solve(simplify(numer(ACon)), Ap))$

$$Ap = \frac{A_{Total}\, K_{Total}\, u_a\, k_1 \left(u_a + u_d\right) P\, k_3}{P \left(u_a + u_d\right) \left(\left(P\, k_4 + K_{Total}\, k_1\right) u_a + P\, k_4\, u_d\right) k_3 + k_2\, K_{Total}^{\,2}\, u_a^{\,2}\, k_1} \tag{2.1}$$

Having solved for the steady state concentration of every variable in terms of parameters, we can find in the equation above for the steady state concentration of Ap that Ap assumes only one fixed value (determined by the parameters). Similarly the other parameters which can be written in terms of the parameters and Ap, also take fixed values - indicating that the system can, for any given choice of parameters, attain only one steady state, i.e. the system is incapable of multi-stationarity.

## 2. Feedback from App (Activating kinase - Positive feedback)

### No Tier/Cascade in App feedback - Presence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$
$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$
$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$
$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot App + u_d \cdot K :$
$dK := +u_a \cdot Ks + k_{PFB} \cdot Ks \cdot App - u_d \cdot K :$

The model is also associated with conservation equations which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation.

$ACon := A + Ap + App - A_{Total} :$
$KCon := K + Ks - K_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d\, K}{App\, k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$Sol := solve(\{dA, dApp\}, \{A, App\}) :$
$assign(Sol)$

With this we have solved all ODEs at steady state. We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now

define the steady state concentrations of the variables and thus the system. We now solve ACon to obtain expressions for steady state concentration of Ap in terms of the parameters and K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system. Note that a positive solution for K, would imply a positive value for the other variables as well.

$$T := simplify(numer(KCon))$$

$$T := k_1\, k_2\, \left(k_{PFB}\, A_{Total} + u_a + u_d\right) K^3 - k_1\, \left(-P\, \left(u_a + u_d\right) k_3 + K_{Total}\, k_2\, \left(k_{PFB}\, A_{Total} + u_a\right)\right) K^2 - P\, k_3\, \left(K_{Total}\, k_1\, u_a - P\, k_4\, \left(u_a + u_d\right)\right) K - P^2\, k_3\, k_4\, u_a\, K_{Total} \quad \textbf{(3.1.1)}$$

Now here we show that the following parameter set admits to three feasible (positive, real) solutions for K for the above equation. Implying that there are three steady states for the system (stable or unstable) showing the presence of multi-stationarity in the model.

$$k_1 := 95.215634132390290 : k_2 := 43.440210403731730 : k_3 := 81.748089333411130 : k_4 := 35.133700810505160 :$$
$$A_{Total} := 38.629906119332080 : K_{Total} := 8.572848678312022 : P := 31.547450162963287 :$$
$$u_a := 2.029597076788661 : u_d := 38.849083119075736 :$$
$$k_{PFB} := 49.129181844316160 :$$

$$Sol := solve(\{T\}, \{K\}, useassumptions) \text{ assuming } K > 0$$
$$Sol := \{K = 0.6212639025\}, \{K = 2.445144802\}, \{K = 4.082872580\} \quad \textbf{(3.1.2)}$$

**This shows that there are three distinct steady states that the system admits - i.e. the system admits to multi-stationarity. Thus direct positive feedback from App generates multi-stationarity in the system.**

## Tiered/cascaded App feedback - Presence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot X + u_d \cdot K :$$
$$dK := +u_a \cdot Ks + k_{PFB} \cdot Ks \cdot X - u_d \cdot K :$$

$$dXs := -d_a \cdot Xs \cdot App + d_d \cdot X :$$
$$dX := +d_a \cdot Xs \cdot App - d_d \cdot X :$$

The model is also associated with conservation conditions which are described below. Here we store

the conservation expressions as ACon, XCon and KCon for the substrate, the covalent modification tier mediating feedback and the kinase. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation.

$$XCon := X + Xs - X_{Total}:$$
$$ACon := A + Ap + App - A_{Total}:$$
$$KCon := K + Ks - K_{Total}:$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{X\, d_d}{App\, d_a}$$

We similarly solve for the other variables using the same command as shown.

$$Ks := solve(dKs, Ks):$$
$$Sol := solve(\{dA, dApp\}, \{A, App\}):$$
$$assign(Sol)$$

With this we have solved all ODEs at steady state. We now have three expresssions, ACon, XCon and KCon - the conservation of the total substrate, the covalent modification layer mediating the feedback and kinase concentration, whose solutions for these variables define the steady state concentrations of the system. We now solve ACon and XCon to obtain an expression for steady state concentration of Ap and X in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$
$$X := solve(XCon, X) = \frac{X_{Total}\, A_{Total}\, K^2\, k_1\, k_2\, d_a}{K^2\, d_a\, k_1\, k_2\, A_{Total} + K^2\, d_d\, k_1\, k_2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4}$$

Thus, now the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables. Note that a positive solution for K, would imply a positive value for the other variables as well.

$$T := simplify(numer(KCon))$$

$$T := k_1\, k_2 \left( \left( u_d + u_a \right) d_d + A_{Total}\, d_a \left( k_{PFB}\, X_{Total} + u_a + u_d \right) \right) K^3 + k_1 \left( \left( -K_{Total}\, u_a\, k_2 + P\, k_3 \left( u_d + u_a \right) \right) d_d - A_{Total}\, K_{Total}\, d_a\, k_2 \left( k_{PFB}\, X_{Total} + u_a \right) \right) K^2 + P \left( -K_{Total}\, u_a\, k_1 + P\, k_4 \left( u_d + u_a \right) \right) k_3\, d_d\, K - P^2\, d_d\, k_3\, k_4\, u_a\, K_{Total} \quad \textbf{(3.2.1)}$$

Now here we show that the following parameter set admits to three solutions for K for the above equation. Implying that there are three steady states for the system (stable or unstable) showing the presence of multi-stationarity in the model.

$$k_1 := 2.870465518053682 : k_2 := 47.904356772646070 : k_3 := 92.022955161014080 : k_4 := 89.103855944507930 :$$
$$A_{Total} := 16.330778745139334 : K_{Total} := 15.435538321499367 : P := 6.327801916805306 : X_{Total} := 19.785838468317202 :$$
$$u_a := 0.462761020511521 : u_d := 43.841634147243570 : d_a := 33.688626061884910 : d_d := 25.091762265929410 :$$
$$k_{PFB} := 7.768970706098548 :$$

$solve(\{T\}, \{K\}, useassumptions)$ assuming $K > 0$

$\{K = 0.1763214795\}, \{K = 2.396296928\}, \{K = 9.202938809\}$ **(3.2.2)**

**This shows that there are three distinct steady states that the system admits - i.e. the system admits to multi-stationarity. Thus positive feedback from App mediated through a covalent modification tier generates multi-stationarity in the system.**

# 3. Feedback from App (Deactivating kinase - Negative feedback)

## No tier/cascade in App feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$
$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$
$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$

$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot App + u_d \cdot K :$
$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot App - u_d \cdot K :$

The model is also associated with conservation conditions for the substrate and the kinase which are described below. Here we store these conservation expressions as ACon and KCon respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation.

$ACon := A + Ap + App - A_{Total} :$
$KCon := K + Ks - K_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\left(App\, k_{NFB} + u_d\right)}{u_a}$$

We similarly solve for the other variables using the same command as shown below.

$Sol := solve(\{dA, dApp\}, \{A, App\}) :$
$assign(Sol)$

With this we have solved all ODEs at steady state. We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. We now solve ACon to

obtain expressions for steady state concentration of Ap in terms of the parameters and K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system.

$$T := simplify(numer(KCon))$$

$$T := k_1\, k_2 \left(k_{NFB}\, A_{Total} + u_a + u_d\right) K^3 - \left(-P \left(u_a + u_d\right) k_3 + K_{Total}\, k_2\, u_a\right) k_1\, K^2 - P\, k_3 \left(K_{Total}\, k_1\, u_a - P\, k_4 \left(u_a + u_d\right)\right) K - P^2\, k_3\, k_4\, u_a\, K_{Total} \qquad \textbf{(4.1.1)}$$

Now if we try to evaluate the system without the feedback (using the *Maple* command *eval*), as shown below - we can make a number of observations.

$$simplify\left(eval\left(T, k_{NFB} = 0\right)\right) = \left(K^2\, k_1\, k_2 + K\, k_1\, P\, k_3 + P^2\, k_3\, k_4\right) \left(\left(u_a + u_d\right) K - K_{Total}\, u_a\right)$$

Observation:

1. Without the negative feedback the third order polynomial in K factorizes and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial (4.1.1) is independent of the feedback term.
3. If the cubic polynomial (4.1.1) were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback does not change these coefficients, and that the leading coefficient and the constant term are positive and negative for all parameter values.

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the two negative roots (real or complex conjugates with negative real parts, when feedback was zero) would have to cross the imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity (Since the polynomial precludes a zero root, because of the constant term of the polynomial always being negative and non-zero).

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a negative feedback strength ($k_{NFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2 should be positive ($C1 > 0$) and negative ($C2 < 0$) at such a $k_{NFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier and since the terms C1 and C2 are independent of $k_{NFB}$).
3. At such $k_{NFB}$, $C0*C3 - C1*C2 = 0$

We now isolate the coefficients of the polynomial (using the *Maple* command *coeff*) to evaluate this expression and we verify using the inbuilt *Maple* command *verify* that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-k_3\, k_1 \left(k_3\, k_4 \left(u_a + u_d\right)^2 P^2 - K_{Total} \left(k_1 \left(u_a + u_d\right) k_3 - k_2\, k_4\, k_{NFB}\, A_{Total}\right) u_a\, P \qquad \textbf{(4.1.2)}$$

$+ K_{Total}^{2}\, k_1\, k_2\, u_a^{2}\Big)\, P$

$verify(C0 \cdot C3 - C1 \cdot C2, 0, less\_than)$ assuming $k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{NFB} > 0, A_{Total} > 0, K_{Total} > 0, P_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$

$$true \qquad \textbf{(4.1.3)}$$

This contradicts our assumption earlier. Thus a conjugate solution of the cubic are necessarily bound to the negative half plane.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that negative feedback from App cannot generate multi-stationarity in the model.**

## Tiered/cascaded App feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$
$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$
$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$

$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot X + u_d \cdot K :$
$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot X - u_d \cdot K :$

$dXs := -d_a \cdot Xs \cdot App + d_d \cdot X :$
$dX := +d_a \cdot Xs \cdot App - d_d \cdot X :$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, XCon and KCon for the substrate, the covalent modification tier mediating feedback and the kinase. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation.

$XCon := X + Xs - X_{Total} :$
$ACon := A + Ap + App - A_{Total} :$
$KCon := K + Ks - K_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{X\, d_d}{App\, d_a}$$

We similarly solve for the other variables using the same command as shown.

$Ks := solve(dKs, Ks) :$
$Sol := solve(\{dA, dApp\}, \{A, App\}) :$
$assign(Sol)$

With this we have solved all ODEs at steady state. We now have three expresssions, ACon, XCon and KCon - the conservation of the total substrate, the covalent modification layer mediating the feedback and kinase concentration, that whose solutions for these variables define the steady state concentrations of the system. We now solve ACon and XCon to obtain an expression for steady state concentration of Ap and X in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$

$$X := solve(XCon, X) = \frac{X_{Total}\, A_{Total}\, K^2\, k_1\, k_2\, d_a}{K^2\, d_a\, k_1\, k_2\, A_{Total} + K^2\, d_d\, k_1\, k_2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables.

$$T := simplify(numer(KCon))$$

$$T := k_2\, k_1 \left(\left(u_d + u_a\right) d_d + A_{Total}\, d_a \left(k_{NFB}\, X_{Total} + u_a + u_d\right)\right) K^3 + k_1 \left(\left(\left(P\, k_3 - k_2\, K_{Total}\right) u_a + u_d\, P\, k_3\right) d_d - A_{Total}\, K_{Total}\, d_a\, k_2\, u_a\right) K^2 + d_d\, P \left(-K_{Total}\, u_a\, k_1 + P\, k_4 \left(u_d + u_a\right)\right) k_3\, K - P^2\, d_d\, k_3\, k_4\, u_a\, K_{Total} \quad \textbf{(4.2.1)}$$

$$simplify\left(eval\left(simplify(T), k_{NFB} = 0\right)\right)$$

$$\left(\left(u_d + u_a\right) K - K_{Total}\, u_a\right) \left(k_1\, k_2 \left(A_{Total}\, d_a + d_d\right) K^2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4\right) \quad \textbf{(4.2.2)}$$

Now if we try to evaluate the system without the feedback (using the *Maple* command *eval*), as shown below - we can make a number of observations.

$$simplify\left(eval\left(T, k_{NFB} = 0\right)\right) =$$
$$\left(\left(u_d + u_a\right) K - K_{Total}\, u_a\right) \left(k_1\, k_2 \left(A_{Total}\, d_a + d_d\right) K^2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4\right)$$

Observation:

1. Without the negative feedback the third order polynomial in K factorizes and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial (4.2.2) is independent of the feedback term.
3. If the cubic polynomial (4.2.2) were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback does not change these coefficients, and that the leading coefficient and the constant term are positive and negative for all parameter values.

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the two negative roots (real or complex conjugates with negative real parts, when feedback was zero) would have to cross the imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity (Since the polynomial precludes a zero root, because of the constant term of the polynomial always being negative and non-zero).

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a negative feedback strength ($k_{NFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2

should be positive (C1 > 0) and negative (C2 < 0) at such a $k_{NFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier and since the terms C1 and C2 are independent of $k_{NFB}$).
3. At such $k_{NFB}$, C0*C3 - C1*C2 = 0

We now isolate the coefficients of the polynomial (using the *Maple* command *coeff*) to evaluate this expression and we verify using the inbuilt *Maple* command *verify* that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-k_3 k_1 d_d \left( \left( \left( P^2 k_3 k_4 - P k_1 k_3 K_{Total} + k_1 k_2 K_{Total}^2 \right) u_a^2 + 2 k_3 P \left( P k_4 - \frac{K_{Total} k_1}{2} \right) u_d u_a \right. \right. \quad \textbf{(4.2.3)}$$
$$\left. \left. + P^2 k_3 k_4 u_d^2 \right) d_d + A_{Total} K_{Total} d_a k_2 u_a \left( P k_4 k_{NFB} X_{Total} + K_{Total} u_a k_1 \right) \right) P$$

$$verify(simplify(C0 \cdot C3 - C1 \cdot C2), 0, less\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{NFB} > 0, A_{Total} > 0, K_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$$
$$true \quad \textbf{(4.2.4)}$$

This contradicts our assumption earlier. Thus conjugate solutions of the cubic are necessarily bound to the negative parameter space.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that negative feedback from App cannot generate multi-stationarity in the model.**

## 4. Feedback from Ap (Activating kinase - Positive feedback analogue)

### No tier/cascade in Ap feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot Ap + u_d \cdot K :$$
$$dK := +u_a \cdot Ks + k_{PFB} \cdot Ks \cdot Ap - u_d \cdot K :$$

The model is also associated with conservation conditions which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase. Note that there is no phosphatase conservation equation. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{Ap\, k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command as shown.

$$Sol := solve(\{dA, dApp\}, \{A, App\}) :$$
$$assign(Sol)$$

With this we have solved all ODEs at steady state. We now have two expresssions, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system. We now solve ACon to obtain an expression for steady state concentration of Ap the variables in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total} K k_1 P k_3}{K^2 k_1 k_2 + K P k_1 k_3 + P^2 k_3 k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables.

$$T := simplify(numer(KCon))$$

$$T := k_1 k_2 (u_a + u_d) K^3 + (P (k_{PFB} A_{Total} + u_a + u_d) k_3 - K_{Total} k_2 u_a) k_1 K^2 - k_3 (K_{Total} (k_{PFB} A_{Total} + u_a) k_1 - P k_4 (u_a + u_d)) P K - P^2 k_3 k_4 u_a K_{Total} \quad \textbf{(5.1.1)}$$

$$simplify(eval(T, k_{PFB} = 0))$$

$$((u_a + u_d) K - u_a K_{Total}) (K^2 k_1 k_2 + K k_1 P k_3 + P^2 k_3 k_4) \quad \textbf{(5.1.2)}$$

Observation:

1. Without the positive feedback the third order polynomial in K simplifies and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial is independent of the feedback term.
3. If the cubic polynomial were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback increases and decreases the coeffecients of the second and first exponents respectively (works against satisfying Descartes rule of signs).

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the negative roots (real or complex conjugates with negative real parts) would have to cross the imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity.

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a positive feedback strength ($k_{PFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2 should be positive (C1 > 0) and negative (C2 < 0) at such a k

$_{PFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier). Even though the coeffecients are affected by the feedback strength, this condition would still have to hold as C2 and C1 would have to be negative and positive at 0 feedback in order for them to satisfy negativity and positivity conditions at higher feedback strengths, since feedback strength increases and decreases the respective coeffecients.
3. At such $k_{PFB}$, C0*C3 - C1*C2 = 0.

We now isolate the coefficients of the polynomial above (using the *Maple* command *coeff*) to evaluate this expression and we verify using the inbuilt *Maple* command *verify* that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-k_3 k_1 P \left( \left( P k_4 \left( u_a + u_d \right) - K_{Total} \left( k_{PFB} A_{Total} + u_a \right) k_1 \right) P \left( k_{PFB} A_{Total} + u_a + u_d \right) k_3 + K_{Total}^2 k_1 k_2 u_a \left( k_{PFB} A_{Total} + u_a \right) \right) \quad \textbf{(5.1.3)}$$

$$verify(simplify(C0 \cdot C3 - C1 \cdot C2), 0, less\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{PFB} > 0, A_{Total} > 0, K_{Total} > 0, P_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$$

$$true \quad \textbf{(5.1.4)}$$

This contradicts our assumption earlier. Thus conjugate solutions of the cubic are necessarily bound to the negative parameter space.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that positive feedback (analogue) from Ap cannot generate multi-stationarity in the model.**

## Tiered/cascaded Ap feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot X + u_d \cdot K :$$
$$dK := +u_a \cdot Ks + k_{PFB} \cdot Ks \cdot X - u_d \cdot K :$$

$$dXs := -d_a \cdot Xs \cdot Ap + d_d \cdot X :$$
$$dX := d_a \cdot Xs \cdot Ap - d_d \cdot X :$$

The model is also associated with conservation conditions which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase. Note that there is no phosphatase conservation equation. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$XCon := X + Xs - X_{Total} :$$

$ACon := A + Ap + App - A_{Total}$ :
$KCon := K + Ks - K_{Total}$ :

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{X\, d_d}{Ap\, d_a}$$

We similarly solve for the other variables using the same command as shown.

$Ks := solve(dKs, Ks)$ :
$Sol := solve(\{dA, dApp\}, \{A, App\})$ :
$assign(Sol)$

With this we have solved all ODEs at steady state. We now have three expresssions, ACon, XCon and KCon - the conservation of the total substrate, covalent tier mediating feedback and kinase concentration, whose solution for the variables define the steady state of the system. We now solve ACon and XCon to obtain an expression for steady state concentration of Ap and X the variables in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$

$$X := solve(XCon, X) = \frac{X_{Total}\, A_{Total}\, K\, k_1\, P\, k_3\, d_a}{K\, P\, d_a\, k_1\, k_3\, A_{Total} + K^2\, d_d\, k_1\, k_2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables.

$T := simplify(numer(KCon))$

$$T := d_d\, k_1\, k_2\, (u_d + u_a)\, K^3 + k_1\, (((u_d + u_a)\, d_d + A_{Total}\, d_a\, (k_{PFB}\, X_{Total} + u_a + u_d))\, P\, k_3 - K_{Total}\, d_d\, k_2\, u_a)\, K^2 + P\, k_3\, (-(u_a\, d_d + A_{Total}\, d_a\, (k_{PFB}\, X_{Total} + u_a))\, K_{Total}\, k_1 + P\, d_d\, k_4\, (u_d + u_a))\, K - P^2\, d_d\, k_3\, k_4\, u_a\, K_{Total} \quad \textbf{(5.2.1)}$$

$simplify(eval(T, k_{PFB} = 0))$

$$(K^2\, d_d\, k_1\, k_2 + P\, k_1\, k_3\, (A_{Total}\, d_a + d_d)\, K + P^2\, d_d\, k_3\, k_4)\, ((u_d + u_a)\, K - K_{Total}\, u_a) \quad \textbf{(5.2.2)}$$

Observation:

1. Without the positive feedback the third order polynomial in K simplifies and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial is independent of the feedback term.
3. If the cubic polynomial were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback increases and decreases the coeffecients of the second and first exponents respectively (works against satisfying Descartes rule of signs).

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the negative roots (real or complex conjugates with negative real parts) would have to cross the imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity.

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a positive feedback strength ($k_{PFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2 should be positive ($C1 > 0$) and negative ($C2 < 0$) at such a $k_{PFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier). Even though the coeffecients are affected by the feedback strength, this condition would still have to hold as C2 and C1 would have to be negative and positive at 0 feedback in order for them to satisfy negativity and positivity conditions at higher feedback strengths, since feedback strength increases and decreases the respective coeffecients.
3. At such $k_{PFB}$, $C0*C3 - C1*C2 = 0$.

We now isolate the coefficients of the polynomial above (using the *Maple* command *coeff*) to evaluate this expression and we verify using the inbuilt *Maple* command *verify* that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-P\,k_3\,k_1\left(P\left(-\left(u_a\,d_d + A_{Total}\,d_a\left(k_{PFB}\,X_{Total} + u_a\right)\right)K_{Total}\,k_1 + P\,d_d\,k_4\left(u_d + u_a\right)\right)\left(\left(u_d + u_a\right)d_d + A_{Total}\,d_a\left(k_{PFB}\,X_{Total} + u_a + u_d\right)\right)k_3 + K_{Total}{}^2\,k_1\,u_a\,k_2\,d_d\left(u_a\,d_d + A_{Total}\,d_a\left(k_{PFB}\,X_{Total} + u_a\right)\right)\right) \quad \textbf{(5.2.3)}$$

$$verify(simplify(C0 \cdot C3 - C1 \cdot C2), 0, less\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{PFB} > 0, A_{Total} > 0, K_{Total} > 0, P_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$$

$$true \quad \textbf{(5.2.4)}$$

This contradicts our assumption earlier. Thus conjugate solutions of the cubic are necessarily bound to the negative parameter space.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that positive feedback (analogue) from Ap cannot generate multi-stationarity in the model.**

## 5. Feedback from Ap (Deactivating kinase - Negative feedback analogue)

### No Tier/Cascade in Ap feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot Ap + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot Ap - u_d \cdot K :$$

The model is also associated with conservation conditions which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase. Note that there is no phosphatase conservation equation. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\left(Ap\, k_{NFB} + u_d\right)}{u_a}$$

We similarly solve for the other variables using the same command as shown.

$$Sol := solve(\{dA, dApp\}, \{A, App\}) :$$
$$assign(Sol)$$

With this we have solved all ODEs at steady state. We have two expresssions, ACon and KCon - the conservation of the total substrate and kinase concentration,whose solution for the variables define the steady state of the system. We now solve ACon to obtain an expression for steady state concentration of Ap the variables in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2 k_1 k_2 + K P k_1 k_3 + P^2 k_3 k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables.

$$T := simplify(numer(KCon))$$

$$T := k_1 k_2 \left(u_a + u_d\right) K^3 + k_1 \left(P \left(k_{NFB} A_{Total} + u_a + u_d\right) k_3 - K_{Total} k_2 u_a\right) K^2 \qquad \textbf{(6.1.1)}$$
$$- k_3 \left(K_{Total} k_1 u_a - P k_4 \left(u_a + u_d\right)\right) P K - P^2 k_3 k_4 u_a K_{Total}$$

$$simplify\left(eval\left(T, k_{NFB} = 0\right)\right) = \left(\left(u_a + u_d\right) K - K_{Total} u_a\right) \left(K^2 k_1 k_2 + K k_1 P k_3 + P^2 k_3 k_4\right)$$

Observation:

1. Without the negative feedback the third order polynomial in K simplifies and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial is independent of the feedback term.
3. If the cubic polynomial were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback does not change the coeffecient of the first exponent, increases the value of the second exponent and that the leading coefficient and the constant term are positive and negative for all parameter values.

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the negative roots (real or complex conjugates with negative real parts) would have to cross the

imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity.

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a negative feedback strength ($k_{NFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2 should be positive ($C1 > 0$) and negative ($C2 < 0$) at such a $k_{NFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier and since C1 is independent of $k_{NFB}$). Even though C2 is affected by the feedback strength, this condition would still have to hold as C2 would have to be negative at 0 feedback in order for it to satisfy negativity conditions at higher feedback strengths, since feedback strength increases C2.
3. At such $k_{NFB}$, $C0*C3 - C1*C2 = 0$.

We now isolate the coefficients of the polynomial above (using the *Maple* command *coeff*) to evaluate this expression and we verify using the inbuilt *Maple* command *verify* that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-P\left(\left(k_{NFB} A_{Total} + u_a + u_d\right) P \left(P k_4 \left(u_a + u_d\right) - K_{Total} k_1 u_a\right) k_3 + K_{Total}^2 k_1 k_2 u_a^2\right) k_3 k_1 \quad \textbf{(6.1.2)}$$

$$verify(simplify(C0 \cdot C3 - C1 \cdot C2), 0, less\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{NFB} > 0, A_{Total} > 0, K_{Total} > 0, P_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$$

$$true \quad \textbf{(6.1.3)}$$

This contradicts our assumption earlier. Thus conjugate solutions of the cubic are necessarily bound to the negative parameter space.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that negative feedback (analogue) from Ap cannot generate multi-stationarity in the model.**

## Tiered/cascaded Ap feedback - Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := k_1 \cdot A \cdot K + k_3 \cdot App \cdot P - k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot Ap \cdot K - k_3 \cdot App \cdot P :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot X + u_d \cdot K :$$
$$dK := u_a \cdot Ks - k_{NFB} \cdot K \cdot X - u_d \cdot K :$$

$$dXs := -d_a \cdot Xs \cdot Ap + d_d \cdot X :$$

$dX := d_a \cdot Xs \cdot Ap - d_d \cdot X$ :

The model is also associated with conservation conditions which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase. Note that there is no phosphatase conservation equation. Each of these expressions is always equal to zero (both in the transient and at steady state).

$XCon := X + Xs - X_{Total}$ :
$ACon := A + Ap + App - A_{Total}$ :
$KCon := K + Ks - K_{Total}$ :

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of other steady state concentration of K and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{X\, d_d}{Ap\, d_a}$$

We similarly solve for the other variables using the same command as shown.

$Ks := solve(dKs, Ks)$ :
$Sol := solve(\{dA, dApp\}, \{A, App\})$ :
$assign(Sol)$

With this we have solved all ODEs at steady state. We now have three expresssions, ACon, XCon and KCon - the conservation of the total substrate and kinase concentration, whose solution for the variables define the steady state of the system. We now solve ACon and XCon to obtain an expression for steady state concentration of Ap the variables in terms of the parameters and steady state concentration of K.

$$Ap := simplify(solve(numer(ACon), Ap)) = \frac{A_{Total}\, K\, k_1\, P\, k_3}{K^2\, k_1\, k_2 + K\, P\, k_1\, k_3 + P^2\, k_3\, k_4}$$

$$X := solve(XCon, X) = \frac{X_{Total}\, A_{Total}\, K\, k_1\, P\, k_3\, d_a}{K\, P\, d_a\, k_1\, k_3\, A_{Total} + K^2\, d_d\, k_1\, k_2 + K\, P\, d_d\, k_1\, k_3 + P^2\, d_d\, k_3\, k_4}$$

Thus the only expresssion left is KCon, the positive feasible solutions of which for K would denote the steady state concentrations of the system and other variables.

$T := simplify(numer(KCon))$

$$T := d_d\, k_1\, k_2\, (u_d + u_a)\, K^3 + (((u_d + u_a)\, d_d + A_{Total}\, d_a\, (k_{NFB}\, X_{Total} + u_a + u_d))\, P\, k_3 - K_{Total}\, d_d\, k_2\, u_a)\, k_1\, K^2 + k_3\, (-K_{Total}\, u_a\, (d_a\, A_{Total} + d_d)\, k_1 + P\, d_d\, k_4\, (u_d + u_a))\, P\, K - P^2\, d_d\, k_3\, k_4\, u_a\, K_{Total} \quad \textbf{(6.2.1)}$$

$simplify(eval(T, k_{NFB} = 0))$

$$(K^2\, d_d\, k_1\, k_2 + P\, k_1\, k_3\, (d_a\, A_{Total} + d_d)\, K + P^2\, d_d\, k_3\, k_4)\, ((u_d + u_a)\, K - K_{Total}\, u_a) \quad \textbf{(6.2.2)}$$

Observation:

1. Without the negative feedback the third order polynomial in K simplifies and the resulting equation has only one positive real root for K and two roots for K that are either negative real or complex conjugates with negative real parts.
2. The constant term in this cubic polynomial is independent of the feedback term.

3. If the cubic polynomial were to have three positive solutions then, the coefficient of the second and first exponent of K in the polynomial would need to be negative and positive respectively even at 0 feedback (by Descartes rule of signs). Note that increasing feedback does not change the coeffecient of the first exponent, increases the value of the second exponent and that the leading coefficient and the constant term are positive and negative for all parameter values.

These observations imply that as the feedback strength changes, if multi-stationarity were to exist, then the negative roots (real or complex conjugates with negative real parts) would have to cross the imaginary axis at some feedback strength to ensure that there can be two/three positive solutions for the polynomial implying multi-stationarity.

We now show that this cannot be the case by contradiction.

If we assume that multi-stationarity exists for some feedback strength.

1. There exists a negative feedback strength ($k_{NFB}$) where the cubic polynomial mentioned earlier has two purely imaginary solutions.
2. If we were to rewrite the cubic polynomial T as $C3*K^3 + C2*K^2 + C1*K^1 + C0 = 0$, C1 and C2 should be positive ($C1 > 0$) and negative ($C2 < 0$) at such a $k_{NFB}$ (this requirement arises from the Descartes rule of signs mentioned earlier and since C1 is independent of $k_{NFB}$). Even though C2 is affected by the feedback strength, this condition would still have to hold as C2 would have to be negative at 0 feedback in order for it to satisfy negativity conditions at higher feedback strengths, since feedback strength increases C2.
3. At such $k_{NFB}$, $C0*C3 - C1*C2 = 0$.

We now isolate the coefficients of the polynomial above (using the *Maple* command coeff) to evaluate this expression and we verify using the inbuilt *Maple* command verify that this can never be the case as C0*C3 - C1*C2 is strictly negative when C1>0.

$$C0 := coeff(T, K, 0) : C1 := coeff(T, K, 1) : C2 := coeff(T, K, 2) : C3 := coeff(T, K, 3) :$$

$$simplify(C0 \cdot C3 - C1 \cdot C2)$$

$$-k_1 P k_3 \left( \left( (u_d + u_a) d_d + A_{Total} d_a \left( k_{NFB} X_{Total} + u_a + u_d \right) \right) P \left( -K_{Total} u_a \left( A_{Total} d_a + d_d \right) k_1 + P d_d k_4 \left( u_d + u_a \right) \right) k_3 + K_{Total}^2 d_d k_1 k_2 u_a^2 \left( A_{Total} d_a + d_d \right) \right) \quad \textbf{(6.2.3)}$$

$$verify(simplify(C0 \cdot C3 - C1 \cdot C2), 0, less\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, u_a > 0, u_d > 0, k_{NFB} > 0, A_{Total} > 0, K_{Total} > 0, P_{Total} > 0, P > 0, C1 > 0, X_{Total} > 0, d_a > 0, d_d > 0$$

$$true \quad \textbf{(6.2.4)}$$

This contradicts our assumption earlier. Thus conjugate solutions of the cubic are necessarily bound to the negative parameter space.

**Hence multi-stationarity with three positive real steady states cannot exist indicating that negative feedback (analogue) from Ap cannot generate multi-stationarity in the model.**

# Ordered DSP with all but one modification step in the unsaturated limit of enzyme action
# Impact of feedback on network behavior (Multi-stationarity)

This worksheet explores the effect of feedback regulation on the propensity for multi-stationarity in the ordered DSP (with common kinase and common phosphatase), when all except one modification step operates in the unsaturated limit of enzymatic action. We first consider such a basic isolated model without any network level regulation and study its potential for multi-stationarity. Further on we systematically study the effect of each feedback on the model's capacity to generate/expand multi-stationarity. The organization of these results and the emerging synthesis is expanded on below.

**Note:** This analysis is carried out for only when the feedback is effected directly on the kinase level through activation and inactivation by the substrate form.

**0. Models and proof for presence/absence of multi-stationarity:** The ordered DSP model (all except one modification step operating in the unsaturated limit of enzyme action) without any feedback (models M3, M4, M5, M6, with AK, ApK, AppP, ApP complex only being present respectively)

**1. Positive feedback from App:** Presence of multi-stationarity in the models M4 and M6
**2. Positive feedback from Ap:** Absence and presence of multi-stationarity in the models M4 and M6 respectively
**3. Negative feedback from App:** Absence of multi-stationarity in the models M4 and M6
**4. Negative feedback from Ap:** Absence of multi-stationarity in the models M4 and M6

The App feedback genearting multi-stationarity in all these models is to be expected as App feedback was able to generate multi-stationarity in the model with no complexes (see section 2.4.1). A summary of these results is also tabulated below as shown.

**1: Summary: Capacity of ordered DSP (with single complex) with feedback to generate multi-stationarity**

| Model (Complex present) | | Without feedback | Feedback from App | | Feedback from Ap | |
|---|---|---|---|---|---|---|
| | | | With Positive Feedback | With Negative Feedback | With Positive Feedback | With Negative Feedback |
| M3 | AK | Yes | Yes | | | |
| M4 | ApK | No | Yes | No | No | No |
| M5 | AppP | Yes | Yes | | | |
| M6 | ApP | No | Yes | No | Yes | No |

These results indicate that the elements of the biochemistry of multisite modification are integral to realizing efficient network level regulation through feedback. This coupled with the nature of the feedback can enable different qualitative behaviors at the systems level, as seen above with positive feedback from Ap enabling multi-stationarity in the system with ApP complex but not in the system with ApK complex. This is contrasted with when the feedback is from App in which case both models show multi-stationarity.

**Instructions**
Each subsection/subsection is self contained - hence please only run the necessary sections modularly (ignore if you are accessing this on the PDF).

## 0. Models (M3, M4, M5, M6) and proof of presence/absence of multi-stationarity

In this subsection we detail the various models (without any feedback and all but one modification step operating in the unsaturated limit of enzyme action) and illustrate their capacity/inability to exhibit multi-stationarity in parameter space. This is done in the following sub-subsections for AK, ApK, AppP and ApP complex (models M3, M4, M5 and M6 respectively).

### Model M3 (with AK complex): Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_{b1}\cdot A\cdot K + k_{ub1}\cdot AK + k_4\cdot Ap\cdot P :$$
$$dAp := -k_4\cdot Ap\cdot P - k_2\cdot Ap\cdot K + k_1\cdot AK + k_3\cdot App\cdot P :$$
$$dApp := -k_3\cdot App\cdot P + k_2\cdot Ap\cdot K :$$

$$dKs := -u_a\cdot Ks + u_d\cdot K :$$
$$dK := +u_a\cdot Ks - u_d\cdot K - k_{b1}\cdot A\cdot K + \left(k_{ub1} + k_1\right)\cdot AK :$$

$$dAK := k_{b1}\cdot A\cdot K - \left(k_{ub1} + k_1\right)\cdot AK :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App + AK - A_{Total} :$$
$$KCon := K + Ks + AK - K_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentrations of App and K. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$Ks := solve(dKs, Ks) = \frac{u_d K}{u_a}$

We similarly solve for the other variables using the same command.

$AK := solve(dAK, AK)$ :
$Sol := solve(\{dA, dAp\}, \{A, Ap\})$ :
$assign(Sol)$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and K for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below admit multi-stationarity in the full system (shown by solving ACon and KCon for App and K)**

$k_1 := 71.899159073249210$ : $k_2 := 89.503347099736500$ : $k_3 := 4.721351558066354$ : $k_4 :=$
$77.302863599793670$ :
$k_{b1} := 46.628869375010076$ : $k_{ub1} := 1$ : $u_a := 39.832563181210130$ : $u_d :=$
$20.981854490936737$ :
$A_{Total} := 1.981655394614496e+02$ : $K_{Total} := 11.543960390650538$ : $P :=$
$1.376396255573174$ :

$solve(\{KCon, ACon\}, useassumptions)$ assuming $K > 0, App > 0$
$\{App = 11.22987505, K = 0.1060078038\}, \{App = 180.4895010, K = 2.519318829\},$ **(1.1.1.1)**
$\{App = 189.5801031, K = 4.755749737\}$

### *Enzyme limiting regime*

**The parameters below admit multi-stationarity in the system, operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and K)**

$k_1 := 71.899159073249210$ : $k_2 := 89.503347099736500$ : $k_3 := 4.721351558066354$ : $k_4 :=$
$77.302863599793670$ :
$k_{b1} := 46.628869375010076$ : $k_{ub1} := 1$ : $u_a := 39.832563181210130$ : $u_d :=$
$20.981854490936737$ :
$A_{Total} := 1.981655394614496e+02$ : $K_{Total} := 11.543960390650538$ : $P :=$
$1.376396255573174$ :

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $K > 0, App > 0$
$\{App = 10.49398551, K = 0.09896769331\}, \{App = 189.1628200, K$ **(1.1.2.1)**

$= 2.790434946\}, \{App = 193.5723773, K = 4.599118628\}$

**In both instances we can see that the system admits three feasible steady states. Hence the model admits to multi-stationarity for some parameter choice.**

## Model M4 (with ApK complex): Absence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$
$dAp := -k_4 \cdot Ap \cdot P - k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$
$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$

$dKs := -u_a \cdot Ks + u_d \cdot K :$
$dK := +u_a \cdot Ks - u_d \cdot K - k_{b2} \cdot Ap \cdot K + (k_{ub2} + k_2) \cdot ApK :$

$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$ACon := A + Ap + App + ApK - A_{Total} :$
$KCon := K + Ks + ApK - K_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of Ap. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$Ks := solve(dKs, Ks) = \frac{u_d K}{u_a}$

We similarly solve for the other variables using the same command as shown.

$ApK := solve(dApK, ApK) :$
$Sol := solve(\{dA, dApp\}, \{A, App\}) :$
$assign(Sol)$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. We begin by solving the enzyme conservation equation to obtain steady state concentration of K as a function of Ap as shown below.

$$K := solve(KCon, K) = \frac{K_{Total}\, u_a \left(k_{ub2} + k_2\right)}{Ap\, k_{b2}\, u_a + k_2\, u_a + k_2\, u_d + k_{ub2}\, u_a + k_{ub2}\, u_d}$$

Re-substituting this in the conservation equation for the substrate results in the following equation.

$T := collect(simplify(numer(ACon)), Ap)$

$$\begin{aligned} T := {} & Ap^3 P^2 k_3 k_4 k_{b2}^2 u_a^2 + \left(2 k_4 \left(u_a \left(k_{ub2} + k_2\right) + u_d \left(k_{ub2} + k_2\right)\right) k_{b2} u_a P\right. \\ & + K_{Total} k_{b2} u_a^2 k_1 \left(k_{ub2} + k_2\right)\Big) P k_3 Ap^2 + \Big(\left(k_4 \left(u_a \left(k_{ub2} + k_2\right) + u_d \left(k_{ub2} + k_2\right)\right)^2 P\right. \\ & - K_{Total} \left(\left(k_{b2} A_{Total} - k_{b2} K_{Total} - k_2 - k_{ub2}\right) u_a - u_d \left(k_{ub2} + k_2\right)\right) k_1 u_a \left(k_{ub2} + k_2\right)\Big) \\ & P k_3 + k_{b2} K_{Total}^2 u_a^2 k_2 \left(k_{ub2} + k_2\right) k_1\Big) Ap - K_{Total} \left(A_{Total} \left(k_{ub2} + k_2\right) u_a\right. \\ & \left. + u_d A_{Total} \left(k_{ub2} + k_2\right)\right) k_1 u_a \left(k_{ub2} + k_2\right) P k_3 \end{aligned} \tag{1.2.1}$$

Now, in order for the system to accommodate multi-stationarity, this polynomial in Ap, should admit to multiple feasible positive real solutions for Ap describing the different steady states. Further, by Descartes rule of signs for the system to have three positive real roots, there should be three sign changes between the coeffecients of the polynomial. However since the leading coeffecient and the coeffecient of the second exponent are both always positive (since the parameters involved in them are all positive), the maximum sign changes cannot be three. Further since the constant term in the polynomial is again negative, the maximum sign changes is at most 1. Thus, there must exist two/zero real negative roots and only one positive real root for the polynomial (1.2.1).

Thus due to a simple violation of the Descartes rule of signs, this model is only capable of possessing one feasible steady state where all concentrations are positive and real.

**Thus multi-stationarity is not possible in this model for any parameter choice.**

Note: There is an alternative proof to show the absence of multi-stationarity in the ordered DSP (with common kinase and common phosphatase) with any or both of the Ap complexes present (ApK, ApP). Please see section 6 in 2.3.3.

## Model M5 (with AppP complex): Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := -k_2 \cdot Ap \cdot K - k_4 \cdot Ap \cdot P + k_1 \cdot A \cdot K + k_3 \cdot AppP :$$
$$dApp := -k_{b3} \cdot App \cdot P + k_{ub3} \cdot AppP + k_2 \cdot Ap \cdot K :$$

$$dKs := -u_a \cdot Ks + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - u_d \cdot K :$$
$$dP := -k_{b3} \cdot App \cdot P + \left(k_{ub3} + k_3\right) \cdot AppP :$$

$dAppP := k_{b3} \cdot App \cdot P - \left(k_{ub3} + k_3\right) \cdot AppP :$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + App + AppP - A_{Total} :$
$KCon := K + Ks - K_{Total} :$
$PCon := P + AppP - P_{Total} :$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$ACon\_Red := A + Ap + App - A_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of App and ratio of the active enzyme forms (introduced below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$Ks := solve(dKs, Ks) = \frac{K\,u_d}{u_a}$

We similarly solve for the other variables using the same command.

$AppP := solve(dAppP, AppP) :$
$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$
$assign(Sol)$

Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase as a function of the other concentrations and parameters.

$K := \epsilon \cdot P :$
$P := solve(PCon, P) :$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and K for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and $\epsilon$)**

$k_1 := 1.899433197735912 : k_2 := 95.393125020766100 : k_3 := 4.535232083692224 : k_4 := 21.419252647660140 :$
$k_{b3} := 28.177389605142256 : k_{ub3} := 1 : u_a := 34.836043243776930 : u_d := 22.091243070031037 :$
$A_{Total} := 79.226842030807520 : K_{Total} := 0.740058332902433 : P_{Total} := 21.402915274236097 :$

$solve(\{KCon, ACon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$
$\{App = 0.01470028542, \epsilon = 0.02274270825\}, \{App = 4.070061623, \epsilon = 0.4595559935\}, \{App = 51.09939397, \epsilon = 5.525205009\}$ **(1.3.1.1)**

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and $\epsilon$)**

$k_1 := 1.899433197735912 : k_2 := 95.393125020766100 : k_3 := 4.535232083692224 : k_4 := 21.419252647660140 :$
$k_{b3} := 28.177389605142256 : k_{ub3} := 1 : u_a := 34.836043243776930 : u_d := 22.091243070031037 :$
$A_{Total} := 79.226842030807520 : K_{Total} := 0.740058332902433 : P_{Total} := 21.402915274236097 :$

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $\epsilon > 0, App > 0$
$\{App = 0.01502865386, \epsilon = 0.02277807765\}, \{App = 2.755950221, \epsilon = 0.3180097117\}, \{App = 73.81609228, \epsilon = 7.972078368\}$ **(1.3.2.1)**

**In both instances we can see that the system admits three feasible steady states. Hence the model admits to multi-stationarity for some parameter choice.**

## Model M6 (with ApP complex): Absence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$
$dAp := -k_2 \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$
$dApp := -k_3 \cdot App \cdot P + k_2 \cdot Ap \cdot K :$

$dKs := -u_a \cdot Ks + u_d \cdot K$ :
$dK := u_a \cdot Ks - u_d \cdot K$ :
$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP$ :

$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP$ :

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + App + ApP - A_{Total}$ :
$KCon := K + Ks - K_{Total}$ :
$PCon := P + ApP - P_{Total}$ :

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of Ap. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\,u_d}{u_a}$$

We similarly solve for the other variables using the same command as shown.

$ApP := solve(dApP, ApP)$ :
$Sol := solve(\{dA, dApp\}, \{A, App\})$ :
$assign(Sol)$

Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase and the ratio of the enzymes as a function of concentration of Ap and parameters.

$K := \epsilon \cdot P$ :
$P := solve(PCon, P)$ :

$$\epsilon := solve(KCon, \epsilon) = \frac{K_{Total}\left(Ap\,k_{b4} + k_4 + k_{ub4}\right) u_a}{P_{Total}\left(k_{ub4} + k_4\right)\left(u_a + u_d\right)}$$

Re-substituting this in the conservation of the substrate results in the following equation.

$T := collect(numer(ACon), Ap)$

$$\begin{aligned} T := {} & k_1 k_2 k_{b4}^2 u_a^2 K_{Total}^2 Ap^3 + \left(2 k_1 k_2 k_4 k_{b4} u_a^2 K_{Total}^2 + 2 k_1 k_2 k_{b4} k_{ub4} u_a^2 K_{Total}^2\right. \\ & + k_1 k_3 k_4 k_{b4} u_a^2 K_{Total} P_{Total} + k_1 k_3 k_4 k_{b4} u_a u_d K_{Total} P_{Total} + k_1 k_3 k_{b4} k_{ub4} u_a^2 K_{Total} P_{Total} \\ & \left. + k_1 k_3 k_{b4} k_{ub4} u_a u_d K_{Total} P_{Total}\right) Ap^2 + \left(-k_1 k_3 k_4 k_{b4} u_a^2 A_{Total} K_{Total} P_{Total}\right. \\ & + k_1 k_3 k_4 k_{b4} u_a^2 K_{Total} P_{Total}^2 - k_1 k_3 k_4 k_{b4} u_a u_d A_{Total} K_{Total} P_{Total} \\ & + k_1 k_3 k_4 k_{b4} u_a u_d K_{Total} P_{Total}^2 - k_1 k_3 k_{b4} k_{ub4} u_a^2 A_{Total} K_{Total} P_{Total} \\ & \left. + k_1 k_3 k_{b4} k_{ub4} u_a^2 K_{Total} P_{Total}^2 - k_1 k_3 k_{b4} k_{ub4} u_a u_d A_{Total} K_{Total} P_{Total}\right. \end{aligned} \tag{1.4.1}$$

$$+ k_1 k_3 k_{b4} k_{ub4} u_a u_d K_{Total} P_{Total}^2 + k_1 k_2 k_4^2 u_a^2 K_{Total}^2 + 2 k_1 k_2 k_4 k_{ub4} u_a^2 K_{Total}^2$$
$$+ k_1 k_2 k_{ub4}^2 u_a^2 K_{Total}^2 + k_1 k_3 k_4^2 u_a^2 K_{Total} P_{Total} + k_1 k_3 k_4^2 u_a u_d K_{Total} P_{Total}$$
$$+ 2 k_1 k_3 k_4 k_{ub4} u_a^2 K_{Total} P_{Total} + 2 k_1 k_3 k_4 k_{ub4} u_a u_d K_{Total} P_{Total}$$
$$+ k_1 k_3 k_{ub4}^2 u_a^2 K_{Total} P_{Total} + k_1 k_3 k_{ub4}^2 u_a u_d K_{Total} P_{Total} + k_3 k_4^2 k_{b4} u_a^2 P_{Total}^2$$
$$+ 2 k_3 k_4^2 k_{b4} u_a u_d P_{Total}^2 + k_3 k_4^2 k_{b4} u_d^2 P_{Total}^2 + k_3 k_4 k_{b4} k_{ub4} u_a^2 P_{Total}^2$$
$$+ 2 k_3 k_4 k_{b4} k_{ub4} u_a u_d P_{Total}^2 + k_3 k_4 k_{b4} k_{ub4} u_d^2 P_{Total}^2 ) Ap$$
$$- k_1 k_3 k_4^2 u_a^2 A_{Total} K_{Total} P_{Total} - k_1 k_3 k_4^2 u_a u_d A_{Total} K_{Total} P_{Total}$$
$$- 2 k_1 k_3 k_4 k_{ub4} u_a^2 A_{Total} K_{Total} P_{Total} - 2 k_1 k_3 k_4 k_{ub4} u_a u_d A_{Total} K_{Total} P_{Total}$$
$$- k_1 k_3 k_{ub4}^2 u_a^2 A_{Total} K_{Total} P_{Total} - k_1 k_3 k_{ub4}^2 u_a u_d A_{Total} K_{Total} P_{Total}$$

Now, in order for the system to accommodate multi-stationarity, this polynomial in Ap, should admit to multiple feasible positive real solutions for Ap describing the different steady states. Further, by Descartes rule of signs for the system to have three positive real roots, there should be three sign changes between the coeffecients of the polynomial. However since the leading coeffecient and the coeffecient of the second exponent are both always positive (since the parameters involved in them are all positive), the maximum sign changes cannot be three. Further since the constant term in the polynomial is again negative, the maximum sign changes is at most 1. Thus, there must exist two/zero real negative roots and only one positive real root for the polynomial (1.4.1).

Thus due to a simple violation of the Descartes rule of signs, this model is only capable of possessing one feasible steady state where all concentrations are positive and real.

**Thus multi-stationarity is not possible in this model for any parameter choice.**

Note: There is an alternative proof to show the absence of multi-stationarity in the ordered DSP (with common kinase and common phosphatase) with any or both of the Ap complexes present (ApK, ApP). Please see section 6 in 2.3.3.

## 1. Positive feedback from App : Presence of multi-stationarity

In this subsection, we detail the various ordered DSP models with all but one modification step operating in the unsaturated limit of enzyme action and with direct App positive feedback, and show their capacity to exhibit multi-stationarity in parameter space. This is done for models containing ApK and ApP complex only (models M4 and M6 respectively). The analysis for models containing AK and AppP complex (models M3 and M5) are not presented since they can already exhibit multi-stationarity without any explicit feedback (refer above). App feedback generating multi-stationarity here in these models is to be expected as App feedback was able to generate multi-stationarity in the model with no complexes (see section 2.4.1).

### With ApK Complex: Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :
```

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions

will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := -k_4 \cdot Ap \cdot P - k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot App + u_d \cdot K :$$
$$dK := u_a \cdot Ks + k_{PFB} \cdot Ks \cdot App - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$

$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App + ApK - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentrations of App and K. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{App\, k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$$ApK := solve(dApK, ApK) :$$
$$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$$
$$assign(Sol)$$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and K for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and K)**

$k_1 := 63.076982750810394 : k_2 := 84.653472127887940 : k_3 := 51.922600626380580 : k_4 := 92.300037949203460 :$
$k_{b2} := 35.148994318095700 : k_{ub2} := 1 :$
$k_{PFB} := 27.661084427201410 : u_a := 0.100875545627312 : u_d := 35.895226014215720 :$
$A_{Total} := 1.749425691457615e+02 : K_{Total} := 17.853564564775844 : P := 19.558214721315263 :$

$solve(\{KCon, ACon\}, useassumptions)$ assuming $K > 0, App > 0$
$\{App = 0.0007615718313, K = 0.06041495289\}, \{App = 0.01918791037, K = 0.3047491015\}, \{App = 1.036155573, K = 2.412678623\}$ **(2.1.1.1)**

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and K)**

$k_1 := 63.076982750810394 : k_2 := 84.653472127887940 : k_3 := 51.922600626380580 : k_4 := 92.300037949203460 :$
$k_{b2} := 35.148994318095700 : k_{ub2} := 1 :$
$k_{PFB} := 27.661084427201410 : u_a := 0.100875545627312 : u_d := 35.895226014215720 :$
$A_{Total} := 1.749425691457615e+02 : K_{Total} := 17.853564564775844 : P := 19.558214721315263 :$

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $K > 0, App > 0$
$\{App = 0.0007616323662, K = 0.06041577722\}, \{App = 0.01914697223, K = 0.3042206558\}, \{App = 1.053458954, K = 2.341620773\}$ **(2.1.2.1)**

**In both instances we can see that the system admits three feasible steady states for some positive feedback strength, while the system without feedback was incapable of multi-stationarity.**

**Thus direct positive feedback from App generates multi-stationarity in the system.**

## With ApP Complex: Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_2 \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot Ap \cdot K :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot App + u_d \cdot K :$$
$$dK := +u_a \cdot Ks + k_{PFB} \cdot Ks \cdot App - u_d \cdot K :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApP - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of App and ratio of the active enzyme forms (introduced below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\,u_d}{App\,k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$$ApP := solve(dApP, ApP) :$$
$$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$$
$$assign(Sol)$$

Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase as a function of the other concentrations and parameters.

$$K := \epsilon \cdot P :$$
$$P := solve(PCon, P) :$$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and $\epsilon$ for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the

expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and K)**

$k_1 := 11.427590135067721 : k_2 := 77.916378347520960 : k_3 := 53.260373560983830 : k_4 := 88.679426386919390 :$
$k_{b4} := 47.883187710468526 : k_{ub4} := 1 :$
$k_{PFB} := 2.278433010708424 : u_a := 1.117157225917797 : u_d := 43.973876782418730 :$
$A_{Total} := 1.050675950807511e+02 : K_{Total} := 0.171595545789285 : P_{Total} := 0.784847933763541 :$

$solve(\{KCon, ACon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ **(2.2.1.1)**
$\{App = 0.001271623604, \epsilon = 0.005860578144\}, \{App = 11.46210363, \epsilon = 0.6347245236\}, \{App = 39.11700410, \epsilon = 1.524066002\}$

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and K)**

$k_1 := 11.427590135067721 : k_2 := 77.916378347520960 : k_3 := 53.260373560983830 : k_4 := 88.679426386919390 :$
$k_{b4} := 47.883187710468526 : k_{ub4} := 1 :$
$k_{PFB} := 2.278433010708424 : u_a := 1.117157225917797 : u_d := 43.973876782418730 :$
$A_{Total} := 1.050675950807511e+02 : K_{Total} := 0.171595545789285 : P_{Total} := 0.784847933763541 :$

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ **(2.2.2.1)**
$\{App = 0.001272435890, \epsilon = 0.005860842862\}, \{App = 11.21938052, \epsilon = 0.6241784527\}, \{App = 39.98845905, \epsilon = 1.545829028\}$

**In both instances we can see that the system admits three feasible steady states for some positive feedback strength, while the system without feedback was incapable of multi-stationarity.**

**Thus direct positive feedback from App generates multi-stationarity in the system.**

## 2. Positive feedback from Ap: Absence and presence of multi-stationarity

In this subsection, we detail the various ordered DSP models with all but one modification step operating in the unsaturated limit of enzyme action and with direct Ap positive feedback, and show their capacity to exhibit multi-stationarity in parameter space. This is done for models containing ApK and ApP complex only (models M4 and M6 respectively). The analysis for models containing AK and AppP complex (models M3 and M5) are not presented since they can already exhibit multi-stationarity without any explicit feedback (refer above).

## With ApK Complex: Absence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := -k_4 \cdot Ap \cdot P - k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot Ap + u_d \cdot K :$$
$$dK := u_a \cdot Ks + k_{PFB} \cdot Ks \cdot Ap - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$

$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App + ApK - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{PFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{PFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$fn := [dA, dAp, dApp, dApK, dKs, dK]:$
$dc := [A, Ap, App, ApK, Ks, K]:$

$J := Jacobian(fn, dc):$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^2$, as the Jacobian contains two linearly dependent rows (arising from the two conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^2}:$$

The constant term in this characteristic polynomial represents the determinant of the system's Jacobian at some steady state. This determinant is zero at the saddle node point, when the stability of the system changes. In the following steps, we show using the inbuilt verify command in Maple that while the variables and parameters involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$C0 := simplify(coeff(C, x, 0))$

$$C0 := P\left(k_4\left(Ap\left(k_{PFB}\,Ap + u_a\right)k_{b2} + \left(k_{ub2} + k_2\right)\left(k_{PFB}\,Ap + u_a + u_d\right)\right)P + k_1\left(k_{b2}\left(k_{PFB}\,Ap + u_a + u_d\right)K^2 + \left(\left(Ap\,(A + Ap + Ks)\,k_{PFB} + u_a\,(A + Ap)\right)k_{b2} + \left(k_{ub2} + k_2\right)\left(k_{PFB}\,Ap + u_a + u_d\right)\right)K - A\,k_{PFB}\,Ks\left(k_{ub2} + k_2\right)\right)\right)k_3 + \left(\left(k_{PFB}\,Ap + u_a + u_d\right)K + k_{PFB}\,Ks\,Ap\right)k_1\,K\,k_{b2}\,k_2 \quad \textbf{(3.1.1)}$$

Now we begin by algebraically solving the system of equations to obtain expressions linking the steady state concentrations of the variables. This is carried out to obtain a simplified version of the expression C0 from above which is dependent on fewer variables. For this purpose we use an inbuilt *Maple* command *solve* as shown below.

$Ks := solve(dKs, Ks):$
$ApK := solve(dApK, ApK):$

$Sol := solve(\{dA, dAp\}, \{A, Ap\}):$
$assign(Sol)$

Now we use the inbuilt *verify* command in *Maple* to show that, while the parameters and variable concentrations are positive, the constant term is strictly greater than zero.

$verify(C0, 0, greater\_than)$ assuming $k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, k_{b2} > 0, k_{ub2} > 0, k_{PFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0$

$$true \quad \textbf{(3.1.2)}$$

**This thus proves that a saddle node bifurcation cannot occur, indicating that the system remains incapable of multi-stationarity, irrespective of the value of $k_{PFB}$.**

**Hence the model is strictly not capable of multi-stationarity.**

Note: A similar proof employing the Jacobian of the system (defined at steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## With ApP Complex: Presence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_2 \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot Ap \cdot K :$$

$$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot Ap + u_d \cdot K :$$
$$dK := u_a \cdot Ks + k_{PFB} \cdot Ks \cdot Ap - u_d \cdot K :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApP - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of App and ratio of the active enzyme forms (introduced below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{K\, u_d}{Ap\, k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$$ApP := solve(dApP, ApP) :$$
$$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$$
$$assign(Sol)$$

Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase as a function of the other concentrations and parameters.

$K := \epsilon \cdot P$ :
$P := solve(PCon, P)$ :

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and $\epsilon$ for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and $\epsilon$)**

$k_1 := 11.427590135067721$ : $k_2 := 77.916378347520960$ : $k_3 := 53.260373560983830$ : $k_4 := 88.679426386919390$ :
$k_{b4} := 47.883187710468526$ : $k_{ub4} := 1$ :
$k_{PFB} := 2.278433010708424$ : $u_a := 1.117157225917797$ : $u_d := 43.973876782418730$ :
$A_{Total} := 1.050675950807511e+02$ : $K_{Total} := 0.171595545789285$ : $P_{Total} := 0.784847933763541$ :

$solve(\{KCon, ACon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ **(3.2.1.1)**
$\{App = 0.002769940158, \epsilon = 0.008653589010\}, \{App = 4.999485134, \epsilon = 0.3949639925\}, \{App = 18.11335309, \epsilon = 0.8462996362\}$

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and $\epsilon$)**

$k_1 := 11.427590135067721$ : $k_2 := 77.916378347520960$ : $k_3 := 53.260373560983830$ : $k_4 := 88.679426386919390$ :
$k_{b4} := 47.883187710468526$ : $k_{ub4} := 1$ :
$k_{PFB} := 2.278433010708424$ : $u_a := 1.117157225917797$ : $u_d := 43.973876782418730$ :
$A_{Total} := 1.050675950807511e+02$ : $K_{Total} := 0.171595545789285$ : $P_{Total} := 0.784847933763541$ :

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $\epsilon > 0, App > 0$ **(3.2.2.1)**
$\{App = 0.002775018985, \epsilon = 0.008658138943\}, \{App = 4.651342338, \epsilon$

$= 0.3783819483\}, \{App = 19.17798527, \epsilon = 0.8752078147\}$

**In both instances we can see that the system admits three feasible steady states for some positive feedback strength, while the system without feedback was incapable of multi-stationarity.**

**Thus direct positive feedback from Ap generates multi-stationarity in the system.**

## 3. Negative feedback from App: Absence of multi-stationarity

In this subsection, we detail the various ordered DSP models with all but one modification step operating in the unsaturated limit of enzyme action and with direct App negative feedback, and show their capacity to exhibit multi-stationarity in parameter space. This is done for models containing ApK and ApP complex only (models M4 and M6 respectively). The analysis for models containing AK and AppP complex (models M3 and M5) are not presented since they can already exhibit multi-stationarity without any explicit feedback (refer above).

### With ApK Complex: Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := -k_4 \cdot Ap \cdot P - k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot App + u_d \cdot K :$$
$$dK := u_a \cdot Ks - k_{NFB} \cdot K \cdot App - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$

$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App + ApK - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive k

$_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$$fn := [dA, dAp, dApp, dApK, dKs, dK]:$$
$$dc := [A, Ap, App, ApK, Ks, K]:$$

$$J := Jacobian(fn, dc):$$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^2$, as the Jacobian contains two linearly dependent rows (arising from the two conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^2}:$$

The constant term in this characteristic polynomial represents the determinant of the Jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when the stability of the system changes. In the following step, we show using the inbuilt verify command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$$C0 := simplify(coeff(C, x, 0))$$

$$C0 := \left(u_a k_{b2} Ap + \left(k_{ub2} + k_2\right)\left(k_{NFB} App + u_a + u_d\right)\right) k_3 k_4 P^2 + K\left(k_1\left(k_{b2}\left(k_{NFB} App + u_a + u_d\right) K + u_a (A + Ap) k_{b2} + \left(k_{ub2} + k_2\right)\left(k_{NFB} App + u_a + u_d\right)\right) k_3 + k_{b2} k_{NFB} Ap k_2 k_4\right) P + K^2 k_1 \left((A + Ap + App) k_{NFB} + u_a + u_d\right) k_{b2} k_2 \quad \textbf{(4.1.1)}$$

Now we begin by algebraically solving the system of equations to obtain expressions linking the steady state concentrations of the variables. This is carried out to obtain a simplified version of the expression C0 from above which is dependent on fewer variables. For this purpose we use an inbuilt *Maple* command *solve* as shown below.

$$Ks := solve(dKs, Ks):$$
$$ApK := solve(dApK, ApK):$$

$$Sol := solve(\{dA, dAp\}, \{A, Ap\}):$$
$$assign(Sol)$$

Now we use the inbuilt *verify* command in *Maple* to show that, while the parameters and variables are positive, the constant term is strictly greater than zero (We can also make this conclusion from directly observing 4.1.1 that C0 is a strictly positive sum of functions of parameters and varaiables that are all

positive themselves at feasible values).

$verify(C0, 0, greater\_than)$ assuming $k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, k_{b2} > 0, k_{ub2} > 0, k_{NFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0$

$$true \qquad \textbf{(4.1.2)}$$

**This thus proves that a saddle node bifurcation cannot occur, indicating that the system remains incapable of multi-stationarity, irrespective of the value of $k_{NFB}$.**

**Hence the model is strictly not capable of multi-stationarity.**

Note: A similar proof employing the Jacobian of the system (defined at a steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## With ApP Complex: Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_2 \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot Ap \cdot K :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot App + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot App - u_d \cdot K :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApP - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$$fn := [dA, dAp, dApp, dApP, dKs, dK, dP]:$$
$$dc := [A, Ap, App, ApP, Ks, K, P]:$$

$$J := Jacobian(fn, dc):$$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^3$, as the Jacobian contains three linearly dependent rows (arising from the three conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^3}:$$

The constant term in this characteristic polynomial represents the determinant of the Jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when the stability of the system changes. In the following step, we show using the inbuilt verify command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$$C0 := simplify(coeff(C, x, 0))$$
$$C0 := k_2 \left(k_{b4} k_{NFB} Ap P + \left((A + Ap + App) k_{NFB} + u_a + u_d\right) \left(k_{b4} Ap + k_4 + k_{ub4}\right)\right) k_1 K^2 + P \left(\left(k_{b4} \left(App k_{NFB} + u_a + u_d\right) P + \left(App (A + Ap + App) k_{NFB} + (Ap + App) \left(u_a + u_d\right)\right) k_{b4} + \left(k_{ub4} + k_4\right) \left(App k_{NFB} + u_a + u_d\right)\right) k_3 k_1 + k_{b4} k_{NFB} Ap k_2 k_4\right) K + k_{b4} P^2 k_3 k_4 \left(App k_{NFB} + u_a + u_d\right) \quad \textbf{(4.2.1)}$$

Now we use the inbuilt *verify* command in *Maple* to show that, while the parameters and variables are positive, the constant term is strictly greater than zero.

$$verify(C0, 0, greater\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, k_{b4} > 0, k_{ub4} > 0, k_{NFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0, A > 0, Ap > 0$$
$$true \quad \textbf{(4.2.2)}$$

**This thus proves that a saddle node bifurcation cannot occur, indicating that the system remains incapable of multi-stationarity, irrespective of the value of $k_{NFB}$.**

**Hence the model is strictly not capable of multi-stationarity.**

Note: A similar proof employing the Jacobian of the system (defined at a steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## 4. Negative feedback from Ap: Absence of multi-stationarity

In this subsection, we detail the various ordered DSP models with all but one modification step operating in the unsaturated limit of enzyme action and with direct Ap negative feedback, and show their capacity to exhibit multi-stationarity in parameter space. This is done for models containing ApK and ApP complex only (models M4 and M6 respectively). The analysis for models containing AK and AppP complex (models M3 and M5) are not presented since they can already exhibit multi-stationarity without any explicit feedback (refer above).

### With ApK Complex: Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot Ap \cdot P :$$
$$dAp := -k_4 \cdot Ap \cdot P - k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot Ap + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot Ap - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$

$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state). Note that there is no phosphatase conservation equation in this model, and P is a 'parameter' in the model rather than a variable.

$$ACon := A + Ap + App + ApK - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$$fn := [dA, dAp, dApp, dApK, dKs, dK]:$$
$$dc := [A, Ap, App, ApK, Ks, K]:$$

$$J := Jacobian(fn, dc):$$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^2$, as the Jacobian contains two linearly dependent rows (arising from the two conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^2}:$$

The constant term in this characteristic polynomial represents the determinant of the Jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when the stability of the system changes. In the following step, we show using the inbuilt verify command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$$C0 := simplify(coeff(C, x, 0))$$
$$\begin{aligned} C0 := {} & P\left(k_4\left(\left(k_{NFB}\,Ap + u_a + u_d\right)k_2 + \left(k_{b2}\,Ap + k_{ub2}\right)u_a + k_{ub2}\left(k_{NFB}\,Ap + u_d\right)\right)P\right. \\ & + \left(k_{b2}\left(u_d + u_a\right)K + \left(u_a + (A + Ap)\,k_{NFB} + u_d\right)k_2 + \left(k_{ub2} + (A + Ap)\,k_{b2}\right)u_a\right. \\ & \left.\left. + \left((A + Ap)\,k_{NFB} + u_d\right)k_{ub2}\right)k_1 K\right)k_3 + k_{b2}\,K^2\,k_1\,k_2\left(u_d + u_a\right) \end{aligned} \quad \textbf{(5.1.1)}$$

Now we begin by solving the system of equations to obtain expresssions linking the steady state concentrations of the variables, primarily to obtain a simplified version of the expression C0 from above which is dependent on fewer variables. For this purpose we use an inbuilt *Maple* command *solve* as shown below.

$$Ks := solve(dKs, Ks):$$
$$ApK := solve(dApK, ApK):$$

$$Sol := solve(\{dA, dAp\}, \{A, Ap\}):$$
$$assign(Sol)$$

Now we use the inbuilt *verify* command in Maple to show that, while the parameters and variables are positive, the constant term is strictly greater than zero (We can also make this conclusion from directly observing from 5.1.1 that C0 is a strictly positive sum of functions of parameters and varaiables that are all positive themselves at feasible values).

$$verify(C0, 0, greater\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, k_{b2} > 0, k_{ub2} > 0, k_{NFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0$$
$$true \quad \textbf{(5.1.2)}$$

**This thus proves that a saddle node bifurcation cannot occur, indicating that the system remains incapable of multi-stationarity, irrespective of the value of $k_{NFB}$.**

**Hence the model is strictly not capable of multi-stationarity.**

Note: A similar proof employing the Jacobian of the system (defined at a steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## With ApP Complex: Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_2 \cdot Ap \cdot K - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot Ap \cdot K :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot Ap + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - k_{NFB} \cdot K \cdot Ap - u_d \cdot K :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the enzymes respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApP - A_{Total} :$$
$$KCon := K + Ks - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$fn := [dA, dAp, dApp, dApP, dKs, dK, dP]:$
$dc := [A, Ap, App, ApP, Ks, K, P]:$

$J := Jacobian(fn, dc):$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt Maple function (CharactericPolynomial). We divide the resulting polynomial by $x^3$, as the Jacobian contains three linearly dependent rows (arising from the three conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^3}:$$

The constant term in this characteristic polynomial represents the determinant of the Jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when the stability of the system changes. In the following step, we show using the inbuilt verify command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$C0 := simplify(coeff(C, x, 0))$

$$C0 := K\left(k_2\left(k_{b4}\,Ap + k_4 + k_{ub4}\right)\left(u_a + u_d\right)K + k_3\,P\left(k_{b4}\left(Ap\,k_{NFB} + u_a + u_d\right)P + \left(k_{NFB}\,Ap^2 + \left((A + App)\,k_{NFB} + u_a + u_d\right)Ap + App\left(u_a + u_d\right)\right)k_{b4} + \left(k_{ub4} + k_4\right)\left(A\,k_{NFB} + Ap\,k_{NFB} + u_a + u_d\right)\right)k_1 + k_{b4}\,P^2\,k_3\,k_4\left(Ap\,k_{NFB} + u_a + u_d\right)\right) \quad \textbf{(5.2.1)}$$

Now we use the inbuilt *verify* command in *Maple* to show that, while the parameters and variables are positive, the constant term is strictly greater than zero.

$verify(C0, 0, greater\_than)$ assuming $k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, k_{b4} > 0, k_{ub4} > 0, k_{NFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0, A > 0, Ap > 0$

$$true \quad \textbf{(5.2.2)}$$

**This thus proves that a saddle node bifurcation cannot occur, indicating that the system remains incapable of multi-stationarity, irrespective of the value of $k_{NFB}$.**

**Hence the model is strictly not capable of multi-stationarity.**

Note: A similar proof employing the Jacobian of the system (defined at steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

# Ordered DSP with only Ap modification steps not in the unsaturated limit of enzyme action
# Impact of feedback on network behavior (Multi-stationarity)

This worksheet explores the effect of feedback regulation on the propensity for multi-stationarity in the ordered DSP (with common kinase and common phosphatase), when the phosphorylation and dephosphorylation of only the unmodified and the fully modified substrate operate in the unsaturated limit of enzyme action. We first consider such a basic isolated model without any feedback, showing that this system is necessarily incapable of exhibiting multi-stationarity. Further on we systematically study the effect of each feedback on models capacity to generate multi-stationarity. The organization of these results and the emerging synthesis is expanded on below.

The analysis is not carried out for the other complex combinations as the presence of any of the other complexes (AK or AppP) directly generates for multi-stability in the model without any explicit feedback (see 2.4.2 for further details)

**Note:** This analysis is carried out for only when the feedback is effected directly on the kinase through activation and or inactivation.

**0. Model:** The ordered DSP model without any feedback (with only Ap complexes present - Model M2)
**1. Absence of multi-stationarity in M2:** Analytic proof establishing that M2 model has only one possible steady state (ruling out multi-stationarity).

**2. Positive feedback from App (activating kinase):** Presence of multi-stationarity
**3. Positive feedback (analog) from Ap (activating kinase):** Presence of multi-stationarity
**4. Negative feedback from App (deactivating kinase):** Absence of multi-stationarity
**5. Negative feedback fanalog) from Ap (deactivating kinase):** Absence of multi-stationarity

**6. Alternate argument for the absence of multi-stationarity in M2.**

We note here that positive feedback from App and Ap generating multi-stationarity in the current model with both Ap complexes is expected as the system with only one complex present (ApP complex) was able to generate multi-stationarity in both instances (see section 2.4.1 and 2.4.2). A summary of these results is tabulated below.

**: Capacity of ordered DSP (all but Ap complexes in the unsaturated limit) with feedback to generate multi-sta**

| Model | Multi-stationarity |
|---|---|
| Isolated ordered DSP (M2) | **No** |
| App Positive Feedback | **Yes (even in the enzyme limiting regime)** |
| App Negative Feedback | **No** |
| Ap Positive Feedback | **Yes (even in the enzyme limiting regime)** |
| Ap Negative Feedback | **No** |

These results indicate that the nature of the feedback at the network regulation level (positive by activating the kinase and its negative analogue) is key to enforcing (and introducing) qualitative behavior such as multi-stationarity. With sufficient sequestration within the biochemistry of the multisite modification, positive

feedback from App and Ap are both able to generate multi-stationarity. However in theses cases negative feedback is unable to increase multi-stationarity.

**Instructions**
Each subsection apart from 0. & 1. is self contained - hence only run the necessary sections modularly (Ignore if you are accessing this on the PDF).

# 0. Model (M2)

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks + u_d \cdot K :$$
$$dK := +u_a \cdot Ks - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the enzymes which are described below. Here we store these conservation expressions as ACon, KCon and PCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApK + ApP - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

# 1. Absence of multi-stationarity in M2

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the the steady state concentrations of variables, primarily we solve the equations to obtain steady state expressions for all variables as a function of steady state concentration of Ap and parameters. For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{u_a}$$

We similarly solve for the other variables using the same command.

$$ApK := solve(dApK, ApK) : ApP := solve(dApP, ApP) :$$
$$Sol := solve(\{dA, dApp\}, \{A, App\}) :$$
$$assign(Sol)$$

At this stage, we introduce a new variable $\epsilon$ as the ratio of the free (unbound) kinase and the phosphatase as shown below. We also introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

$$K := \epsilon \cdot P :$$
$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) : k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

We now have three equations, ACon, PCon, KCon - the conservation of the total substrate and enzyme concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. We now solve these PCon and KCon to obtain expressions for the steady state concentrations of P and $\epsilon$ in terms of Ap and the parameters.

$$P := solve(PCon, P) :$$
$$\epsilon := solve(KCon, \epsilon) :$$

Resubstituting these into the remaining expression for ACon (Stored as T=0), it simplifies into the following third degree polynomial in Ap (whose solutions are the distinct solutions of the system at various steady states). Note that when the solutions for Ap are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

$$T := collect(numer(ACon), Ap)$$

$$\begin{aligned} T := {} & \left(c_2^2 c_4 k_3 k_4 u_a^2 P_{Total}^2 + c_2 c_4^2 k_1 k_2 u_a^2 K_{Total}^2 + c_2 c_4 k_1 k_3 u_a^2 K_{Total} P_{Total}\right) Ap^3 + \left( \right. \\ & -c_2 c_4 k_1 k_3 u_a^2 A_{Total} K_{Total} P_{Total} + c_2 c_4 k_1 k_3 u_a^2 K_{Total}^2 P_{Total} + c_2 c_4 k_1 k_3 u_a^2 K_{Total} P_{Total}^2 \\ & + 2 c_2 c_4 k_1 k_2 u_a^2 K_{Total}^2 + 2 c_2 c_4 k_3 k_4 u_a^2 P_{Total}^2 + 2 c_2 c_4 k_3 k_4 u_a u_d P_{Total}^2 \\ & + c_2 k_1 k_3 u_a^2 K_{Total} P_{Total} + c_4 k_1 k_3 u_a^2 K_{Total} P_{Total} + c_4 k_1 k_3 u_a u_d K_{Total} P_{Total} \left.\right) Ap^2 + \left( \right. \\ & -c_2 k_1 k_3 u_a^2 A_{Total} K_{Total} P_{Total} + c_2 k_1 k_3 u_a^2 K_{Total}^2 P_{Total} - c_4 k_1 k_3 u_a^2 A_{Total} K_{Total} P_{Total} \\ & + c_4 k_1 k_3 u_a^2 K_{Total} P_{Total}^2 - c_4 k_1 k_3 u_a u_d A_{Total} K_{Total} P_{Total} + c_4 k_1 k_3 u_a u_d K_{Total} P_{Total}^2 \\ & + c_2 k_1 k_2 u_a^2 K_{Total}^2 + c_4 k_3 k_4 u_a^2 P_{Total}^2 + 2 c_4 k_3 k_4 u_a u_d P_{Total}^2 + c_4 k_3 k_4 u_d^2 P_{Total}^2 \\ & + k_1 k_3 u_a^2 K_{Total} P_{Total} + k_1 k_3 u_a u_d K_{Total} P_{Total} \left.\right) Ap - k_1 k_3 u_a^2 A_{Total} K_{Total} P_{Total} \\ & - k_1 k_3 u_a u_d A_{Total} K_{Total} P_{Total} \end{aligned} \tag{2.1}$$

Now we can ascertain by the structure of the polynomial in Ap above that the coefficient of the leading exponent of Ap and the constant term are positive and negative always. This implies the system either has 1 or 3 positive real roots. Thus in order for this system to admit to 3 positive roots (multi-stationarity) - the system should necessarily satisfy Descartes rule of signs for the signs of the remaining coefficients. In particular this stipulates that coefficient of the second exponent be negative while the coefficient of the first exponent is positive, please note that this is only a necessary and not sufficient condition for the system to admit to multi-stationarity.

However, this is not possible due to the factorization of the terms in these two polynomials. We show this below using the inbuilt *Maple* command *verify*.

In order to establish this, we first collect the various coefficients of the polynomial into variables C3, C2, C1, C0 as shown below.

$C3 := coeff(numer(ACon), Ap, 3) : C2 := coeff(numer(ACon), Ap, 2) : C1 := coeff(numer(ACon), Ap, 1) : C0 := coeff(numer(ACon), Ap, 0) :$

We can now use the command *verify* as shown below to show that while C1 is positive, C2 is also necessarily positive.

$verify(C2, 0, greater\_than)$ assuming $k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, c_1 > 0, c_2 > 0, c_3 > 0, c_4 > 0, u_a > 0, u_d > 0, Ap > 0, K_{Total} > 0, P_{Total} > 0, A_{Total} > 0, C1 > 0$

$$true \tag{2.2}$$

This violates Descartes rule of signs - thus the polynomial can admit only one real positive root.

**Thus the system is necessarily incapable of multi-stationarity.**

Note: An alternate proof can be used to establish absence of multi-stationarity in the model using an argument of monotonicity (refer to section 6).

## 2. Positive feedback from App (activating kinase): Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modeled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$
$dAp := -k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$
$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$

$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot App + u_d \cdot K :$
$dK := u_a \cdot Ks + k_{PFB} \cdot Ks \cdot App - u_d \cdot K - k_{b2} \cdot Ap \cdot K + (k_{ub2} + k_2) \cdot ApK :$
$dP := -k_{b4} \cdot Ap \cdot P + (k_{ub4} + k_4) \cdot ApP :$

$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$
$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$

The model is also associated with conservation conditions for the substrate and the kinase and the phosphatase enzyme which are described below. Here we store these conservation expressions as ACon, KCon, PCon for the substrate, the kinase and the phosphatase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + App + ApK + ApP - A_{Total} :$
$KCon := K + Ks + ApK - K_{Total} :$
$PCon := P + ApP - P_{Total} :$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the

enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$ACon\_Red := A + Ap + App - A_{Total}$ :

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of App and ratio of the active enzyme forms (introduced below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{App\ k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$ApK := solve(dApK, ApK)$ : $ApP := solve(dApP, ApP)$ :

$Sol := solve(\{dA, dAp\}, \{A, Ap\})$ :
$assign(Sol)$

We now have three equations, ACon, PCon, KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase as a function of the other concentrations and parameters.

$K := \epsilon \cdot P$ :
$P := solve(PCon, P)$ :

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and $\epsilon$ for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

## *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and $\epsilon$)**

$k_1 := 73.766774870814170$ : $k_2 := 50.728659481758854$ : $k_3 := 13.238074900624039$ : $k_4 := 19.193486191836730$ :
$k_{b2} := 43.980000871648210$ : $k_{b4} := 45.993124426309450$ : $k_{ub2} := 1$ : $k_{ub4} := 1$ :
$k_{PFB} := 11.387592549349863$ : $u_a := 0.122434842985345$ : $u_d := 27.102650018489605$ :
$A_{Total} := 200$ : $K_{Total} := 2.406185147705586$ : $P_{Total} := 18.423952551406850$ :

$solve(\{ACon, KCon\}, useassumptions)$ assuming $App > 0, \epsilon > 0$
$\{App = 0.002233951537, \epsilon = 0.001462381593\}, \{App = 0.004669048683, \epsilon$ **(3.1.1)**

$= 0.002124569052\}, \{App = 52.33037476, \epsilon = 0.3440940225\}$

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and ϵ)**

$k_1 := 73.766774870814170 : k_2 := 50.728659481758854 : k_3 := 13.238074900624039 : k_4 := 19.193486191836730 :$
$k_{b2} := 43.980000871648210 : k_{b4} := 45.993124426309450 : k_{ub2} := 1 : k_{ub4} := 1 :$
$k_{PFB} := 11.387592549349863 : u_a := 0.122434842985345 : u_d := 27.102650018489605 :$
$A_{Total} := 1.563694105533203e+02 : K_{Total} := 2.406185147705586 : P_{Total} := 18.423952551406850 :$

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $\epsilon > 0$
$\{App = 0.0009073046475, \epsilon = 0.001028227996\}, \{App = 0.01006766155, \epsilon = 0.003432159244\}, \{App = 45.47443770, \epsilon = 0.3431998319\}$ **(3.2.1)**

**In both instances we can see that the system admits three feasible steady states for some positive feedback strength, while the system without feedback was incapable of multi-stationarity. Thus direct positive feedback from App generates multi-stationarity in the system.**

We note here that positive feedback from App generating multi-stationarity in the current model with both Ap complexes is expected as the system with only no complex present was able to generate multi-stationarity (see section 2.4.1).

## 3. Positive feedback (analogue) from Ap (activating kinase): Presence of multi-stationarity

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modelled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$
$dAp := -k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$
$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$

$dKs := -u_a \cdot Ks - k_{PFB} \cdot Ks \cdot Ap + u_d \cdot K :$
$dK := u_a \cdot Ks + k_{PFB} \cdot Ks \cdot Ap - u_d \cdot K - k_{b2} \cdot Ap \cdot K + (k_{ub2} + k_2) \cdot ApK :$
$dP := -k_{b4} \cdot Ap \cdot P + (k_{ub4} + k_4) \cdot ApP :$

$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$
$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$

The model is also associated with conservation conditions for the substrate and the kinase and the phosphatase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + App + ApK + ApP - A_{Total} :$
$KCon := K + Ks + ApK - K_{Total} :$
$PCon := P + ApP - P_{Total} :$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$ACon\_Red := A + Ap + App - A_{Total} :$

Now we begin by algebraically solving the system of equations at steady state to obtain expressions linking the steady state concentrations of variables, primarily we solve the equations to obtain expressions for steady state concentrations of all variables as a function of steady state concentration of App and ratio of the active enzyme forms (introduced below). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Ks := solve(dKs, Ks) = \frac{u_d K}{Ap\, k_{PFB} + u_a}$$

We similarly solve for the other variables using the same command.

$ApK := solve(dApK, ApK) : ApP := solve(dApP, ApP) :$
$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$
$assign(Sol)$

We now have three equations, ACon, PCon, KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Here, we introduce a new variable $\epsilon$ as the ratio of the free enzymes and continue solving the remaining conservation equation for the phosphatase to evaluate the steady state concentration of phosphatase as a function of the other concentrations and parameters.

$K := \epsilon \cdot P :$
$P := solve(PCon, P) :$

We now have two equations, ACon and KCon - the conservation of the total substrate and kinase concentration, whose solutions for the variables now define the steady state concentrations of the variables and thus the system. Note that when solutions for App and $\epsilon$ for ACon and KCon are positive the remaining variable concentrations are also positive (as can be easily established from the expressions linking them).

We now show that the system is capable of exhibiting multi-stationarity by providing parameters and demonstrating the multiple steady states of the system, both in the full system and in the reduced enzyme limiting regime. To do this, we solve the corresponding equations, ACon and ACon_Red respectively.

### *Full system*

**The parameters below show the presence of multi-stationarity in the full system (shown by solving ACon and KCon for App and ϵ)**

$k_1 := 12 : k_2 := 20 : k_3 := 4 : k_4 := 30 :$
$k_{b2} := 0.5 : k_{b4} := 20 : k_{ub2} := 1 : k_{ub4} := 1 :$
$k_{PFB} := 5 : u_a := 0.01 : u_d := 20 :$
$A_{Total} := 61 : K_{Total} := 3.082419324 : P_{Total} := 30 :$

$solve(\{KCon, ACon\}, useassumptions)$ assuming $\epsilon > 0, App > 0$
$\{App = 0.00002060244053, \epsilon = 0.002166843258\}, \{App = 0.001267986795, \epsilon = 0.01814703028\}, \{App = 0.02859253356, \epsilon = 0.09828219670\}$ **(4.1.1)**

### *Enzyme limiting regime*

**The parameters below show the presence of multi-stationarity in the system operating in the enzyme limiting regime (shown by solving ACon_Red and KCon for App and ϵ)**

$k_1 := 12.671815778685290 : k_2 := 15.350910681055495 : k_3 := 4.171797154872850 : k_4 := 25.612969178615450 :$
$k_{b2} := 0.439837895897674 : k_{b4} := 16.389916794389020 : k_{ub2} := 1 : k_{ub4} := 1 :$
$k_{PFB} := 4.145883769718306 : u_a := 0.097091819880133 : u_d := 24.974445763783336 :$
$A_{Total} := 82.892339149290550 : K_{Total} := 1.634287530580490 : P_{Total} := 28.667598537434646 :$

$solve(\{KCon, ACon\_Red\}, useassumptions)$ assuming $\epsilon > 0, App > 0$
$\{App = 2.556982106\ 10^{-6}, \epsilon = 0.0006229980969\}, \{App = 0.004942009844, \epsilon = 0.02768570467\}, \{App = 0.8947837481, \epsilon = 0.4296509802\}$ **(4.2.1)**

**In both instances we can see that the system admits three feasible steady states for some positive feedback strength, while the system without feedback was incapable of multi-stationarity. Thus direct positive feedback from Ap generates multi-stationarity in the system.**

We note here that positive feedback from Ap generating multi-stationarity in the current model with both Ap complexes is expected as the system with only one complex present (ApP complex) was able to generate multi-stationarity (see section 2.4.2).

## 4. Negative feedback from App (deactivating kinase): Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$

The system is modelled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot App + u_d \cdot K :$$
$$dK := u_a \cdot Ks - k_{NFB} \cdot K \cdot App - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase and the phosphatase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApK + ApP - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$$fn := [dA, dAp, dApp, dApK, dApP, dKs, dK, dP] :$$
$$dc := [A, Ap, App, ApK, ApP, Ks, K, P] :$$

$$J := Jacobian(fn, dc) :$$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^3$, as the Jacobian contains three linearly dependent rows (arising from the three conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^3} :$$

The constant term in this characteristic polynomial represents the determinant of the jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when

the stability of the system changes. In the following step, we show using the inbuilt verify command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. We first isolate for the constant term and label it as C0.

$$C0 := simplify(coeff(C, x, 0))$$

$$C0 := \left(\left(c_2\, Ap\, \left(k_{ub2}+k_2\right)\left(k_{NFB}\,k_2\,Ap+\left(\left(u_a+u_d+k_{ub4}+k_4\right)k_3+k_{NFB}\,k_2\right)P+\left(A\,k_{NFB}+k_{NFB}\,App+k_4+k_{ub4}+u_a+u_d\right)k_2+\left(k_{NFB}\,App+u_a+u_d\right)\left(k_{ub4}+k_4\right)\right)k_1\,K^2+\left(\left(c_2\left(k_3\left(u_a+k_{ub4}+k_4\right)P+\left(k_{ub4}+k_4\right)u_a\right)\left(k_{ub2}+k_2\right)Ap^2+\left(k_3\,u_a\,c_2\left(k_{ub2}+k_2\right)P^2+\left(\left(k_{ub2}+k_2\right)\left(u_a+k_{ub4}+k_4\right)(A+App)\,c_2+\left(k_{NFB}\,App+k_4+k_{ub4}+u_a+u_d\right)k_2+\left(k_{NFB}\,App+k_4+k_{ub4}+u_a+u_d\right)k_{ub2}+\left(k_{NFB}\,App+u_a+u_d\right)\left(k_{ub4}+k_4\right)\right)k_3\,P+\left(k_{ub2}+k_2\right)\left(k_{ub4}+k_4\right)\left(A\,c_2\,u_a+k_{NFB}\,App+u_a+u_d\right)\right)Ap+P\,k_3\left(\left(k_{ub2}+k_2\right)\left(k_{NFB}\,App+u_a+u_d\right)P+App\left(\left(A\,k_{NFB}+k_{NFB}\,App+k_4+k_{ub4}+u_a+u_d\right)k_2+\left(A\,k_{NFB}+k_{NFB}\,App+k_4+k_{ub4}+u_a+u_d\right)k_{ub2}+\left(k_{ub4}+k_4\right)\left(A\,k_{NFB}+k_{NFB}\,App+u_a+u_d\right)\right)\right)\right)k_1+c_2\left(k_{NFB}\,k_2\left(k_{ub4}+k_4\right)Ap+\left(\left(k_{ub4}+k_4\right)\left(u_a+u_d\right)k_3+k_{NFB}\,k_2\,k_4\right)P+k_2\left(k_{NFB}\,App+u_a+u_d\right)\left(k_{ub4}+k_4\right)\right)Ap\left(k_{ub2}+k_2\right)\right)K+\left(\left(k_{ub4}+k_4\right)Ap+k_4\,P+App\left(k_{ub4}+k_4\right)\right)\left(k_{ub2}+k_2\right)P\,k_3\left(Ap\,c_2\,u_a+k_{NFB}\,App+u_a+u_d\right)\right)c_4+K\left(k_{ub2}+k_2\right)\left(c_2\left(k_{NFB}\,k_2\,Ap+k_3\left(k_{NFB}\,App+u_a+u_d\right)P+k_2\left(A\,k_{NFB}+k_{NFB}\,App+u_a+u_d\right)\right)K+P\,k_3\left(A\,c_2\,u_a+Ap\,c_2\,u_a+k_{NFB}\,App+u_a+u_d\right)\right)k_1\right)\left(k_{ub4}+k_4\right) \quad (5.1)$$

Now we use the inbuilt *verify* command in *Maple* to show that, while the parameters and variables are positive, the constant term is strictly greater than zero.

$$verify(C0, 0, greater\_than) \text{ assuming } k_1>0,\ k_2>0,\ k_3>0,\ k_4>0,\ k_{b1}>0,\ k_{b2}>0,\ k_{b3}>0,\ k_{b4}>0,\ k_{ub1}>0,\ k_{ub2}>0,\ k_{ub3}>0,\ k_{ub4}>0,\ k_{NFB}>0,\ P>0,\ K>0,\ App>0,\ u_a>0,\ u_d>0,\ Ap>0,\ A>0$$

$$true \quad (5.2)$$

**This thus proves that a saddle node bifurcation cannot occur, implying that the system remains incapable of multi-stationarity (noting that the system was incapable of multi-stationarity without feedback), irrespective of the value of $k_{NFB}$.**

Note: A similar proof employing the Jacobian of the system (defined at steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## 5. Negative feedback (analogue) from Ap (deactivating kinase) : Absence of multi-stationarity

We initialize the *Maple* file with the *restart* command and load the relevant libraries of inbuilt *Maple* functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart: with(LinearAlgebra): with(VectorCalculus): with(Student[LinearAlgebra]):
```

The system is modelled as a set of ODEs as shown below. Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be

equal to zero.

$$dA := -k_1 \cdot A \cdot K + k_4 \cdot ApP :$$
$$dAp := -k_{b2} \cdot Ap \cdot K + k_{ub2} \cdot ApK - k_{b4} \cdot Ap \cdot P + k_{ub4} \cdot ApP + k_1 \cdot A \cdot K + k_3 \cdot App \cdot P :$$
$$dApp := -k_3 \cdot App \cdot P + k_2 \cdot ApK :$$

$$dKs := -u_a \cdot Ks + k_{NFB} \cdot K \cdot Ap + u_d \cdot K :$$
$$dK := u_a \cdot Ks - k_{NFB} \cdot K \cdot Ap - u_d \cdot K - k_{b2} \cdot Ap \cdot K + \left(k_{ub2} + k_2\right) \cdot ApK :$$
$$dP := -k_{b4} \cdot Ap \cdot P + \left(k_{ub4} + k_4\right) \cdot ApP :$$

$$dApP := k_{b4} \cdot Ap \cdot P - \left(k_{ub4} + k_4\right) \cdot ApP :$$
$$dApK := k_{b2} \cdot Ap \cdot K - \left(k_{ub2} + k_2\right) \cdot ApK :$$

The model is also associated with conservation conditions for the substrate and the kinase and the phosphatase enzyme which are described below. Here we store these conservation expressions as ACon and KCon for the substrate and the kinase respectively. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + ApK + ApP - A_{Total} :$$
$$KCon := K + Ks + ApK - K_{Total} :$$
$$PCon := P + ApP - P_{Total} :$$

We now proceed to show that this system is incapable of multi-stationarity. To begin with, we note that the basic system with no feedback (accounting to $k_{NFB} = 0$) is incapable of multi-stationarity and thus is strictly incapable of multi-stationarity. Thus in order for the network to exhibit multi-stationarity for some positive $k_{NFB}$, there needs to necessarily exist a saddle node bifurcation point, where the system looses stability.

We show that this is not possible, by showing that the Determinant of the Jacobian (which is indicative of the stability of the system can never be zero) and thus multi-stationarity is not possible in the system.

Note: The Determinant of the Jacobian here is the product of the Eigenvalues of the system at any given steady state and since at the saddle node point the stability of the system changes, one of the Eigenvalue is 0 and thus the Determinant must be 0.

In order to do this, we first calculate the Jacobian of the system (as a function of the variables). This is carried out symbolically using the inbuilt *Maple* command, *Jacobian*. This Jacobian can then be evaluated at some steady state, by inputting the concentrations of the various variables at the steady state.

$$fn := [dA, dAp, dApp, dApK, dApP, dKs, dK, dP] :$$
$$dc := [A, Ap, App, ApK, ApP, Ks, K, P] :$$

$$J := Jacobian(fn, dc) :$$

We then evaluate the characteristic polynomial (generated as a polynomial in the variable x) of this Jacobian using the inbuilt *Maple* function (*CharactericPolynomial*). We divide the resulting polynomial by $x^3$, as the Jacobian contains three linearly dependent rows (arising from the three conservations built into the model).

$$C := \frac{CharacteristicPolynomial(J, x)}{x^3} :$$

The constant term in this characteristic polynomial represents the determinant of the jacobian of the system (to be ascertained at some steady state). This determinant is zero at the saddle node point, when the stability of the system changes. In the following step, we show using the inbuilt *verify* command in Maple that while the variables involved are positive (since they represent feasible steady state concentrations and parameter values) the determinant is always positive. Here C0 represents the determinant of the Jacobian.

$$C0 := simplify(coeff(C, x, 0))$$

$$C0 := \left(\left(c_2 k_1 \left(P \left(u_a + u_d + k_{ub4} + k_4\right) k_3 + \left(u_a + u_d + k_{ub4} + k_4\right) k_2 + \left(u_a + u_d\right)\left(k_{ub4} + k_4\right)\right) Ap \left(k_{ub2} + k_2\right) K^2 + \left(k_1 \left(\left(\left(\left(k_4 + u_a + k_{ub4}\right) c_2 + k_{NFB}\right) k_2 + k_{ub2}\left(k_4 + u_a + k_{ub4}\right) c_2 + k_{NFB}\left(k_4 + k_{ub2} + k_{ub4}\right)\right) P k_3 + \left(k_{ub2} + k_2\right)\left(c_2 u_a + k_{NFB}\right)\left(k_{ub4} + k_4\right)\right) Ap^2 + \left(\left(k_1 \left(k_{ub2} + k_2\right)\left(c_2 u_a + k_{NFB}\right) P + \left(\left(\left(k_4 + u_a + k_{ub4}\right)(A + App) c_2 + A k_{NFB} + k_{NFB} App + k_4 + u_a + u_d + k_{ub4}\right) k_2 + k_{ub2}\left(k_4 + u_a + k_{ub4}\right)(A + App) c_2 + \left(A k_{NFB} + k_{NFB} App + k_4 + k_{ub4} + u_a + u_d\right) k_{ub2} + \left(k_{ub4} + k_4\right)\left(A k_{NFB} + k_{NFB} App + u_a + u_d\right)\right) k_1 + c_2\left(k_{ub2} + k_2\right)\left(u_a + u_d\right)\left(k_{ub4} + k_4\right)\right) P k_3 + \left(\left(A u_a c_2 + A k_{NFB} + u_a + u_d\right) k_1 + c_2 k_2\left(u_a + u_d\right)\right)\left(k_{ub4} + k_4\right)\left(k_{ub2} + k_2\right)\right) Ap + k_1 \left(\left(k_{ub2} + k_2\right)\left(u_a + u_d\right) P + App \left(\left(u_a + u_d + k_{ub4} + k_4\right) k_2 + \left(u_a + u_d + k_{ub4} + k_4\right) k_{ub2} + \left(u_a + u_d\right)\left(k_{ub4} + k_4\right)\right)\right) k_3 P\right) K + \left(\left(k_{ub4} + k_4\right) Ap + k_4 P + App\left(k_{ub4} + k_4\right)\right)\left(\left(c_2 u_a + k_{NFB}\right) Ap + u_a + u_d\right) k_3\left(k_{ub2} + k_2\right) P\right) c_4 + k_1 K\left(c_2\left(u_a + u_d\right)\left(P k_3 + k_2\right) K + \left(\left(c_2 u_a + k_{NFB}\right) Ap + A u_a c_2 + A k_{NFB} + u_a + u_d\right) k_3 P\right)\left(k_{ub2} + k_2\right)\right)\left(k_{ub4} + k_4\right) \quad (6.1)$$

$$verify(C0, 0, greater\_than) \text{ assuming } k_1 > 0, k_2 > 0, k_3 > 0, k_4 > 0, c_1 > 0, c_2 > 0, c_3 > 0, c_4 > 0, k_{ubl} > 0, k_{ub2} > 0, k_{ub3} > 0, k_{ub4} > 0, k_{NFB} > 0, P > 0, K > 0, App > 0, u_a > 0, u_d > 0, A > 0, Ap > 0$$

$$true \quad (6.2)$$

**This thus proves that a saddle node bifurcation cannot occur, implying that the system remains incapable of multi-stationarity (noting that the system was incapable of multi-stationarity without feedback), irrespective of the value of $k_{NFB}$.**

Note: A similar proof employing the Jacobian of the system (defined at steady state) and Brouwers degree can be used to establish absence of multi-stationarity (see Conradi and Mincheva, 2014).

## 6. Alternate proof for absence of multi-stationarity in model M2

We present an alternative proof (to the one presented in section 1) to establish absence of multi-stationarity in model M2 (due Suwamnmajo and Krishnan, (Exploring the intrinsic behaviour of multisite phosphorylation systems as part of signalling pathwaysJ. R. Soc. Interface.152018010920180109). Here we sketch the proof of the argument; which follows on from the systematic reduction in subsection 1, resulting in equation 2.1, which involved expressing the steady state concentrations of all substrates and complexes in terms of that of Ap. To recapitulate, instead of simplifying the equation 2.1 we write the substrate conservation as it is here below.

$$ACon := \frac{c_4 Ap P_{Total} k_4 \left(Ap c_2 u_a + u_a + u_d\right)}{\left(Ap c_4 + 1\right) K_{Total} u_a k_1} + Ap + \frac{c_2 Ap K_{Total} u_a k_2 \left(Ap c_4 + 1\right)}{\left(Ap c_2 u_a + u_a + u_d\right) P_{Total} k_3}$$

$$+ \frac{c_2\, Ap\, K_{Total}\, u_a}{Ap\, c_2\, u_a + u_a + u_d} + \frac{c_4\, Ap\, P_{Total}}{Ap\, c_4 + 1} = A_{Total}$$

The left hand side of this is the conservation of the substrates and the two Ap complexes (A + Ap + App + ApK + ApP) while the right hand side is the total concentration of the substrate in the model.

We note that the equation is a function in just one variable and that each term of the left hand side (as a function of Ap) is strictly monotonic with Ap (increases as Ap increases). Thus entire left hand side of this equation is a monotonic function of Ap while the right hand side is a constant concentration of total substrate in the system.

This implies that the number of viable solutions of Ap where the left hand side equals the right hand side is at most 1 (in the positive real range of Ap values). Thus the system can only accomdate one positive steady state concentration of Ap.

**Hence the system (model M2) is strictly incapable of multi-stationarity.**

# Ordered DSP Network - Role of Feedback in perturbing Biphasic behavior (Along S)

In this Maple worksheet we analytically establish the following results regarding the role of feedback regulation on biphasic dose-response behavior in the maximally phosphorylated substrate as a function of signal concentration (where signal activates the kinase), in the ordered DSP system (with common kinase and common phosphatase).

1. If the modular isolated DSP network presented with biphasic dose-response in App with signal, then the network with feedback (any) will too (except in the case of positive feedback where biphasic behavior might be precluded - see below).
2. The network with positive/negative feedback through activation/deactivation of kinase will have the same App and $\epsilon$ value at the peak/biphasic point.
3. The saturation value of App (at high concentrations of signal) is also the same irrespective of the strength or nature of feedback ($k_{PFB}/k_{NFB}$).
4. This signal value at which App shows biphasic transition (dubbed peak S) changes with the feedback strength in the following way
   a. For PFB - As feedback strength increases, the peak S moves towards zero and at a certain point the biphasic 'peak' disappears by moving into negative signal ranges.
   b. For NFB - As feedback strength increases the peak S moves towards higher concentrations (approaches infinity).

The analytical proofs for each of these assertions is presented in detail in the following sections for the corresponding models of feedback.

The general logic of the proof follows from the mathematical definition of biphasic behavior in the DSP network with signal concentration, i.e. the existence of some steady state of concentrations where $\frac{dApp}{dS} = 0$. By delineating the necessary conditions from the mathematical model for this condition to be true, we can arrive at the specific proofs (as shown below).

## Isolated DSP with activation and inactivation (No Feedback)

We first begin by analyzing the system with **no feedback**, which acts as our standard system for this analysis which is then perturbed by the different feedbacks (positive or negative from App or Ap - refer to figure 1 in the main text for schematic).

We initiate this section by refreshing Maple and invoking the necessary libraries needed for computation.

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The model of the DSP system with activation and inactivation of the kinase can then be represented by the following system of ODEs

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$dK := -k_{b1}\,A\,K + \left(k_{ub1} + k_1\right) AK - k_{b2}\,Ap\,K + \left(k_{ub2} + k_2\right) ApK + u_a \cdot S \cdot Ks - u_d \cdot K + u_\beta \cdot Ks :$
$dKs := -u_a \cdot S \cdot Ks + u_d \cdot K - u_\beta \cdot Ks :$
$dP := -k_{b3}\,App\,P + \left(k_{ub3} + k_3\right) AppP - k_{b4}\,Ap\,P + \left(k_{ub4} + k_4\right) ApP :$

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$
$PCon := P + AppP + ApP - P_{Total} :$
$KCon := K + AK + ApK + Ks - K_{Total} :$

We now solve the system at steady state and obtain expressions for steady state concentrations of species in terms of each other. By successively solving the ODE and conservation equations we wish to obtain expressions for all steady state concentrations in terms of steady state concentrations of App, and $\epsilon$, where $\epsilon$ is defined as the ratio of the active kinase and the phosphatase ($\frac{K}{P}$ - explained below). Here we use an inbuilt *Maple* command *solve* to solve the ODEs and conservation equations, and an example of this is shown below.

$$Ks := solve(dKs, Ks) = \frac{u_d\,K}{S\,u_a + u_\beta}$$

We similarly solve for the other variables using the same command as shown.

$AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP := solve(dApP, ApP) :$
$Sol := solve(\{dA, dAp\}, \{A, Ap\}) :$
$assign(Sol)$

As mentioned earlier, we now introduce a new ratio, $\epsilon = \frac{K}{P}$, and rewrite the kinetic constants ($k_{b1}$, $k_{b2}$, $k_{b3}$, and $k_{b4}$) using new parameters. This is done for the sake of brevity and easy tractability of the expressions obtained.

$K := \epsilon \cdot P :$
$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) : k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$

Now we solve for the steady state concentration of the phosphatase using the conservation expression for the enzyme (PCon = 0)

$P := solve(simplify(PCon), P) :$

This results in the following expressions linking the steady state concentrations of different variables and that of App and $\epsilon$.

$$A = \frac{k_4\,c_4\,c_3\,App\,k_3}{\epsilon^2\,k_2\,c_2\,k_1\,c_1}$$

$$Ap = \frac{c_3\,App\,k_3}{\epsilon\,k_2\,c_2}$$

$$AK = \frac{k_4\,c_4\,c_3\,App\,k_3\,P_{Total}}{k_1\left(App\,c_2\,c_3\,\epsilon\,k_2 + c_4\,c_3\,App\,k_3 + \epsilon\,k_2\,c_2\right)}$$

$ApK =$

$$AppP = \frac{\frac{c_2\, c_3\, App\, k_3\, \epsilon\, P_{Total}}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}$$

$$ApP = \frac{\frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}$$

$$Ks = \frac{\frac{c_4\, c_3\, App\, k_3\, P_{Total}}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}}{App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2}$$

$$Ks = \frac{\frac{u_a\, \epsilon^2\, P_{Total}\, k_2\, c_2}{\left(App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2\right)}}{\left(App\, c_2\, c_3\, \epsilon\, k_2 + App\, c_3\, c_4\, k_3 + c_2\, \epsilon\, k_2\right)\ \left(S\, u_a + u_\beta\right)}$$

Here we make note of a couple of things
1. We have algebraically solved the system of ODEs at steady state and the only unsolved equations left that define the system are ACon and KCon (the total substrate conservation and kinase conservation)
2. That for any positive real value (non-zero) of both App and $\epsilon$ at steady state, the steady state concentrations of the other variables are also positive and real.

Thus this implies that the steady state concentrations of the system are defined by the positive solution of set {App,$\epsilon$} in the equations{ACon,KCon}

Now we rewrite these two expressions as follows (F, G).

$F := solve(KCon, S)$ :
$G := ACon$ :

Where F is rearranged and solved for S. We have used the inbuilt *Maple* function *solve* here to perform this operation. Thus,

$F = S$ :
$G = 0$ :

Each of the above expressions as mentioned earlier are functions of App. Now differentiating both of these expressions, we have.

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dF}{dApp} \cdot \frac{dApp}{dS}$$

$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dG}{dApp} \cdot \frac{dApp}{dS}$$

Now since we are interested in conditions describing the presence of biphasic behavior, we begin with the assumption that a biphasic behavior in dose-response exists in this system, i.e. $\frac{dApp}{dS} = 0$ for some finite real positive steady state concentration of App and $\epsilon$. Thus the above expressions simplify to

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$
$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$

**Note**

This reduction was possible since the denominator of $\frac{dF}{dApp}$ and $\frac{dG}{dApp}$ are both non-zero when App and $\epsilon$ are non-zero and positive, as shown below.

$$\frac{dF}{dApp} = simplify(diff(F, App)) =$$

$$-\left(k_1 c_3 u_d P_{Total} k_2 \epsilon^2 c_2 \left(\left(c_2 \left(K_{Total} k_2 - k_3 P_{Total}\right)\epsilon + K_{Total} c_4 k_3\right) k_1 - c_4 k_3 k_4 P_{Total}\right)\right) \Big/ \left(\left(\left(-\epsilon^2 P_{Total} k_2 c_2 + \left(-App\, c_3 k_3 P_{Total} + K_{Total} k_2 \left(App\, c_3 + 1\right)\right) c_2 \epsilon + App\, K_{Total} c_3 c_4 k_3\right) k_1 - k_4 c_4 c_3 App\, k_3 P_{Total}\right)^2 u_a\right)$$

$$F = simplify(F) =$$

$$\left(\left(P_{Total} c_2 k_2 \left(u_d + u_\beta\right)\epsilon^2 - u_\beta \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 \epsilon - App\, u_\beta K_{Total} c_3 c_4 k_3\right) k_1 + App\, c_3 c_4 k_3 k_4 u_\beta P_{Total}\right) \Big/ \left(u_a \left(\left(-\epsilon^2 P_{Total} k_2 c_2 + \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 \epsilon + App\, K_{Total} c_3 c_4 k_3\right) k_1 - k_4 c_4 c_3 App\, k_3 P_{Total}\right)\right)$$

$$\frac{dG}{dApp} = simplify(diff(G, App)) =$$

$$\left(k_2^2 c_2^3 k_1 \left(App^2 c_3^2 k_2 + \left(\left(2\, App + P_{Total}\right) k_2 + k_3 P_{Total}\right) c_3 + k_2\right) c_1 \epsilon^4 + k_2^2 c_2^2 k_3 c_3 \left(App^2 c_3^2 k_1 + 2\, App\, k_1 \left(App\, c_4 + 1\right) c_3 + 2\, App\, c_4 k_1 + P_{Total} c_4 k_1 + P_{Total} c_4 k_4 + k_1\right) c_1 \epsilon^3 + 2 k_2 c_4 c_2 \left(\frac{k_2 k_4 \left(App\, c_3 + 1\right)^2 c_2}{2} + App\, c_1 c_3 k_1 k_3 \left(App\, c_3 + \frac{1}{2} App\, c_4 + 1\right)\right) k_3 c_3 \epsilon^2 + \left(2 k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) c_4^2 App\, k_3^2 c_3^2 \epsilon + App^2 c_3^3 c_4^3 k_3^3 k_4\right) \Big/ \left(k_2 c_2 \epsilon^2 \left(c_2 k_2 \left(App\, c_3 + 1\right)\epsilon + App\, c_3 c_4 k_3\right)^2 k_1 c_1\right)$$

Note the denominator of $\frac{dF}{dApp}$ is non-zero as signal = F shares the same denominator and for a feasible positive signal this denominator cannot be 0.

The only other way in which the denominator of F = signal could still be 0 and yield a finite Signal is if the numerator is also 0. However the numerator of F can be written as the sum of a linear function of the denominator and a constant (non-zero term). This is shown below to be the case.

$$Nr\_F := \left(P_{Total} c_2 k_2 \left(u_d + u_\beta\right)\epsilon^2 - u_\beta \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 \epsilon - App\, u_\beta K_{Total} c_3 c_4 k_3\right) k_1 + App\, c_3 c_4 k_3 k_4 u_\beta P_{Total} :$$

$$Dr\_F := u_a \left(\left(-\epsilon^2 P_{Total} k_2 c_2 + \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 \epsilon + App\, K_{Total} c_3 c_4 k_3\right) k_1 - k_4 c_4 c_3 App\, k_3 P_{Total}\right) :$$

$$simplify\left(Nr\_F - \left(-\frac{Dr\_F \cdot u_\beta}{u_a} + c_2 \epsilon^2 k_1 k_2 u_d P_{Total}\right)\right) = 0$$

Hence it is also not possible for both the numerator and denominator to be simultaneously zero. Thus the denominator of $\frac{dF}{dApp}$ is non-zero allowing us proceed with the aforementioned simplification.

Now from the nature of the equation above, we can straight away discern that the steady state for a biphasic behavior, $\frac{d\epsilon}{dS}$ cannot be equal to zero. Now, we evaluate $\frac{dG}{d\epsilon}$. We use the inbuilt Maple function *diff* to evaluate these terms, as shown below.

$$\frac{dG}{d\epsilon} = simplify(diff(G, \epsilon)) =$$

$$-\left(App\, k_3 \left(k_2 c_2^2 c_1 \left(\left(App^2 c_3^2 k_1 + App \left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total} \left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App\, P_{Total} c_3 c_4 k_1 k_3\right)\epsilon^3 + 2 k_2 c_4 c_2 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right)\left(App\, c_3 + 1\right)\epsilon^2 + c_4^2 \left(4 k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) App\, k_3 c_3 \epsilon + 2\, App^2 c_3^2 c_4^3 k_3^2 k_4\right) c_3\right) \Big/ \left(k_2 c_2 \epsilon^3 \left(c_2 k_2 \left(App\, c_3 + 1\right)\epsilon + App\, c_3 c_4 k_3\right)^2 k_1 c_1\right)$$

In the above expression we note that the denominator is always positive (as $\epsilon$ and App can only take positive finite values). Thus the numerator must be independently equal to 0 for biphasic behavior to exist. We isolate the numerator of the above expression as shown below using the inbuilt Maple command *numer*.

$$T := simplify(numer(diff(G, \epsilon))) :$$

Hence we have the above condition - relabeled as expression T = 0 - which must be satisfied at the point of biphasic behavior in the dose-response graph. Thus if we summarize the constraints that define the biphasic transition, we have
1. G = 0 (ACon) [a requirement of steady state]
2. F = S (KCon) [a requirement of steady state]
3. T = 0 (diff(ACon, $\epsilon$)) [following from our assumption of the presence of biphasic behavior in the dose-response graph of App vs S]

$$KCon = \frac{\epsilon^2 P_{Total} k_2 c_2}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2} + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2} + \frac{u_d \epsilon^2 P_{Total} k_2 c_2}{\left(App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2\right) \left(S u_a + u_\beta\right)} - K_{Total}$$

$$ACon = \frac{k_4 c_4 c_3 App k_3}{\epsilon^2 k_2 c_2 k_1 c_1} + \frac{c_3 App k_3}{\epsilon k_2 c_2} + App + \frac{k_4 c_4 c_3 App k_3 P_{Total}}{k_1 \left(App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2\right)} + \frac{c_2 c_3 App k_3 \epsilon P_{Total}}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2} + \frac{c_3 App P_{Total} \epsilon k_2 c_2}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2} + \frac{c_4 c_3 App k_3 P_{Total}}{App c_2 c_3 \epsilon k_2 + App c_3 c_4 k_3 + c_2 \epsilon k_2} - A_{Total}$$

$$T = -App k_3 \left(k_2 c_2^2 c_1 \left(\left(App^2 c_3^2 k_1 + App \left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total} \left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App P_{Total} c_3 c_4 k_1 k_3\right) \epsilon^3 + 2 k_2 c_4 c_2 \left(k_2 k_4 \left(App c_3 + 1\right) c_2 + App c_1 c_3 k_1 k_3\right) \left(App c_3 + 1\right) \epsilon^2 + c_4^2 \left(4 k_2 k_4 \left(App c_3 + 1\right) c_2 + App c_1 c_3 k_1 k_3\right) App k_3 c_3 \epsilon + 2 App^2 c_3^2 c_4^3 k_3^2 k_4\right) c_3$$

Since we begin by assuming that the basic system is capable of biphasic behavior, we assume that this system of coupled polynomial equations has solutions $App_0$ and $\epsilon_0$ where these three conditions are satisfied.

## DSP with feedback (Ap/App) through activation or inactivation of kinase

We now analyze the system with feedback (positive and negative from App/Ap) through the activation/deactivation of kinase, with an intermediate tier of covalent modification cycle mediating the feedback (Xs to X inter-conversion in this model). We begin by defining a generic master model in this analytical work that can be suitably adapted to the four different models of feedback presented in the paper (see figure 1 and below). The purpose of considering a generic master model of the above sort is to minimize the need for repetition in the proofs as the insights we obtain and the logic followed in one model is equally valid in the others.

We initiate this section by refreshing *Maple* and invoking the necessary libraries needed for computation.

$$restart : with(LinearAlgebra) : with(VectorCalculus) : with(Student[LinearAlgebra]) :$$

The model of the feedback from a substrate form 'M' activating and deactivating kinase in the DSP system can be represented by the following system of ODEs

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1}\, A\, K + (k_{ub1} + k_1)\, AK - k_{b2}\, Ap\, K + (k_{ub2} + k_2)\, ApK + u_a \cdot S \cdot Ks - (X \cdot (-k_{PFB} \cdot Ks + k_{NFB} \cdot K)) - u_d \cdot K + u_\beta \cdot Ks :$$
$$dKs := -u_a \cdot S \cdot Ks + X \cdot (-k_{PFB} \cdot Ks + k_{NFB} \cdot K) + u_d \cdot K - u_\beta \cdot Ks :$$
$$dP := -k_{b3}\, App\, P + (k_{ub3} + k_3)\, AppP - k_{b4}\, Ap\, P + (k_{ub4} + k_4)\, ApP :$$

$$dXs := -d_a \cdot Xs \cdot M + d_d \cdot X :$$
$$dX := +d_a \cdot Xs \cdot M - d_d \cdot X :$$

Where M is either Ap/App corresponding to whether the feedback is coming from the partially modified substrate or the fully modified substrate and $k_{PFB}$ and $k_{NFB}$ represent the strength of the positive and negative feedback respectively. Thus, the four models presented in the main text (figure 1) can be arrived at by the following choice of parameters for M and the feedback strengths.

**Table 1: Variations of the ODE description to represent specific feedback architectures**

| Model | M | $k_{PFB}$ | kNFB |
|---|---|---|---|
| App Positive Feedback | App | $k_{PFB}$ | 0 |
| App Negative Feedback | App | 0 | $k_{NFB}$ |
| Ap Positive Feedback | Ap | $k_{PFB}$ | 0 |
| Ap Negative Feedback | Ap | 0 | $k_{NFB}$ |

The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state). These equations are independent of the nature or presence of feedback.

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$
$$XCon := X + Xs - X_{Total} :$$

Similar to the analysis earlier with the no feedback we solve the system at steady state and obtain expressions linking the steady state concentrations of various species in terms of each other (However

unlike earlier we do not solve for the modified substrate forms, A/Ap/App at this stage). For this purpose we use an inbuilt *Maple* command *solve* as shown with the example below.

$$Xs := solve(dXs, Xs) = \frac{d_d X}{M d_a}$$

We similarly solve for the other variables using the same command as shown.

$$Ks := solve(dKs, Ks) :$$
$$AK := solve(dAK, AK) : ApK := solve(dApK, ApK) : AppP := solve(dAppP, AppP) : ApP := solve(dApP, ApP) :$$

We now introduce a new ratio, $\epsilon = \frac{K}{P}$ and also rewrite the kinetic constants ($k_{b1}$, $k_{b2}$, $k_{b3}$, and $k_{b4}$) using new parameters. This is done for the sake of brevity and easy tractability of the expressions obtained.

$$K := \epsilon \cdot P :$$
$$k_{b1} := c_1 \cdot \left(k_1 + k_{ub1}\right) : k_{b2} := c_2 \cdot \left(k_2 + k_{ub2}\right) : k_{b3} := c_3 \cdot \left(k_3 + k_{ub3}\right) : k_{b4} := c_4 \cdot \left(k_4 + k_{ub4}\right) :$$

Now we solve for the steady state of active phosphatase and the substrate involved in the covalent modification tier using the conservation expression for the enzyme (PCon), and the substrate respectively (XCon) as shown below.

$$P := solve(simplify(PCon), P) :$$
$$X := solve(XCon, X) :$$

This results in the following steady state correlations between the various variables in terms of the substrate concentrations and $\epsilon$.

$$AK = \frac{c_1 A \epsilon P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$
$$ApK = \frac{c_2 Ap \epsilon P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$
$$AppP = \frac{c_3 App P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$
$$ApP = \frac{c_4 Ap P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$
$$Ks = \frac{\epsilon P_{Total}\left(\frac{X_{Total} M d_a k_{NFB}}{M d_a + d_d} + u_d\right)}{\left(Ap\, c_4 + App\, c_3 + 1\right)\left(S u_a + \frac{X_{Total} M d_a k_{PFB}}{M d_a + d_d} + u_\beta\right)}$$
$$X = \frac{X_{Total} M d_a}{M d_a + d_d}$$
$$Xs = \frac{d_d X_{Total}}{M d_a + d_d}$$
$$K = \frac{\epsilon P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$
$$P = \frac{P_{Total}}{Ap\, c_4 + App\, c_3 + 1}$$

Now depending on the nature of the feedback (i.e. if M is A / App) we perform a slightly different analysis to proceed with our proof as shown below.

## App feedback (M = App)

Here we establish that the feedback (positive or otherwise) is from App by assigning the variable to M.

$M := App$ :

We now solve the remaining equations (ODEs) at steady to obtain expressions for A and Ap in terms of the steady state concentrations of App and $\epsilon$.

$Sol := solve(\{dA, dAp\}, \{A, Ap\})$ :
$assign(Sol)$ :

Thus the previous expressions for variable concentrations at steady state get modified as shown below.

$$simplify(AK) = \frac{App\, c_3\, c_4\, k_3\, k_4\, P_{Total}}{k_1 \left(c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4\right)}$$

$$simplify(ApK) = \frac{App\, c_3\, k_3\, P_{Total}\, c_2\, \epsilon}{c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(AppP) = \frac{c_3\, App\, P_{Total}\, c_2\, \epsilon\, k_2}{c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(ApP) = \frac{c_4\, App\, c_3\, k_3\, P_{Total}}{c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(Ks) = \frac{\left(d_a \left(k_{NFB}\, X_{Total} + u_d\right) App + d_d\, u_d\right) \epsilon^2\, P_{Total}\, k_2\, c_2}{\left(c_3 \left(c_2\, \epsilon\, k_2 + c_4\, k_3\right) App + c_2\, \epsilon\, k_2\right) \left(d_a \left(S\, u_a + k_{PFB}\, X_{Total} + u_\beta\right) App + d_d \left(S\, u_a + u_\beta\right)\right)}$$

$$X = \frac{X_{Total}\, App\, d_a}{d_a\, App + d_d}$$

$$Xs = \frac{d_d\, X_{Total}}{d_a\, App + d_d}$$

$$simplify(K) = \frac{k_2\, c_2\, P_{Total}\, \epsilon^2}{c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(P) = \frac{P_{Total}\, c_2\, \epsilon\, k_2}{c_2\, k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4}$$

Here, similar to the scenario earlier with no feedback, we make note of a couple of insights.
1. We have algebraically solved the system of ODEs at steady state and the only unsolved equations left that define the system are ACon and KCon (the total substrate conservation and kinase conservation)
2. That for any positive real value (non-zero) of both App and $\epsilon$ at steady state, the steady state concentrations of the other variables are also positive and real.

Thus this implies that the steady state concentrations of the system are defined by the positive solution set {App,$\epsilon$} of the equations {ACon, KCon}

Now we rewrite these two expressions as follows (F, G).

$F := simplify(solve(KCon, S))$ :

$G := ACon$ :

Where we rearrange KCon to find an expression F for S (signal) in terms of the other variables (App and $\epsilon$). Thus,

$F = S$ :
$G = 0$ :

Each of the above expressions are functions in App and $\epsilon$. Now differentiating both of these expressions, we have.

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dF}{dApp} \cdot \frac{dApp}{dS}$$

$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dG}{dApp} \cdot \frac{dApp}{dS}$$

Now since we are isolating for biphasic behavior, we begin with the assumption that biphasic behavior exists, i.e. $\frac{dApp}{dS} = 0$ for some finite value of App and $\epsilon$. Thus the above expressions simplify to

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$
$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$

***Note***

This reduction was possible since the denominator of $\frac{dF}{dApp}$ and $\frac{dG}{dApp}$ are both non-zero when App and $\epsilon$ are non-zero and positive, as shown below.

$$\frac{dF}{dApp} = simplify(diff(F, App)) =$$
$$\left(\left(-P_{Total}^{2} X_{Total} c_2^{2} d_a d_d k_2^{2} \left(k_{NFB} + k_{PFB}\right) \epsilon^4 - P_{Total} c_2^{2} k_2 \left(\left(App^2 \left(X_{Total} k_{NFB} + u_d\right) d_a^{2} + 2\, App\, d_d \left(-X_{Total} k_{PFB} + u_d\right) d_a + u_d d_d^{2}\right) \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 - K_{Total} X_{Total} d_a d_d k_2 \left(k_{NFB} + 2 k_{PFB}\right)\right) \epsilon^3 - \left(d_a \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right)^2 k_{PFB} d_d X_{Total} c_2 + P_{Total} \left(App^2 \left(X_{Total} k_{NFB} + u_d\right) d_a^{2} + 2\, App\, d_d \left(-X_{Total} k_{PFB} + u_d\right) d_a + u_d d_d^{2}\right) k_3 c_4 K_{Total} k_2 c_3\right) c_2 \epsilon^2 - 2 d_a \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) k_3 c_2 c_4 k_{PFB} d_d X_{Total} App\, K_{Total} c_3 \epsilon - App^2 k_{PFB} K_{Total}^{2} X_{Total} c_3^{2} c_4^{2} d_a d_d k_3^{2}\right) k_1^{2} + P_{Total} k_3 k_4 c_4 c_3 \left(P_{Total} \left(App^2 \left(X_{Total} k_{NFB} + u_d\right) d_a^{2} + 2\, App\, d_d \left(-X_{Total} k_{PFB} + u_d\right) d_a + u_d d_d^{2}\right) c_2 k_2 \epsilon^2 + 2 d_a \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 k_{PFB} d_d X_{Total} App\, \epsilon + 2\, App^2 k_{PFB} K_{Total} X_{Total} c_3 c_4 d_a d_d k_3\right) k_1 - App^2 c_3^{2} c_4^{2} d_a d_d k_3^{2} k_4^{2} k_{PFB} P_{Total}^{2} X_{Total}\right) \Big/ \left(\left(\left(\epsilon^2 P_{Total} c_2 k_2 - \left(App \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} k_2\right) c_2 \epsilon - App\, K_{Total} c_3 c_4 k_3\right) k_1 + App\, c_3 c_4 k_3 k_4 P_{Total}\right)^{2} \left(d_a App + d_d\right)^{2} u_a\right)$$
$$F = simplify(F) =$$
$$\left(\left(d_a \left(X_{Total} k_{PFB} + u_\beta\right) \left(c_2 \left(K_{Total} k_2 - k_3 P_{Total}\right) \epsilon + K_{Total} c_4 k_3\right) c_3 App^2 + \left(-d_a \left(u_\beta + \left(k_{NFB} + k_{PFB}\right) X_{Total} + u_d\right) c_2 P_{Total} k_2 \epsilon^2 + c_2 \left(u_\beta d_d \left(K_{Total} k_2 - k_3 P_{Total}\right) c_3 + K_{Total} d_a k_2 \left(X_{Total} k_{PFB} + u_\beta\right)\right) \epsilon + u_\beta K_{Total} c_3 c_4 d_d k_3\right) App - \epsilon c_2 \left(P_{Total} \left(u_d + u_\beta\right) \epsilon\right.$$

$$-u_\beta K_{Total}\right) d_d k_2\right) k_1-\left(d_a\left(X_{Total} k_{PFB}+u_\beta\right) App+d_d u_\beta\right) k_3 k_4 c_4 P_{Total} App\, c_3\Big)\Big/\Big(\Big(\Big(-\left(c_2\left(K_{Total} k_2-k_3 P_{Total}\right)\epsilon+K_{Total} c_4 k_3\right) c_3 App+\epsilon c_2 k_2\left(\epsilon P_{Total}-K_{Total}\right)\right) k_1+App\, c_3 c_4 k_3 k_4 P_{Total}\Big)\left(d_a App+d_d\right) u_a\Big)$$

$$\frac{dG}{dApp}=simplify(diff(G, App))=$$

$$\Big(\left(App^2 c_3^2 k_2+\left(\left(2\,App+P_{Total}\right) k_2+k_3 P_{Total}\right) c_3+k_2\right) k_2^2 c_1 k_1 c_2^3\epsilon^4+k_2^2\left(App^2 c_3^2 k_1+2\,App\, k_1\left(App\, c_4+1\right) c_3+2\,App\, k_1 c_4+P_{Total} c_4 k_1+P_{Total} c_4 k_4+k_1\right) c_1 c_3 k_3 c_2^2\epsilon^3+2\,c_4\left(\frac{k_2 k_4\left(App\, c_3+1\right)^2 c_2}{2}+App\, c_1 c_3 k_1 k_3\left(App\, c_3+\frac{1}{2}\,App\, c_4+1\right)\right) k_2 c_3 k_3 c_2\epsilon^2+c_4^2 c_3^2\left(2 k_2 k_4\left(App\, c_3+1\right) c_2+App\, c_1 c_3 k_1 k_3\right) App\, k_3^2\epsilon+App^2 c_3^3 c_4^3 k_3^3 k_4\Big)\Big/\left(k_2 c_1\epsilon^2 k_1 c_2\left(c_2 k_2\left(App\, c_3+1\right)\epsilon+App\, c_3 k_3 c_4\right)^2\right)$$

Note the denominator of $\frac{dF}{dApp}$ is non-zero as signal = F shares the same denominator and for a feasible positive signal this denominator cannot be 0.
The only other way in which the denominator of F = signal could still be 0 and give describe finite Signal is if the numerator is also 0. However the numerator of F can be written as the sum of a linear function of the denominator and constant terms that are never zero (non-zero terms). This is shown below to be the case.

$$Nr\_F:=\left(d_a\left(X_{Total} k_{PFB}+u_\beta\right)\left(c_2\left(K_{Total} k_2-k_3 P_{Total}\right)\epsilon+K_{Total} c_4 k_3\right) c_3 App^2+\left(-d_a\left(u_\beta+\left(k_{NFB}+k_{PFB}\right) X_{Total}+u_d\right) c_2 P_{Total} k_2\epsilon^2+c_2\left(u_\beta d_d\left(K_{Total} k_2-k_3 P_{Total}\right) c_3+K_{Total} d_a k_2\left(X_{Total} k_{PFB}+u_\beta\right)\right)\epsilon+u_\beta K_{Total} c_3 c_4 d_d k_3\right) App-\epsilon c_2\left(P_{Total}\left(u_d+u_\beta\right)\epsilon-u_\beta K_{Total}\right) d_d k_2\right) k_1-\left(d_a\left(X_{Total} k_{PFB}+u_\beta\right) App+d_d u_\beta\right) k_3 k_4 c_4 P_{Total} App\, c_3:$$

$$Dr\_F:=u_a\left(\left(\left(-\left(c_2\left(K_{Total} k_2-k_3 P_{Total}\right)\epsilon+K_{Total} c_4 k_3\right) c_3 App+\epsilon c_2 k_2\left(\epsilon P_{Total}-K_{Total}\right)\right) k_1+App\, c_3 c_4 k_3 k_4 P_{Total}\right)\left(d_a App+d_d\right) u_a\right):$$

$$simplify\left(simplify\left(Nr\_F-\left(-\frac{Dr\_F\cdot\left(u_\beta+\frac{X_{Total}\cdot k_{PFB}\cdot\left(d_a\cdot App\right)}{\left(d_a App+d_d\right)}\right)}{u_a^2}\right.\right.\right.$$

$$\left.\left.\left.-App\, c_2 d_a\epsilon^2 k_1 k_2 k_{NFB} P_{Total} X_{Total}-\left(App\, d_a u_d+d_d u_d\right) k_1 P_{Total} k_2 c_2\epsilon^2\right)\right)\right)$$

$$=0$$

Hence it is also not possible for both the numerator and denominator to be simultaneously zero. Thus the denominator of $\frac{dF}{dApp}$ is non-zero allowing us proceed with the aforementioned simplification.

Now from the nature of the equation above, we can straight away discern that the steady state for a biphasic behavior, $\frac{d\epsilon}{dS}$ cannot be equal to zero. Now, we evaluate $\frac{dG}{d\epsilon}$. We use the inbuilt Maple function *diff* to evaluate these terms, as shown below.

$$\frac{dG}{d\epsilon}=simplify(diff(G,\epsilon))=$$

$$-\left(\left(k_2 c_1 \left(\left(App^2 c_3^2 k_1 + App \left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total}\left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App\, P_{Total} c_3 c_4 k_1 k_3\right) c_2^2 \epsilon^3 + 2 c_4 k_2 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) c_2 \left(App\, c_3 + 1\right) \epsilon^2 + c_4^2 c_3 App\, k_3 \left(4 k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) \epsilon + 2 App^2 c_3^2 c_4^3 k_3^2 k_4\right) c_3 App\, k_3\right) \Big/ \left(k_2 c_1 \epsilon^3 k_1 c_2 \left(c_2 k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3 k_3 c_4\right)^2\right)$$

In the above expression we note that the denominator is always positive (as $\epsilon$ and App can only take positive finite values). Thus the numerator must be independently equal to 0 for biphasic behavior to exist. We isolate the numerator of the above expression as shown below using the inbuilt Maple command *numer*.

$$T := simplify(numer(diff(G, \epsilon))):$$

Hence we have a above condition - relabeled as expression T - must be satisfied at the point of biphasic transition. Thus if we summarize the constraints that define the biphasic transition, we have
1. G = 0 (ACon) [a requirement of steady state]
2. F = S (KCon) [a requirement of steady state]
3. T = 0 (diff(ACon, $\epsilon$)) [following from our assumption of the presence of biphasic behavior in the dose-response graph of App vs S]

Here we can discern two insights again.

First that the expressions that we have here (G = 0 and T = 0) are the same expressions that we observed in the **no feedback** model (Note that this is expected as they arise from the conservation expression of the substrate and thus are not directly linked to the feedback). Thus the same $App_0$ and $\epsilon_0$ that satisfied G and T for the system with no feedback will satisfy the conditions here (based on our assumption that $App_0$ and $\epsilon_0$ are the steady state concentrations of the respective variables at which the system with no feedback obtained a biphasic peak in it's dose-response graph).

$$simplify(G) = \left(k_2 c_1 k_1 c_2^2 \left(\left(App^2 c_3 + \left(1 + \left(-A_{Total} + P_{Total}\right) c_3\right) App - A_{Total}\right) k_2 + App\, P_{Total} c_3 k_3\right) \epsilon^3 + k_2 \left(k_1 \left(c_3 + c_4\right) App + \left(1 + \left(-A_{Total} + P_{Total}\right) c_4\right) k_1 + P_{Total} c_4 k_4\right) c_1 c_3 App\, k_3 c_2 \epsilon^2 + c_4 c_3 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) App\, k_3 \epsilon + App^2 c_3^2 c_4^2 k_3^2 k_4\right) \Big/ \left(k_2 c_1 \epsilon^2 k_1 c_2 \left(c_2 k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3 k_3 c_4\right)\right)$$

$$simplify(T) = -\left(k_2 c_1 \left(\left(App^2 c_3^2 k_1 + App \left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total}\left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App\, P_{Total} c_3 c_4 k_1 k_3\right) c_2^2 \epsilon^3 + 2 c_4 k_2 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) c_2 \left(App\, c_3 + 1\right) \epsilon^2 + c_4^2 c_3 App\, k_3 \left(4 k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) \epsilon + 2 App^2 c_3^2 c_4^3 k_3^2 k_4\right) k_2 c_3 App\, k_3 c_2$$

Second, we can note that F (or rather KCon = 0) has the same first three terms (K + AK + ApK) with only the fourth term (representing the inactive kinase) being different between the system with no feedback and the system with feedback. This expression KCon = 0 is rewritten below. Also note that $K_{Total}$ in this analysis is fixed while the bifurcation is carried out along signal (S) strength. Now if we assume that App and $\epsilon$ take the concentrations $App_0$ and $\epsilon_0$. Then the first three terms are the same as is $K_{Total}$ (between the system with and without feedback).

$$(KCon) = \frac{\epsilon P_{Total}}{\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1} + \frac{App\, c_3 c_4 k_3 k_4 P_{Total}}{c_2 \epsilon k_1 k_2 \left(\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1\right)}$$

$$+\frac{App\,c_3\,k_3\,P_{Total}}{k_2\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)}$$
$$+\frac{\epsilon\,P_{Total}\left(\frac{X_{Total}\,App\,d_a\,k_{NFB}}{d_a\,App+d_d}+u_d\right)}{\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)\left(S\,u_a+\frac{X_{Total}\,App\,d_a\,k_{PFB}}{d_a\,App+d_d}+u_\beta\right)}-K_{Total}$$

Now if there exists and S at which this expression is satisfied then that S represents peak S and in turn we would have satisfied the three conditions defining the presence of a biphasic peak.

Let us for the moment assume that we have negative feedback, i.e. $k_{NFB} = 0$. Then we know that as negative feedback strength increases, the signal strength at which the biphasic peak is observed would need to increase accordingly to satisfy the expression. This increase is unbounded and feasible, i.e. as $k_{NFB}$ approaches Infinity, signal strength would have to approach Infinity to satisfy the expression. We can also see by the form of negative feedback strength appearing in the numerator and S in the denominator (of the Ks term) that a scaling of S will guarantee satisfying the expression i.e. there exists a S value for every negative feedback strength satisfying the equation (for the given $App_0$ and $\epsilon_0$).

On the other hand with positive feedback, i.e. $k_{PFB} = 0$. Increasing positive feedback implies that the signal strength where the biphasic peak occurs should decrease to satisfy the expression, until it can decrease no further and the biphasic peak moves towards negative signal strengths (as was observed in the computational results). However this also reveals that if the basic kinetics produced biphasic behavior with some S value then for an given strength of positive feedback it is possible (but not guaranteed) to observe biphasic 'peak' at a lower S value. However the requirement of S being positive places a constraint on guaranteeing the presence of the behavior for arbitrary ranges of $k_{PFB}$. But note that should the biphasic exist again it would have to exist with the concentrations $App_0$ and $\epsilon_0$ at the peak.

**This allows us to conclude that**
**1. If the modular DSP network presented with biphasic response, then the network with feedback will too (Though with constraints with feedback ranges in PFB - see below).**
**2. The network with positive/negative feedback through activation/deactivation of kinase will have the same App and $\epsilon$ value at the biphasic point (Peak concentration of App in the system)**
**4. This peak S changes with the feedback strength in the following way**
**a. For PFB - As feedback strength increases the peak S moves towards 0, and eventually the system is no longer able to exhibit biphasic behavior.**
**b. For NFB - As feedback strength increases the peak S moves towards infinity.**

Now to unpack characteristics of the saturation value of App (When $K_{Total}$ approaches Infinity) we go back to the steady state description of the model which is now defined by the positive solution set $\{App, \epsilon\}$ of $\{F = S, G = 0\}$. Please note that T=0 is no longer relevant to define the saturation value of App, and the expression for G (ACon) is independent of S and hence remains the same.

$$((KCon)) = \frac{\epsilon\,P_{Total}}{\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1}+\frac{App\,c_3\,c_4\,k_3\,k_4\,P_{Total}}{c_2\,\epsilon\,k_1\,k_2\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)}$$
$$+\frac{App\,c_3\,k_3\,P_{Total}}{k_2\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)}$$

$$+\frac{\epsilon\, P_{Total}\left(\frac{X_{Total}\, App\, d_a\, k_{NFB}}{d_a\, App + d_d} + u_d\right)}{\left(\frac{App\, c_3\, k_3\, c_4}{c_2\, \epsilon\, k_2} + App\, c_3 + 1\right)\left(S\, u_a + \frac{X_{Total}\, App\, d_a\, k_{PFB}}{d_a\, App + d_d} + u_\beta\right)} - K_{Total}$$

Now since S approaches Infinity - Ks (the fourth term in the KCon expression above) approaches zero, erasing the role of the feedback ($k_{NFB}/k_{PFB}$). Thus the saturation concentration of App now defined as solutions of KCon and ACon, which is the same as with or without feedback. Thus,

**3. The saturation value of App is also the same as the system without feedback irrespective of the value of $k_{PFB}/K_{NFB}$.**
Infact, the saturation value will be the same irrespective of the presence or absence of biphasic behavior in the system.

Hence we have proved the four assertions made earlier regarding the nature of the biphasic behavior when the feedback arises from App.

## Ap feedback (M = Ap)

The procedure is similar to that pursued earlier (in the case where the feedback was from App). We begin by first noting M = Ap.

$M := Ap$ :

We now solve the remaining equations (ODEs) at steady to obtain expressions for A and Ap in terms of the steady state concentrations of App and $\epsilon$.

$Sol := solve(\{dA, dAp\}, \{A, Ap\})$ :
$assign(Sol)$ :

Thus the correlations get modified as shown below.

$$simplify(AK) = \frac{App\, c_3\, c_4\, k_3\, k_4\, P_{Total}}{k_1\left(c_2\, k_2\left(App\, c_3 + 1\right)\epsilon + App\, c_3\, k_3\, c_4\right)}$$

$$simplify(ApK) = \frac{App\, c_3\, k_3\, P_{Total}\, c_2\, \epsilon}{c_2\, k_2\left(App\, c_3 + 1\right)\epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(AppP) = \frac{c_3\, App\, P_{Total}\, c_2\, \epsilon\, k_2}{c_2\, k_2\left(App\, c_3 + 1\right)\epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(ApP) = \frac{c_4\, App\, c_3\, k_3\, P_{Total}}{c_2\, k_2\left(App\, c_3 + 1\right)\epsilon + App\, c_3\, k_3\, c_4}$$

$$simplify(Ks) = \left(\left(App\, d_a\, k_3\left(k_{NFB}\, X_{Total} + u_d\right) c_3 + c_2\, d_d\, \epsilon\, k_2\, u_d\right)\epsilon^2\, P_{Total}\, c_2\, k_2\right) \Big/ \left(\left(App\left(c_2\, \epsilon\, k_2 + c_4\, k_3\right) c_3 + c_2\, \epsilon\, k_2\right)\left(App\, d_a\, k_3\left(S\, u_a + k_{PFB}\, X_{Total} + u_\beta\right) c_3 + \epsilon\, c_2\, d_d\, k_2\left(S\, u_a + u_\beta\right)\right)\right)$$

$$X = \frac{X_{Total}\, App\, c_3\, k_3\, d_a}{\left(\frac{App\, c_3\, k_3\, d_a}{c_2\, \epsilon\, k_2} + d_d\right) c_2\, \epsilon\, k_2}$$

$$Xs = \frac{d_d\, X_{Total}}{\frac{App\, c_3\, k_3\, d_a}{c_2\, \epsilon\, k_2} + d_d}$$

$$simplify(K) = \frac{\epsilon^2 P_{Total} c_2 k_2}{c_2 k_2 (App c_3 + 1) \epsilon + App c_3 k_3 c_4}$$
$$simplify(P) = \frac{P_{Total} c_2 \epsilon k_2}{c_2 k_2 (App c_3 + 1) \epsilon + App c_3 k_3 c_4}$$

Here, similar to the scenario earlier with no feedback, we make note of a couple of insights.
1. We have algebraically solved the system of ODEs at steady state and the only unsolved equations left that define the system are ACon and KCon (the total substrate conservation and kinase conservation)
2. That for any positive real value (non-zero) of both App and $\epsilon$ at steady state, the steady state concentrations of the other variables are also positive and real.

Thus this implies that the steady state concentrations of the system are defined by the positive solution set {App,$\epsilon$} of the equations {ACon, KCon}

Now we rewrite these two expressions as follows (F, G).

$F := simplify(solve(KCon, S))$ :
$G := ACon$ :

Where we rearrange KCon to find an expression F for S (signal) in terms of the other variables (substrate concentration and $\epsilon$). Thus,

$F = S$ :
$G = 0$ :

Each of the above expressions are functions in App. Now differentiating both of these expressions, we have.

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dF}{dApp} \cdot \frac{dApp}{dS}$$
$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dG}{dApp} \cdot \frac{dApp}{dS}$$

Now since we are isolating for biphasic behavior, we begin with the assumption that biphasic behavior exists, i.e. $\frac{dApp}{dS} = 0$ for some finite value of App and $\epsilon$. Thus the above expressions simplify to

$$\frac{dF}{dS} = 1 = \frac{dF}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$
$$\frac{dG}{dS} = 0 = \frac{dG}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$

***Note***

This reduction was possible since the denominator of $\frac{dF}{dApp}$ and $\frac{dG}{dApp}$ are both non-zero when App and $\epsilon$ are non-zero and positive, as shown below.

$$\frac{dF}{dApp} = simplify(diff(F, App)) =$$
$$-\left(c_2 \left(\left(c_2^2 P_{Total} d_d k_2^2 \left(-\left(u_d c_2 d_d - X_{Total} d_a \left(k_{NFB} + k_{PFB}\right)\right) P_{Total} k_3 + u_d K_{Total} c_2 d_d k_2\right) \epsilon^4 + 2 c_2^2 k_3 P_{Total} \left(-App P_{Total} c_3 d_a \left(-k_{PFB} X_{Total} + u_d\right) k_3 + K_{Total} \left(App d_a \left(-k_{PFB} X_{Total}\right.\right.\right.\right.\right.\right.$$

$$+ u_d\right) c_3 - \left(\frac{X_{Total}\left(k_{NFB} + 2\,k_{PFB}\right) d_a}{2} + \frac{u_d\, c_4\, d_d}{2}\right) k_2\Bigg)\, d_d\, k_2\, \epsilon^3 + c_2\, d_a\, k_3 \left(-c_3^{\,2}\,\left(-k_{PFB}\, X_{Total}\, d_d\, c_2 + d_a\left(X_{Total}\, k_{NFB} + u_d\right)\right) App^2\, P_{Total}^{\,2}\, k_3^{\,2} + \left(-2\, k_{PFB}\, X_{Total}\, d_d\left(App\, c_3 + 1\right) c_2 + App\, d_a\left(X_{Total}\, k_{NFB} + u_d\right) c_3 + 2\, c_4\, d_d\left(-k_{PFB}\, X_{Total} + u_d\right)\right) K_{Total}\, c_3\, App\, P_{Total}\, k_2\, k_3 + k_{PFB}\, K_{Total}^{\,2}\, X_{Total}\, c_2\, d_d\, k_2^{\,2}\left(App\, c_3 + 1\right)^2\right) \epsilon^2 + d_a\, K_{Total}\, c_3\, k_3^{\,2}\left(c_3\, App\, P_{Total}\left(-2\, k_{PFB}\, X_{Total}\, d_d\, c_2 + d_a\left(X_{Total}\, k_{NFB} + u_d\right)\right) k_3 + 2\, k_{PFB}\, K_{Total}\, X_{Total}\, c_2\, d_d\, k_2\left(App\, c_3 + 1\right)\right) App\, c_4\, \epsilon + App^2\, k_{PFB}\, K_{Total}^{\,2}\, X_{Total}\, c_3^{\,2}\, c_4^{\,2}\, d_a\, d_d\, k_3^{\,3}\right) k_1^{\,2} - k_4\, k_3\left(u_d\, P_{Total}\, c_2^{\,2}\, d_d^{\,2}\, k_2^{\,2}\, \epsilon^3 + 2\, App\, P_{Total}\, c_2\, c_3\, d_a\, d_d\, k_2\, k_3\left(-k_{PFB}\, X_{Total} + u_d\right) \epsilon^2 + d_a\, c_3\, k_3\left(c_3\, App\, P_{Total}\left(-2\, k_{PFB}\, X_{Total}\, d_d\, c_2 + d_a\left(X_{Total}\, k_{NFB} + u_d\right)\right) k_3 + 2\, k_{PFB}\, K_{Total}\, X_{Total}\, c_2\, d_d\, k_2\left(App\, c_3 + 1\right)\right) App\, \epsilon + 2\, App^2\, k_{PFB}\, K_{Total}\, X_{Total}\, c_3^{\,2}\, c_4\, d_a\, d_d\, k_3^{\,2}\right) P_{Total}\, c_4\, k_1 + App^2\, c_3^{\,2}\, c_4^{\,2}\, d_a\, d_d\, k_3^{\,3}\, k_4^{\,2}\, k_{PFB}\, P_{Total}^{\,2}\, X_{Total}\right) c_3\, k_2\, \epsilon\Bigg) \Bigg/ \left(u_a\left(App\, c_3\, k_3\, d_a + d_d\, c_2\, \epsilon\, k_2\right)^2 \left(\left(P_{Total}\, c_2\, k_2\, \epsilon^2 - c_2\left(-App\, P_{Total}\, c_3\, k_3 + K_{Total}\, k_2\left(App\, c_3 + 1\right)\right) \epsilon - App\, K_{Total}\, c_3\, c_4\, k_3\right) k_1 + App\, c_3\, c_4\, k_3\, k_4\, P_{Total}\right)^2\right)$$

$F = simplify(F) =$

$$\Bigg(-\left(\left(-c_2\left(K_{Total}\, k_2 - k_3\, P_{Total}\right) \epsilon - K_{Total}\, c_4\, k_3\right) k_1 + k_4\, P_{Total}\, c_4\, k_3\right) d_a\, k_3\, App^2\left(k_{PFB}\, X_{Total} + u_\beta\right) c_3^{\,2} - c_2\left(\left(\left(-u_\beta\, d_d\left(K_{Total}\, k_2 - k_3\, P_{Total}\right) c_2 + P_{Total}\, d_a\left(u_\beta + \left(k_{NFB} + k_{PFB}\right) X_{Total} + u_d\right) k_3\right) \epsilon - \left(\left(k_{PFB}\, X_{Total} + u_\beta\right) d_a + u_\beta\, c_4\, d_d\right) k_3\, K_{Total}\right) k_1 + k_4\, u_\beta\, P_{Total}\, c_4\, d_d\, k_3\right) App\, k_2\, \epsilon\, c_3 - c_2^{\,2}\, k_1\, d_d\left(P_{Total}\left(u_d + u_\beta\right) \epsilon - u_\beta\, K_{Total}\right) k_2^{\,2}\, \epsilon^2\Bigg) \Bigg/ \left(u_a\left(App\, c_3\, k_3\, d_a + d_d\, c_2\, \epsilon\, k_2\right)\left(\left(\left(-c_2\left(K_{Total}\, k_2 - k_3\, P_{Total}\right) \epsilon - K_{Total}\, c_4\, k_3\right) k_1 + k_4\, P_{Total}\, c_4\, k_3\right) App\, c_3 + \epsilon\, c_2\, k_1\, k_2\left(\epsilon\, P_{Total} - K_{Total}\right)\right)\right)$$

$$\frac{dG}{dApp} = simplify(diff(G, App)) =$$

$$\Bigg(c_1\, k_2^{\,2}\, c_2^{\,3}\left(App^2\, c_3^{\,2}\, k_2 + \left(\left(2\, App + P_{Total}\right) k_2 + k_3\, P_{Total}\right) c_3 + k_2\right) k_1\, \epsilon^4 + c_1\, k_2^{\,2}\left(App^2\, c_3^{\,2}\, k_1 + 2\, App\, k_1\left(App\, c_4 + 1\right) c_3 + 2\, App\, c_4\, k_1 + P_{Total}\, c_4\, k_1 + P_{Total}\, c_4\, k_4 + k_1\right) c_2^{\,2}\, c_3\, k_3\, \epsilon^3 + 2\, c_4\, k_2\, c_2\left(\frac{k_2\, k_4\left(App\, c_3 + 1\right)^2 c_2}{2} + App\, c_1\, c_3\, k_1\, k_3\left(App\, c_3 + \frac{1}{2}\, App\, c_4 + 1\right)\right) c_3\, k_3\, \epsilon^2 + \left(2\, k_2\, k_4\left(App\, c_3 + 1\right) c_2 + App\, c_1\, c_3\, k_1\, k_3\right) c_4^{\,2}\, App\, c_3^{\,2}\, k_3^{\,2}\, \epsilon + App^2\, c_3^{\,3}\, c_4^{\,3}\, k_3^{\,3}\, k_4\Bigg) \Bigg/ \left(c_1\, \epsilon^2\, k_2\, c_2\left(c_2\, k_2\left(App\, c_3 + 1\right) \epsilon + App\, c_3\, k_3\, c_4\right)^2 k_1\right)$$

The only other way in which the denominator of F = signal could still be 0 and give describe finite Signal is if the numerator is also 0. However the numerator of F can be written as the sum of a linear function of the denominator and constant terms that are never zero (non-zero terms). This is shown below to be the case.

$$Nr\_F := -\left(\left(-c_2\left(K_{Total}\, k_2 - k_3\, P_{Total}\right) \epsilon - K_{Total}\, c_4\, k_3\right) k_1 + k_4\, P_{Total}\, c_4\, k_3\right) d_a\, k_3\, App^2\left(k_{PFB}\, X_{Total} + u_\beta\right) c_3^{\,2} - c_2\left(\left(\left(-u_\beta\, d_d\left(K_{Total}\, k_2 - k_3\, P_{Total}\right) c_2 + P_{Total}\, d_a\left(u_\beta + \left(k_{NFB} + k_{PFB}\right) X_{Total} + u_d\right) k_3\right) \epsilon - \left(\left(k_{PFB}\, X_{Total} + u_\beta\right) d_a + u_\beta\, c_4\, d_d\right) k_3\, K_{Total}\right) k_1 + k_4\, u_\beta\, P_{Total}\, c_4\, d_d\, k_3\right) App\, k_2\, \epsilon\, c_3 - c_2^{\,2}\, k_1\, d_d\left(P_{Total}\left(u_d + u_\beta\right) \epsilon - u_\beta\, K_{Total}\right) k_2^{\,2}\, \epsilon^2 :$$

$$Dr\_F := u_a \left(App\, c_3 k_3 d_a + d_d c_2 \epsilon k_2\right) \left(\left(\left(-c_2 \left(K_{Total} k_2 - k_3 P_{Total}\right) \epsilon - K_{Total} c_4 k_3\right) k_1 + k_4 P_{Total} c_4 k_3\right) App\, c_3 + \epsilon c_2 k_1 k_2 \left(\epsilon P_{Total} - K_{Total}\right)\right) :$$

$$simplify\left(simplify\left(Nr\_F - \left(-\frac{Dr\_F \cdot \left(u_\beta + \frac{X_{Total} \cdot k_{PFB} \cdot App\, c_3 k_3 d_a}{App\, c_3 k_3 d_a + d_d c_2 \epsilon k_2}\right)}{u_a} - k_2 c_2 P_{Total} \left(c_2 d_d \epsilon k_2 u_d + App\, c_3 d_a k_3 \left(X_{Total} k_{NFB} + u_d\right)\right) \epsilon^2 k_1\right)\right)\right) = 0$$

Hence it is also not possible for both the numerator and denominator to be simultaneously zero. Thus the denominator of $\frac{dF}{dApp}$ is non-zero allowing us proceed with the aforementioned simplification.

Now from the nature of the equation above, we can straight away discern that the steady state for a biphasic behavior, $\frac{d\epsilon}{dS}$ cannot be equal to zero. Now, we evaluate $\frac{dG}{d\epsilon}$. We use the inbuilt Maple function *diff* to evaluate these terms, as shown below.

$$simplify(diff(G, \epsilon)) = -\left(\left(c_1 k_2 c_2^2 \left(\left(App^2 c_3^2 k_1 + App \left(P_{Total} c_4 k_4 + 2 k_1\right) c_3 + P_{Total} \left(k_1 + k_4\right) c_4 + k_1\right) k_2 - App\, P_{Total} c_3 c_4 k_1 k_3\right) \epsilon^3 + 2 c_4 k_2 c_2 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) \left(App\, c_3 + 1\right) \epsilon^2 + c_4^2 \left(4 k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) App\, c_3 k_3 \epsilon + 2 App^2 c_3^2 c_4^3 k_3^2 k_4\right) App\, c_3 k_3\right) \Big/ \left(c_1 \epsilon^3 k_2 c_2 \left(c_2 k_2 \left(App\, c_3 + 1\right) \epsilon + App\, c_3 k_3 c_4\right)^2 k_1\right)$$

In the above expression we note that the denominator is always positive (as $\epsilon$ and App can only take positive finite values). Thus the numerator must be independently equal to 0 for biphasic behavior to exist. We isolate the numerator of the above expression as shown below using the inbuilt Maple command *numer*.

$$T := simplify(numer(diff(G, \epsilon))) :$$

Hence we have a above condition - relabeled as expression T - must be satisfied at the point of biphasic transition. Thus if we summarize the constraints that define the biphasic transition, we have
1. G = 0 (ACon) [a requirement of steady state]
2. F = S (KCon) [a requirement of steady state]
3. T = 0 (diff(ACon, $\epsilon$)) [following from our assumption of the presence of biphasic behavior in the dose-response graph of App vs S]

Here we can note two insights.

First that the expressions that we have here (G = 0 and T = 0) are the same expressions that we observed in the **no feedback** model (Note that this is expected as they arise from the conservation expression of the substrate and thus are not directly linked to the feedback). Thus the same $App_0$ and $\epsilon_0$ that satisfied G and T for the system with no feedback will satisfy the conditions here (based on our assumption that $App_0$ and $\epsilon_0$ are the steady state concentrations of the respective variables at which the system with no feedback obtained a biphasic peak in it's dose-response graph).

$$simplify(G) = \left(c_1 \left(\left(App^2 c_3 + \left(1 + \left(-A_{Total} + P_{Total}\right) c_3\right) App - A_{Total}\right) k_2 + App\, P_{Total} c_3 k_3\right) k_2 c_2^2 k_1 \epsilon^3 + c_1 \left(k_1 \left(c_3 + c_4\right) App + \left(1 + \left(-A_{Total} + P_{Total}\right) c_4\right) k_1 + P_{Total} c_4 k_4\right) k_2 c_2 App\, c_3 k_3 \epsilon^2 + c_4 \left(k_2 k_4 \left(App\, c_3 + 1\right) c_2 + App\, c_1 c_3 k_1 k_3\right) App\, c_3 k_3 \epsilon + App^2 c_3^2 c_4^2 k_3^2 k_4\right) \Big/$$

$$\left(c_1\,\epsilon^2\,k_2\,c_2\left(c_2\,k_2\left(App\,c_3+1\right)\epsilon+App\,c_3\,k_3\,c_4\right)k_1\right)$$

$$simplify(T)=$$
$$-\left(c_1\,k_2\,c_2^{\,2}\left(\left(App^2\,c_3^{\,2}\,k_1+App\left(P_{Total}\,c_4\,k_4+2\,k_1\right)c_3+P_{Total}\left(k_1+k_4\right)c_4+k_1\right)k_2\right.\right.$$
$$\left.-App\,P_{Total}\,c_3\,c_4\,k_1\,k_3\right)\epsilon^3+2\,c_4\,k_2\,c_2\left(k_2\,k_4\left(App\,c_3+1\right)c_2+App\,c_1\,c_3\,k_1\,k_3\right)\left(App\,c_3$$
$$\left.+1\right)\epsilon^2+c_4^{\,2}\left(4\,k_2\,k_4\left(App\,c_3+1\right)c_2+App\,c_1\,c_3\,k_1\,k_3\right)App\,c_3\,k_3\,\epsilon+2\,App^2\,c_3^{\,2}\,c_4^{\,3}\,k_3^{\,2}\,k_4\Big)$$
$$k_2\,c_2\,App\,c_3\,k_3$$

Second, we can note that F (or rather KCon = 0) has the same first three terms (K + AK + ApK) with only the fourth term (representing the inactive kinase) being different between the system with no feedback and the system with feedback. This expression KCon = 0 is rewritten below. Also note that $K_{Total}$ in this analysis is fixed while the bifurcation is carried out along signal (S) strength. Now if we assume that App and $\epsilon$ take the concentrations $App_0$ and $\epsilon_0$. Then the first three terms are the same as is $K_{Total}$ (between the system with and without feedback).

$$(KCon)=$$
$$\frac{\epsilon\,P_{Total}}{\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1}+\frac{App\,c_3\,c_4\,k_3\,k_4\,P_{Total}}{c_2\,\epsilon\,k_1\,k_2\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)}$$
$$+\frac{App\,c_3\,k_3\,P_{Total}}{k_2\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)}$$
$$+\frac{\epsilon\,P_{Total}\left(\frac{X_{Total}\,App\,c_3\,k_3\,d_a\,k_{NFB}}{\left(\frac{App\,c_3\,k_3\,d_a}{c_2\,\epsilon\,k_2}+d_d\right)c_2\,\epsilon\,k_2}+u_d\right)}{\left(\frac{App\,c_3\,k_3\,c_4}{c_2\,\epsilon\,k_2}+App\,c_3+1\right)\left(S\,u_a+\frac{X_{Total}\,App\,c_3\,k_3\,d_a\,k_{PFB}}{\left(\frac{App\,c_3\,k_3\,d_a}{c_2\,\epsilon\,k_2}+d_d\right)c_2\,\epsilon\,k_2}+u_\beta\right)}-K_{Total}$$

Let us for the moment assume that we have negative feedback, i.e. $k_{NFB} = 0$. Then we know that as negative feedback strength increases, the signal strength at which the biphasic peak is observed would need to increase accordingly to satisfy the expression. This increase is unbounded and feasible, i.e. as $k_{NFB}$ approaches Infinity, signal strength would have to approach Infinity to satisfy the expression. We can also see by the form of negative feedback strength appearing in the numerator and S in the denominator (of the Ks term) that a scaling of S will guarantee satisfying the expression i.e. there exists a S value for every negative feedback strength satisfying the equation (for the given $App_0$ and $\epsilon_0$).

On the other hand with positive feedback, i.e. $k_{PFB} = 0$. Increasing positive feedback implies that the signal strength where the biphasic peak occurs should decrease to satisfy the expression, until it can decrease no further and the biphasic peak moves towards negative signal strengths (as was observed in the computational results). However this also reveals that if the basic kinetics produced biphasic behavior with some S value then for an given strength of positive feedback it is possible (but not guaranteed) to observe biphasic 'peak' at a lower S value. However the requirement of S being positive places a constraint on guaranteeing the presence of the behavior for arbitrary ranges of $k_{PFB}$. But note that should the biphasic exist again it would have to exist with the concentrations $App_0$ and $\epsilon_0$ at the peak.

**This allows us to conclude that**

**1. If the modular DSP network presented with biphasic dose-response, then the network with feedback will too (Except for PFB - see below).**
**2. The network with positive/negative feedback through activation/deactivation of kinase will have the same App and $\epsilon$ value at the biphasic point (Peak concentration of App in the system)**
**4. This peak $K_{Total}$ changes with the feedback strength in the following way**
**a. For PFB - As feedback strength increases the peak S moves towards 0, and eventually the system is no longer able to exhibit biphasic behavior.**
**b. For NFB - As feedback strength increases the peak S moves towards infinity.**

Now to unpack characteristics of the saturation value of App (When $K_{Total}$ approaches Infinity) we go back to the steady state description of the model which is now defined by the positive solution set $\{App, \epsilon\}$ of $\{F = S, G = 0\}$. Please note that T=0 is no longer relevant to define the saturation value of App, and the expression for G (ACon) is independent of S and hence remains the same.

$$((KCon)) = \frac{\epsilon P_{Total}}{\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1} + \frac{App\, c_3 c_4 k_3 k_4 P_{Total}}{c_2 \epsilon k_1 k_2 \left(\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1\right)} + \frac{App\, c_3 k_3 P_{Total}}{k_2 \left(\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1\right)} + \frac{\epsilon P_{Total} \left(\frac{X_{Total}\, App\, c_3 k_3 d_a k_{NFB}}{\left(\frac{App\, c_3 k_3 d_a}{c_2 \epsilon k_2} + d_d\right) c_2 \epsilon k_2} + u_d\right)}{\left(\frac{App\, c_3 k_3 c_4}{c_2 \epsilon k_2} + App\, c_3 + 1\right)\left(S\, u_a + \frac{X_{Total}\, App\, c_3 k_3 d_a k_{PFB}}{\left(\frac{App\, c_3 k_3 d_a}{c_2 \epsilon k_2} + d_d\right) c_2 \epsilon k_2} + u_\beta\right)} - K_{Total}$$

Now since S approaches Infinity - Ks (the fourth term in the KCon expression above) approaches zero, erasing the role of the feedback ($k_{NFB}/k_{PFB}$). Thus the saturation concentration of App now defined as solutions of KCon and ACon, which is the same as with or without feedback. Thus,

**3. The saturation value of App is also the same as the system without feedback irrespective of the value of $k_{PFB}/K_{NFB}$.**
Infact, the saturation value will be the same irrespective of the presence or absence of biphasic behavior in the system.

Hence we have proved the four assertions made earlier regarding the nature of the biphasic behavior when the feedback arises from Ap.

## Feedforward network regulation models (iFFW and cFFW)

This folder deals with results pertaining to the ordered DSP network (with common kinase and common phosphatase) with incoherent and coherent feedforward network regulation by a signal realized through activation/inactivation of the kinase and the phosphatase enzymes (see Fig. 1B(iii, iv)).

We establish that the steady state behavior of the networks (in each case) can be mapped on to the steady state behavior of the isolated ordered DSP (at a different choice of parameters) without any regulation. We initially trace this mapping in steady state behavior and then extend results pertaining to steady state behavior of isolated ordered DSP model for the coherent and incoherent feedforward models.

In addition to this, complementing the computational results presented in the main text with regard to isolas, we show analytically the capacity of the incoherent feedforward system (iFFW) to present isolas in the dose response curve with the signal, while the same is shown to not be possible, analytically, in the coherent feedforward system (cFFW). This analysis is carried out in the enzyme limiting regime (total substrate concentration is much higher than the total enzyme concentrations).

In each manuscript, detailed mathematical models are presented in the beginning (in section 0.) along with instructions on how to run the file followed by specific results.

# Incoherent feedforward model (iFFW)

Incoherent feedforward regulation is a basic way of realizing adaptation in cellular networks. This document details results pertaining to the ordered DSP network (with common kinase and common phosphatase) with incoherent feedforward network regulation by a signal realized through activation of both the kinase and the phosphatase enzymes (see Fig. 1B(iii)).

The document contains key analytical proofs concerning the behavioral characteristics of the iFFW model; Namely 1). we map the steady state behavior of the system on to that of the isolated DSP and in turn characterize its propensity for multi-stationarity, and 2). we analytically show the presence of isolas in the system (when it operates in the enzyme limiting regime, complementing the computational result in the figure 5 in the main text). These results and the detailed mathematical model for the network are organized into subsections as follows.

**0. Model:** The Incoherent feedforward model (iFFW)
**1. Mapping the steady state behavior of iFFW to that of the isolated DSP:** The steady state behavior of the iFFW model can be mapped onto the steady state behavior of the well studied ordered DSP network in isolation (without any enzyme activation inactivation step, i.e. where all free unbound enzyme is assumed to be active). [Run subsection 0 before running this section]
**2. Propensity for multistability:** We delineate strict regions in parameter space where multistability is precluded, enabled and guaranteed with incoherent feedforward regulation.
**3. Presence of isolas:** Parameters values are listed at which the presence of isolas in the dose response curve with signal for the maximally modified substrate form ($A_{pp}$) is shown (in the enzyme limiting regime). [Run subsection 0 before running this section]

The summary of the result is presented in the beginning of each section along with information of the limit (or simplification) used in the model.

**Instructions**
Run the first section (0.) before running only the required section (ignore if you are accessing this on the PDF).

# 0. The Incoherent Feedfoward Model (iFFW)

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

$$restart : with(VectorCalculus) : with(LinearAlgebra) : with(Student[LinearAlgebra]) :$$

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dKs := -u_a \cdot Ks \cdot S + u_d \cdot K - u_\beta \cdot Ks :$$
$$dPs := -k_{IFFW} \cdot Ps \cdot S + d_d \cdot P - d_\beta \cdot Ps :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot Ks \cdot S - u_d \cdot K + u_\beta \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP + k_{IFFW} \cdot Ps \cdot S - d_d \cdot P + d_\beta \cdot Ps :$$

In this model, $u_\beta$ is the kinetic constant associated with the basal activation of the kinase and the phosphatase. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP + Ps - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzyme. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

# 1. Mapping steady state behavior of iFFW to isolated ordered DSP

**Result:** Mapping steady state behavior of iFFW model to that of the isolated DSP model (at a different choice of parameters)
**Limits or simplification:** None

This subsection shows that that the steady state behavior of the iFFW model can be mapped onto the behavior of the isolated DSP (at some different choice of parameters). In order to establish this we show that through a suitable change of variables, it is possible to simplify the iFFW model at steady state to that of the isolated DSP model.

Before proceeding further we present the full model of the isolated DSP with no enzyme activation or inactivation first (system of ODEs and the conservation equations) as shown below. Note that since all free enzyme are assumed to be active in this model, there is no inactive forms of the enzymes 'Ks' and 'Ps'.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1}\, A\, K + (k_{ub1} + k_1)\, AK - k_{b2}\, Ap\, K + (k_{ub2} + k_2)\, ApK :$$
$$dP := -k_{b3}\, App\, P + (k_{ub3} + k_3)\, AppP - k_{b4}\, Ap\, P + (k_{ub4} + k_4)\, ApP :$$

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK - K_{Total} :$$

Now we return to the full model detailed in subsection 0. and we introduce a number of new parameters and variables (which are a function of other existing parameters and variables) as shown below.

Namely we introduce
1. Six new parameters ($k_{b1-pr}$, $k_{b2-pr}$, $k_{b3-pr}$, $k_{b4-pr}$, $\alpha$, $\beta$) which are functions of the old parameters (see below).
2. Two new variables $K_{pr}$ and $P_{pr}$ in place of K and P.

$$\alpha := \frac{u_d}{S \cdot u_a + u_\beta} : \beta := \frac{d_d}{k_{IFFW} \cdot S + d_\beta} :$$

$$k_{b1} := k_{b1-pr} \cdot (1 + \alpha) :$$
$$k_{b2} := k_{b2-pr} \cdot (1 + \alpha) :$$
$$k_{b3} := k_{b3-pr} \cdot (1 + \beta) :$$
$$k_{b4} := k_{b4-pr} \cdot (1 + \beta) :$$

$$K := \frac{K_{pr}}{1 + \alpha} : P := \frac{P_{pr}}{1 + \beta} :$$

Using these newly defined parameters, we solve for the steady state concentrations of the inactive enzymes in the iFFW model (by solving the respective differential equations) as shown below. This is carried out using the *solve* command in Maple which algebraically solves a given equation (first argument) for the variable of concern (second argument). Here we use the ordinary differential equation of the respective inactive enzymes to solve for their respective steady state concentrations as functions of other parameters and species steady state concentrations.

$$Ks := solve(simplify(dKs), Ks) = \frac{K_{pr}\, u_d}{S\, u_a + u_d + u_\beta}$$
$$Ps := solve(simplify(dPs), Ps) = \frac{P_{pr}\, d_d}{S\, k_{IFFW} + d_d + d_\beta}$$

Feeding these expressions back into the ODEs defining the rest of the iFFW model, we notice that the model now bears an exact similarity to the isolated DSP model, with the binding constants replaced by a new constants ($k_{b1-pr}$, $k_{b2-pr}$, $k_{b3-pr}$, $k_{b4-pr}$) and the free enzymes replaced by the new variables $K_{pr}$ and $P_{pr}$. Note that the substitution is done only for the purposes of solving at the steady state.

$$dA = -A\, K_{pr}\, k_{b1\text{-}pr} + AK\, k_{ub1} + ApP\, k_4$$
$$dAp = -Ap\, K_{pr}\, k_{b2\text{-}pr} - Ap\, P_{pr}\, k_{b4\text{-}pr} + AK\, k_1 + ApK\, k_{ub2} + ApP\, k_{ub4} + AppP\, k_3$$
$$dApp = -App\, P_{pr}\, k_{b3\text{-}pr} + ApK\, k_2 + AppP\, k_{ub3}$$
$$dAK = A\, K_{pr}\, k_{b1\text{-}pr} - (k_{ub1} + k_1)\, AK$$
$$dApK = Ap\, K_{pr}\, k_{b2\text{-}pr} - (k_{ub2} + k_2)\, ApK$$
$$dAppP = App\, P_{pr}\, k_{b3\text{-}pr} - (k_{ub3} + k_3)\, AppP$$
$$dApP = Ap\, P_{pr}\, k_{b4\text{-}pr} - (k_{ub4} + k_4)\, ApP$$
$$simplify(dK) = (-A\, k_{b1\text{-}pr} - Ap\, k_{b2\text{-}pr})\, K_{pr} + (k_{ub2} + k_2)\, ApK + (k_{ub1} + k_1)\, AK$$
$$simplify(dP) = (-Ap\, k_{b4\text{-}pr} - App\, k_{b3\text{-}pr})\, P_{pr} + (k_{ub3} + k_3)\, AppP + (k_{ub4} + k_4)\, ApP$$

$$simplify(KCon) = AK + ApK + K_{pr} - K_{Total}$$
$$simplify(PCon) = ApP + AppP + P_{pr} - P_{Total}$$
$$simplify(ACon) = A + Ap + App + AK + ApK + AppP + ApP - A_{Total}$$

Note that the solution of the right hand side of this system of ODEs and the conservation equations would represent steady state solution for the whole system. The exact similarity in the system of ODEs between the isolated network and the reduced form of the iFFW above implies that should there exist a steady state for the isolated system for some choice of parameters (satisfying the system of ODEs and the conservation equation), the same exists for the iFFW model at different choices of parameters (and vice versa).

Moreover this also ensures that the qualitative behavioral landscape of both networks is the same at steady state. This allows us to make a number of comments regarding the iFFW model by leveraging the existing results from literature on the behavior of the isolated DSP as shown in the subsequent subsection.

## 2. Propensity for multistability

*restart* :

**Result:** We delineate sharp boundaries between regions in parameter space where multi-stationarity is guaranteed to occur, where it may and where only monostability is possible.
**Limits or simplification:** None

This subsection shows the presence of clearly defined regions in parameter space where multistability is guaranteed to occur and where it isn't. A key insight here that informs all these results is that a steady state mapping is possible between the iFFW model and the isolated DSP as shown above in subsection 1. This allows us to leverage existing result (Conradi C, Mincheva M. 2014 Catalytic constants enable the emergence of bistability in dual phosphorylation. J. R. Soc. Interface 11: 20140158) about parameters enabling multi-stationarity in the isolated model and extrapolate it for the iFFW model.

We first highlight the result due to Conradi et al., 2014 here for the isolated DSP before moving onto how it is augmented due to our transformation mapping the steady state of iFFW model to the isolated

DSP.

1. $k_1 \cdot k_3 - k_2 \cdot k_4$

2. $k_1 \cdot k_3 \cdot \left( \frac{1}{c_2} + \frac{1}{c_4} \right) - k_2 \cdot k_4 \cdot \left( \frac{1}{c_1} + \frac{1}{c_3} \right)$

*Where* $c_i = \frac{k_{bi}}{(k_{ubi} + k_i)}$ :

**Conradi et al., 2014 showed that**
1. If ex1. < 0 then multistability is guaranteed for some total concentration of substrate and enzymes.
2. If ex1. > 0 and ex2. > 0 then the system is guaranteed to be monostable (irrespective of the values of other parameters and total concentrations of substrate and enzymes).
3. If ex1. > 0 and ex2. < 0 - Then multistability 'may' be possible but isn't guaranteed in parameter space.

Due to our mapping of the steady states of iFFW to the isolated DSP - the qualitative landscape defined by these equations for multistability would hold good for modified parameters in the iFFW model too. In particular, eq1. would remain unchanged as the catalytic constants were not involved in the mapping (refer to subsection 1). However eq2. would change as shown below to involve parameters from the activation and inactivation of the enzymes, and the signal. We first relist the transformations (from subsection 1.) and then show how the above equations are modified.

$$c_1 := \frac{k_{b1-pr}}{(k_{ub1} + k_1)} : c_2 := \frac{k_{b2-pr}}{(k_{ub2} + k_2)} : c_3 := \frac{k_{b3-pr}}{(k_{ub3} + k_3)} : c_4 := \frac{k_{b4-pr}}{(k_{ub4} + k_4)} :$$

$$k_{b1-pr} := \frac{k_{b1}}{(1 + \alpha)} : k_{b2-pr} := \frac{k_{b2}}{(1 + \alpha)} : k_{b3-pr} := \frac{k_{b3}}{(1 + \beta)} : k_{b4-pr} := \frac{k_{b4}}{(1 + \beta)} :$$

$$\alpha := \frac{u_d}{S \cdot u_a + u_\beta} : \beta := \frac{d_d}{k_{IFFW} \cdot S + d_\beta} :$$

The two equations thus become.

1. $simplify\left(k_1 \cdot k_3 - k_2 \cdot k_4\right) = k_1 k_3 - k_2 k_4$

2. $simplify\left(k_1 \cdot k_3 \cdot \left( \frac{1}{c_2} + \frac{1}{c_4} \right) - k_2 \cdot k_4 \cdot \left( \frac{1}{c_1} + \frac{1}{c_3} \right)\right) =$

$$\frac{1}{k_{b2}\left(S u_a + u_\beta\right) k_{b4}\left(k_{IFFW} S + d_\beta\right) k_{b1} k_{b3}} \left( \left( \left(S u_a + u_\beta\right)\left(k_{IFFW} S + d_d + d_\beta\right)\left(k_1 k_3 \left(k_{ub4} + k_4\right) k_{b3} - k_{b4} k_2 k_4 \left(k_{ub3} + k_3\right)\right) k_{b2} + k_{b3} k_{b4} k_1 k_3 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub2} + k_2\right)\right) k_{b1} - k_{b2} k_{b3} k_{b4} k_2 k_4 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub1} + k_1\right)\right)$$

Thus we can make a similar assertion regarding the propensity for multistability in this system.

1. If ex1. < 0 then multi-stationarity and multistability is guaranteed for some total concentration of substrate and enzymes.
2. If ex1. > 0 and ex2. > 0 then the system is guaranteed to be monostable (irrespective of other values of other parameters).
3. If ex1. > 0 and ex2. < 0 then multi-stationarity and multistability 'may' be possible.

Further results can be extracted from this analytical expression for when a system is guaranteed to stay monostable (as some parameter is varied). These could be pursued either by analytical deduction or semi-analytical approaches if the value of a majority of other parameters are known (see section 4 of this workbook).

## 3. Presence of Isolas

**Result:** We show the presence of isolas in the dose response curve with Signal for the maximally modified substrate form ($A_{pp}$), when the system operates in the enzyme limiting regime (complementing the result shown in figure 5 which corresponds to the full model with no assumptions on enzymatic regimes).
**Limits or simplification:** Enzyme limiting regime (Total substrate concentration is much larger than the total enzyme concentrations)

In this section we provide parameter sets where isolas occur in the dose response curve with signal for the maximally modified substrate form while operating in the enzyme limiting regime. In order to do this, we leverage the basic characteristic of an isola (i.e. a closed curve in steady state dose response plot), that the system at steady state must have two distinct set of steady state concentrations that are solutions to the following equation.

$$\frac{dApp}{dS} = 0$$

Note: The presence of just one such solution would indicate the presence of a simple biphasic behavior in the dose-response curve.

In order to show that such an isola exists, we first solve for the steady state of the system by successively solving (algebraically) the system of ODEs and the conservation equations, for steady state concentrations of variables in terms of each other as shown below. We again use the in built Maple *solve* command to perform this operation.

```
Ks := solve(dKs, Ks) :
Ps := solve(dPs, Ps) :

AK := solve(dAK, AK) :
ApK := solve(dApK, ApK) :
AppP := solve(dAppP, AppP) :
ApP := solve(dApP, ApP) :

Sol := solve({dA, dAp}, {A, Ap}) :
assign(Sol)
```

At this stage, we introduce a new variable $\epsilon$ as the ratio of the free (unbound) kinase and the phosphatase as shown below. We also introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

```
k_b1 := c_1·(k_1 + k_ub1) : k_b2 := c_2·(k_2 + k_ub2) : k_b3 := c_3·(k_3 + k_ub3) : k_b4 := c_4·(k_4 + k_ub4) :
K := ϵ·P :
```

We now solve the conservation of the phosphatase to similarly ascertain the steady state concentration of free active phosphatase in terms of other App, $\epsilon$ and parameters.

```
P := simplify(solve(((PCon)), P)) :
```

We now have solved all ODEs (and PCon) to obtain the steady state concentration of every variable as a function of concentrations of App and $\epsilon$. The only two conservation equation that is yet to be solved are ACon_Red and KCon (corresponding to the conservation equation of the substrate in the enzyme limiting regime, and the enzyme kinase).

$$A = \frac{k_4\, c_4\, c_3\, App\, k_3}{\epsilon^2\, k_2\, c_2\, k_1\, c_1}$$

$$Ap = \frac{c_3\, App\, k_3}{\epsilon\, k_2\, c_2}$$

$$AK = \frac{k_4\, c_4\, c_3\, App\, k_3\, P_{Total}\left(S\, k_{IFFW} + d_\beta\right)}{k_1\left(k_2\left(App\left(S\, k_{IFFW} + d_\beta\right) c_3 + S\, k_{IFFW} + d_\beta + d_d\right) c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)\right)}$$

$$ApK = \frac{c_2\, c_3\, App\, k_3\, \epsilon\, P_{Total}\left(S\, k_{IFFW} + d_\beta\right)}{k_2\left(App\left(S\, k_{IFFW} + d_\beta\right) c_3 + S\, k_{IFFW} + d_\beta + d_d\right) c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)}$$

$$AppP = \frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2\left(S\, k_{IFFW} + d_\beta\right)}{k_2\left(App\left(S\, k_{IFFW} + d_\beta\right) c_3 + S\, k_{IFFW} + d_\beta + d_d\right) c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)}$$

$$ApP = \frac{c_4\, c_3\, App\, k_3\, P_{Total}\left(S\, k_{IFFW} + d_\beta\right)}{k_2\left(App\left(S\, k_{IFFW} + d_\beta\right) c_3 + S\, k_{IFFW} + d_\beta + d_d\right) c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)}$$

$$P = \frac{P_{Total}\, \epsilon\, k_2\, c_2\left(S\, k_{IFFW} + d_\beta\right)}{k_2\left(App\left(S\, k_{IFFW} + d_\beta\right) c_3 + S\, k_{IFFW} + d_\beta + d_d\right) c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)}$$

Now as mentioned earlier, in order to have an isola we need two distinct solutions for App, S (and $\epsilon$) where $\frac{dApp}{dS} = 0$.
Thus if we assume there is a isola, then ACon_Red (which has no Signal parameter), would take the following form upon differentiation with S

$$\frac{dACon\_Red}{dS} = 0 = \frac{dACon\_Red}{dApp} \cdot \frac{dApp}{dS} + \frac{dACon\_Red}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$

where,

$$\frac{dACon\_red}{dApp} = simplify(diff(ACon\_Red, App)) = \frac{k_1\, \epsilon\left(\epsilon\, k_2\, c_2 + c_3\, k_3\right) c_1 + k_4\, c_4\, c_3\, k_3}{c_1\, c_2\, k_1\, k_2\, \epsilon^2}$$

$$\frac{dACon\_Red}{d\epsilon} = simplify(diff(ACon\_Red, \epsilon)) = -\frac{c_3\, App\, k_3\left(k_1\, \epsilon\, c_1 + 2\, c_4\, k_4\right)}{\epsilon^3\, k_2\, c_2\, k_1\, c_1}$$

But since $\frac{dApp}{dS} = 0$, and since $\frac{dACon\_Red}{dApp}$ **and** $\frac{dACon\_Red}{d\epsilon}$ are non-zero positive functions of other parameters and variables, the equation $\frac{dACon\_Red}{dS}$ above necessarily implies that $\frac{d\epsilon}{dS} = 0$
Thus, in order to obtain an isola, there exists a steady state satisfying $\frac{d\epsilon}{dS} = \frac{dApp}{dS} = 0$
Thus now differentiating the other conservation equation, KCon, with respect to Signal, we get

$$\frac{dKCon}{dS} = 0 = \frac{dKCon}{dApp} \cdot \frac{dApp}{dS} + \frac{dKCon}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dKCon}{dS}$$

where,

$$\frac{dKCon}{dApp} = simplify(diff(KCon, App)) =$$
$$\left(P_{Total}\, c_3\left(-c_2\, k_1\, k_2\left(S\, k_{IFFW} + d_\beta\right)\left(S\, u_a + u_d + u_\beta\right)\epsilon^2 + k_1\, k_3\left(u_a\, k_{IFFW}\left(c_2 - c_4\right) S^2 + \left(\left(d_d\, u_a + k_{IFFW}\, u_\beta + u_a\, d_\beta\right) c_2 - c_4\left(u_a\, d_\beta + k_{IFFW}\left(u_d + u_\beta\right)\right)\right) S + u_\beta\left(d_\beta + d_d\right) c_2 - d_\beta\, c_4\left(u_d + u_\beta\right)\right)\right.\right.$$
$$\left.\left.\epsilon + k_4\, c_4\, k_3\left(S\, k_{IFFW} + d_d + d_\beta\right)\left(S\, u_a + u_\beta\right)\right) k_2\left(S\, k_{IFFW} + d_\beta\right)\epsilon\, c_2\right) \Big/ \left(k_1\left(S\, u_a + u_\beta\right)\left(\left(k_{IFFW}\left(c_3\, App + 1\right) S + d_\beta\, App\, c_3 + d_\beta + d_d\right) k_2\, c_2\, \epsilon + c_4\, c_3\, App\, k_3\left(S\, k_{IFFW} + d_\beta\right)\right)^2\right)$$

$$\frac{dKCon}{d\epsilon} = simplify(diff(KCon, \epsilon)) =$$

$$-\left(P_{Total}\left(S\,k_{IFFW}+d_{\beta}\right)c_2\left(-k_1\left(App\left(S\,k_{IFFW}+d_{\beta}\right)c_3+S\,k_{IFFW}+d_{\beta}+d_d\right)\left(S\,u_a+u_d+u_{\beta}\right)\epsilon^2\,c_2\,k_2{}^2+k_3\,c_4\,c_3\,App\left(k_4\,App\left(S\,k_{IFFW}+d_{\beta}\right)\left(S\,u_a+u_{\beta}\right)c_3+u_a\,k_{IFFW}\left(-2\,\epsilon\,k_1+k_4\right)S^2+\left(-2\,k_1\left(u_a\,d_{\beta}+k_{IFFW}\left(u_d+u_{\beta}\right)\right)\epsilon+k_4\left(d_d\,u_a+k_{IFFW}\,u_{\beta}+u_a\,d_{\beta}\right)\right)S-2\,d_{\beta}\,k_1\left(u_d+u_{\beta}\right)\epsilon+k_4\,u_{\beta}\left(d_{\beta}+d_d\right)\right)k_2-App^2\,c_3{}^2\,c_4\,k_1\,k_3{}^2\left(S\,k_{IFFW}+d_{\beta}\right)\left(S\,u_a+u_{\beta}\right)\right)\Big/\left(k_1\left(S\,u_a+u_{\beta}\right)\left(k_2\left(App\left(S\,k_{IFFW}+d_{\beta}\right)c_3+S\,k_{IFFW}+d_{\beta}+d_d\right)c_2\,\epsilon+c_4\,c_3\,App\,k_3\left(S\,k_{IFFW}+d_{\beta}\right)\right)^2\right)$$

But since $\frac{dApp}{dS}=\frac{d\epsilon}{dS}=0$, and since the denominators of $\frac{dKCon}{dApp}$ and $\frac{dKCon}{d\epsilon}$ are non-zero positive functions of other parameters and variables, the equation $\frac{dKCon}{dS}$ above necessarily implies that $\frac{dKCon}{dS}=0$

We introduce this condition as a new expression T = 0, as follows.

$T := simplify(diff(KCon, S))$ :

**Thus for an isola to exist, the following three conditions; 1) T = 0, 2) ACon_Red = 0 and 3) KCon = 0 must be satisfied by a set of App, $\epsilon$ and S concentrations.**

### Enzyme Limiting

Now we show that for the following parameter set, the system admits to isolas in the system (when the substrate concentrations are much higher than the enzyme concentrations - enzyme limiting regime)

$k_1 := 0.01 : k_2 := 1 : k_3 := 0.05 : k_4 := 1$ :
$k_{b1} := 1 : k_{b2} := 1 : k_{b3} := 1 : k_{b4} := 1$ :
$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1$ :

$u_a := 3 : u_d := 4 : u_{\beta} := 0.01 : u_{\beta} := 0.01$ :
$k_{IFFW} := 0.25 : d_d := 100 : d_{\beta} := 1$ :

$K_{Total} := 2 : P_{Total} := 5 : A_{Total} := 40$ :

$$c_1 := \frac{k_{b1}}{k_1+k_{ub1}} : c_2 := \frac{k_{b2}}{k_2+k_{ub2}} : c_3 := \frac{k_{b3}}{k_3+k_{ub3}} : c_4 := \frac{k_{b4}}{k_4+k_{ub4}} :$$

We use the inbuilt solve command in Maple to solve the three conditions algebraically to find feasible positive concentrations of S, $\epsilon$ and App; thus proving the presence of isolas in the enzyme limiting regime.

$Sol := solve(\{T, KCon, ACon\_Red\}, \{\epsilon, S, App\}, useassumptions)$ assuming $\epsilon > 0, App > 0, S > 0$

$Sol := \{App = 8.277788649, S = 0.4267648310, \epsilon = 1.132774651\}, \{App = 32.64082959, S = 1.000344891, \epsilon = 4.834708893\}, \{App = 39.22342986, S = 2.543149481, \epsilon = 18.17561275\}$ **(4.1.1)**

Hence we have shown the presence of isolas by showing the presence of multiple points (3) where the system admits to feasible steady states satisfying the a necessary condition of isolas. This data point can be easily checked for validity using computation (bifurcation analysis).

The bifurcation analysis reveals that the three extremum points predicted here correspond to 1). The two extremum points in the dose-response graph (substrate with signal concentration) in the isola's isolated branch of steady states 2). The other branch of steady state which presents with a simple biphasic response; resulting in a dose-response curve of App with signal similar to the one presented in figure 5G for the full system with incoherent feedforward regulation.

# Coherent feedforward model (cFFW)

The coherent feedforward regulation is associated with reinforcement characteristics and often results in qualitative strengthening of the various behaviors exhibited by a network. This document details results pertaining to the ordered DSP network (with common kinase and common phosphatase) with coherent feedforward network regulation realized through activation and deactivation of the kinase and the phosphatase enzyme by a common signal respectively (see Fig. 1B(iv)).

The document contains key analytical proofs concerning the behavioral characteristics of the cFFW model; Namely 1). we map the steady state behavior of the system on to that of the isolated DSP and in turn characterize its propensity for multi-stationarity, and 2). we analytically show the absence of isolas in the system when it operates in the enzyme limiting regime (Total substrate concentration is much greater than the total enzyme concentrations). These results and the detailed mathematical model for the network are organized into subsections as follows.

**0. Model:** The coherent feedforward model (cFFW)
**1. Mapping the steady state behavior of cFFW to that of the isolated DSP:** The steady state behavior of the cFFW model can be mapped onto the steady state behavior of the well studied ordered DSP network in isolation (without any enzyme activation inactivation step, i.e. where all free unbound enzyme is assumed to be active). [Run subsection 0 before running this section]
**2. Propensity for multistability:** We delineate strict regions in parameter space where multistability is precluded, enabled and guaranteed with coherent feedforward regulation.
**3. Absence of isolas:** Isolas in the dose response curve with Signal for the maximally modified substrate form ($A_{pp}$) can be ruled out in the cFFW model (in the enzyme limiting regime). [Run subsection 0 before running this section]

The summary of the result is presented in the beginning of each section along with information of the limit (or simplification) used in the model.

**Instructions**
Run the first section (0.) before running only the required section (ignore if you are accessing this on the PDF).

# 0. The Coherent Feedfoward Model (cFFW)

We initialize the Maple file with the *restart* command and load the relevant libraries of inbuilt Maple functions (*LinearAlgebra, VectorCalculus, Student[LinearAlgebra]*)

```
restart :
with(VectorCalculus) : with(LinearAlgebra) : with(Student[LinearAlgebra]) :
```

The system is modeled as a set of ODEs using the kinetic nomenclature described in the main text (Models and Methods section). Here dA represents d[A]/dt and similarly in the case of the other variables. At steady state thus, the right hand sides of each of these expressions will be equal to zero.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dKs := -u_a \cdot Ks \cdot S + u_d \cdot K - u_b \cdot Ks :$$
$$dPs := k_{CFFW} \cdot P \cdot S + d_d \cdot P - d_b \cdot Ps :$$

$$dK := -k_{b1} A K + (k_{ub1} + k_1) AK - k_{b2} Ap K + (k_{ub2} + k_2) ApK + u_a \cdot Ks \cdot S - u_d \cdot K + u_b \cdot Ks :$$
$$dP := -k_{b3} App P + (k_{ub3} + k_3) AppP - k_{b4} Ap P + (k_{ub4} + k_4) ApP - k_{CFFW} \cdot P \cdot S - d_d \cdot P + d_b \cdot Ps :$$

In this model, $u_b$ and $d_b$ the kinetic constant associated with the basal activation of the kinase and the phosphatase in the absence of signal. The model is also associated with conservation conditions which are described below. Here we store the conservation expressions as ACon, PCon and KCon for the substrate and the respective enzymes. Each of these expressions is always equal to zero (both in the transient and at steady state).

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP + Ps - P_{Total} :$$
$$KCon := K + AK + ApK + Ks - K_{Total} :$$

Some of the results that follow rely on an assumption that the system is operating in an enzyme limiting regime, i.e. The total concentration of substrate is significantly higher than the total concentration of the enzymes. Thus under this condition ACon from above can be approximated as ACon_Red (shown below) by ignoring the complex concentrations. Again this expression is always equal to zero (in this limit), both in the transient and at steady state.

$$ACon\_Red := A + Ap + App - A_{Total} :$$

# 1. Mapping steady state behavior of cFFW to isolated ordered DSP

**Result:** Mapping steady state behavior of cFFW model to that of the isolated DSP model (at a different choice of parameters)

**Limits or simplification:** None

This subsection shows that that the steady state behavior of the cFFW model can be mapped onto the behavior of the isolated DSP (at some different choice of parameters). In order to establish this we show that through a suitable change of variables, it is possible to simplify the cFFW model at steady state to that of the isolated DSP model.

Before proceeding further we present the full model of the isolated DSP with no enzyme activation or inactivation first (system of ODEs and the conservation equations) as shown below. Note that since all unbound enzymes are assumed to be active in this model, there is no 'Ks' and 'Ps'.

$$dA := k_4 \cdot ApP + k_{ub1} \cdot AK - k_{b1} \cdot A \cdot K :$$
$$dAp := k_1 \cdot AK + k_3 \cdot AppP + k_{ub2} \cdot (ApK) + k_{ub4} \cdot (ApP) - k_{b2} \cdot (Ap) \cdot (K) - k_{b4} \cdot Ap \cdot P :$$
$$dApp := k_2 \cdot ApK + k_{ub3} \cdot AppP - k_{b3} \cdot App \cdot P :$$

$$dAK := k_{b1} \cdot A \cdot K - (k_{ub1} + k_1) \cdot AK :$$
$$dApK := k_{b2} \cdot Ap \cdot K - (k_{ub2} + k_2) \cdot ApK :$$
$$dAppP := k_{b3} \cdot App \cdot P - (k_{ub3} + k_3) \cdot AppP :$$
$$dApP := k_{b4} \cdot Ap \cdot P - (k_{ub4} + k_4) \cdot ApP :$$

$$dK := -k_{b1}\, A\, K + (k_{ub1} + k_1)\, AK - k_{b2}\, Ap\, K + (k_{ub2} + k_2)\, ApK :$$
$$dP := -k_{b3}\, App\, P + (k_{ub3} + k_3)\, AppP - k_{b4}\, Ap\, P + (k_{ub4} + k_4)\, ApP :$$

$$ACon := A + Ap + App + AK + ApK + AppP + ApP - A_{Total} :$$
$$PCon := P + AppP + ApP - P_{Total} :$$
$$KCon := K + AK + ApK - K_{Total} :$$

Now we return to the full model detailed in subsection 0. and we introduce a number of new parameters and variables (which are a function of other existing parameters and variables) as shown below.

Namely we introduce
1. Six new parameters ($k_{b1-pr}$, $k_{b2-pr}$, $k_{b3-pr}$, $k_{b4-pr}$, $\alpha$, $\beta$) which are functions of the old parameters (see below).
2. Two new variables $K_{pr}$ and $P_{pr}$ in place of K and P.

$$\alpha := \frac{u_d}{S \cdot u_a + u_b} : \beta := \frac{k_{CFFW} \cdot S + d_d}{d_b} :$$
$$K := \frac{K_{pr}}{1 + \alpha} : P := \frac{P_{pr}}{1 + \beta} :$$
$$k_{b1} := k_{b1-pr} \cdot (1 + \alpha) :$$
$$k_{b2} := k_{b2-pr} \cdot (1 + \alpha) :$$
$$k_{b3} := k_{b3-pr} \cdot (1 + \beta) :$$
$$k_{b4} := k_{b4-pr} \cdot (1 + \beta) :$$

Using these newly defined parameters, we solve for the steady state concentrations of the inactive enzymes in the cFFW model (by solving the respective differential equations) as shown below. This is carried out using the *solve* command in Maple which algebraically solves a given equation (first argument) for the variable of concern (second argument). Here we use the ordinary differential equation of the respective inactive enzymes to solve for their respective steady state concentrations as functions of other parameters and species steady state concentrations.

$$Ks := solve(simplify(dKs), Ks) = \frac{K_{pr}\, u_d}{S\, u_a + u_b + u_d}$$
$$Ps := solve(simplify(dPs), Ps) = \frac{P_{pr}\left(S\, k_{CFFW} + d_d\right)}{S\, k_{CFFW} + d_b + d_d}$$

Feeding these expressions back into the ODEs defining the rest of the cFFW model we notice that the model now bears a striking similarity to the isolated DSP model, with the binding constants replaced by a new constants ($k_{b1-pr}, k_{b2-pr}, k_{b3-pr}, k_{b4-pr}$) and and the free enzymes replaced by the new variables $K_{pr}$ and $P_{pr}$. However the requirement of S being positive places a constraint on guarenteeing the presence of the behavior for arbritrary ranges of $k_{PFB}$.

$$dA = -A\, K_{pr}\, k_{b1\text{-}pr} + AK\, k_{ub1} + ApP\, k_4$$
$$dAp = -Ap\, K_{pr}\, k_{b2\text{-}pr} - Ap\, P_{pr}\, k_{b4\text{-}pr} + AK\, k_1 + ApK\, k_{ub2} + ApP\, k_{ub4} + AppP\, k_3$$
$$dApp = -App\, P_{pr}\, k_{b3\text{-}pr} + ApK\, k_2 + AppP\, k_{ub3}$$
$$dAK = A\, K_{pr}\, k_{b1\text{-}pr} - \left(k_{ub1} + k_1\right) AK$$
$$dApK = k_{b2\text{-}pr}\, Ap\, K_{pr} - \left(k_{ub2} + k_2\right) ApK$$
$$dAppP = k_{b3\text{-}pr}\, App\, P_{pr} - \left(k_{ub3} + k_3\right) AppP$$
$$dApP = k_{b4\text{-}pr}\, Ap\, P_{pr} - \left(k_{ub4} + k_4\right) ApP$$
$$simplify(dK) = \left(-A\, k_{b1\text{-}pr} - Ap\, k_{b2\text{-}pr}\right) K_{pr} + \left(k_{ub2} + k_2\right) ApK + \left(k_{ub1} + k_1\right) AK$$
$$simplify(dP) = \left(-Ap\, k_{b4\text{-}pr} - App\, k_{b3\text{-}pr}\right) P_{pr} + \left(k_{ub3} + k_3\right) AppP + \left(k_{ub4} + k_4\right) ApP$$

$$simplify(KCon) = AK + ApK + K_{pr} - K_{Total}$$
$$simplify(PCon) = ApP + AppP + P_{pr} - P_{Total}$$
$$simplify(ACon) = A + Ap + App + AK + ApK + AppP + ApP - A_{Total}$$

Note that the solution of this system of ODEs and the conservation equations would represent steady state solution for the whole system. The exact similarity in the system of ODEs between the isolated network and the reduced form of the cFFW above implies that should there exist a steady state for the isolated system for some choice of parameters (satisfying the system of ODEs and the conservation equation), the same exists for the cFFW model at different choices of parameters (and vice versa).

Moreover this also ensures that the qualitative behavioral landscape of both networks is the same at steady state. This allows us to make a number of comments regarding the cFFW model by leveraging the existing results from literature on the behavior of the isolated DSP as shown in the subsequent subsection.

## 2. Propensity for multistability

*restart* :

**Result:** We delineate sharp boundaries between regions in parameter space where multi-stationarity is guaranteed to occur, where it may and where only mononstability is possible.
**Limits or simplification:** None

This subsection shows the presence of clearly defined regions in parameter space where multistability is guaranteed to occur and where it isn't. A key insight here that informs all these results is that a steady state mapping is possible between the cFFW model and the isolated DSP as shown above in subsection 1. This allows us to leverage existing result (Conradi C, Mincheva M. 2014 Catalytic constants enable the emergence of bistability in dual phosphorylation. J. R. Soc. Interface 11: 20140158) about parameters enabling multi-stationarity in the isolated model and extrapolate it for the cFFW model.

We first highlight the result due to Conradi et al., 2014 here for the isolated DSP before moving onto how it is augmented due to our transformation mapping the steady state of cFFW model to the isolated

DSP.

1. $k_1 \cdot k_3 - k_2 \cdot k_4$

2. $k_1 \cdot k_3 \cdot \left( \frac{1}{c_2} + \frac{1}{c_4} \right) - k_2 \cdot k_4 \cdot \left( \frac{1}{c_1} + \frac{1}{c_3} \right)$

*Where* $c_i = \frac{k_{bi}}{(k_{ubi} + k_i)}$ :

**Conradi et al., 2014 showed that**
1. If ex1. < 0 then multistability is guaranteed for some total concentration of substrate and enzymes.
2. If ex1. > 0 and ex2. > 0 then the system is guaranteed to be monostable (irrespective of the values of other parameters and total concentrations of substrate and enzymes).
3. If ex1. > 0 and ex2. < 0 - Then multistability 'may' be possible but isn't guaranteed in parameter space.

Due to our mapping of the steady states of cFFW to the isolated DSP - the qualitative landscape defined by these equations for multistability would hold good for modified parameters in the cFFW model too. In particular, eq1. would remain unchanged as the catalytic constants were not involved in the mapping (refer to subsection 1). However eq2. would change as shown below to involve parameters from the activation and inactivation of the enzymes, and the signal. We first relist the transformations (from subsection 1.) and then show how the above equations are modified.

$$c_1 := \frac{k_{b1-pr}}{(k_{ub1} + k_1)} : c_2 := \frac{k_{b2-pr}}{(k_{ub2} + k_2)} : c_3 := \frac{k_{b3-pr}}{(k_{ub3} + k_3)} : c_4 := \frac{k_{b4-pr}}{(k_{ub4} + k_4)} :$$

$$k_{b1-pr} := \frac{k_{b1}}{(1+\alpha)} : k_{b2-pr} := \frac{k_{b2}}{(1+\alpha)} : k_{b3-pr} := \frac{k_{b3}}{(1+\beta)} : k_{b4-pr} := \frac{k_{b4}}{(1+\beta)} :$$

$$\alpha := \frac{u_d}{S \cdot u_a + u_b} : \quad \beta := \frac{k_{CFFW} \cdot S + d_d}{d_b} :$$

The two equations thus become.

1. $simplify\left(k_1 \cdot k_3 - k_2 \cdot k_4\right) = k_1 k_3 - k_2 k_4$

2. $simplify\left(k_1 \cdot k_3 \cdot \left( \frac{1}{c_2} + \frac{1}{c_4} \right) - k_2 \cdot k_4 \cdot \left( \frac{1}{c_1} + \frac{1}{c_3} \right)\right) =$

$$\frac{1}{k_{b2}\,(S\,u_a + u_b)\,k_{b4}\,d_b\,k_{b1}\,k_{b3}} \left( \left( \left( k_{CFFW}\,S + d_b + d_d \right) \left( S\,u_a + u_b \right) \left( k_1\,k_3\,\left( k_{ub4} + k_4 \right) k_{b3} \right. \right. \right.$$
$$\left. - k_{b4}\,k_2\,k_4\,\left( k_{ub3} + k_3 \right) \right) k_{b2} + k_{b3}\,k_{b4}\,d_b\,k_1\,k_3\,\left( k_{ub2} + k_2 \right) \left( S\,u_a + u_b + u_d \right) \Big) k_{b1}$$
$$\left. - k_{b2}\,k_{b3}\,k_{b4}\,d_b\,k_2\,k_4\,\left( k_{ub1} + k_1 \right) \left( S\,u_a + u_b + u_d \right) \right)$$

Thus we can make a similar assertion regarding the propensity for multistability in this system.

1. If ex1. < 0 then multi-stationarity and multistability is guaranteed for some total concentration of substrate and enzymes.
2. If ex1. > 0 and ex2. > 0 then the system is guaranteed to be monostable (irrespective of other values of other parameters).
3. If ex1. > 0 and ex2. < 0 then multi-stationarity and multistability 'may' be possible.

Further results can be extracted from this analytical expression for when a system is guaranteed to stay monostable (as some parameter is varied). These could be pursued either by analytical deduction or semi-analytical approaches if the value of a majority of other parameters are known (see section 4 of this workbook).

## 3. Absence of Isolas

**Result:** Isolas in the dose response curve with Signal for the maximally modified substrate form ($A_{pp}$) can be ruled out in the cFFW model in the enzyme limiting regime.
**Limits or simplification:** Enzyme limiting regime (Total substrate concentration is greater than the total enzyme concentrations)

In this section we prove the absence of such an isola across parameter space when the system operates in an enzyme limiting regime. In order to do this, we leverage the basic characteristic of an isola, i.e. a closed curve in steady state dose response plot; That the system at steady state must have two distinct set of steady state concentrations that are solutions to the following equation.

$$\frac{dApp}{dS} = 0$$

Note: The presence of just one such solution would indicate the presence of a simple biphasic behavior in the dose-response curve.

In order to show that such multiple solutions cannot exist (that an isola cannot exist), we first solve for the steady state of the system by successively solving (algebraically) the system of ODEs and the conservation equations, for steady state concentrations of variables in terms of each other as shown below. We again use the in built Maple *solve* command to perform this operation.

```
Ks := solve(dKs, Ks) :
Ps := solve(dPs, Ps) :

AK := solve(dAK, AK) :
ApK := solve(dApK, ApK) :
AppP := solve(dAppP, AppP) :
ApP := solve(dApP, ApP) :

Sol := solve({dA, dAp}, {A, Ap}) :
assign(Sol)
```

At this stage, we introduce a new variable $\epsilon$ as the ratio of the free (unbound) kinase and the phosphatase as shown below. We also introduce the following parameters ($c_1$, $c_2$, $c_3$, and $c_4$). This is done for the sake of brevity and easy tractability of the expressions obtained.

```
k_b1 := c_1·(k_1 + k_ub1) : k_b2 := c_2·(k_2 + k_ub2) : k_b3 := c_3·(k_3 + k_ub3) : k_b4 := c_4·(k_4 + k_ub4) :
K := ϵ·P :
```

We now solve the conservation of the phosphatase to similarly ascertain the steady state concentration of free active phosphatase in terms of App, $\epsilon$ and parameters.

```
P := simplify(solve((numer(PCon)), P)) :
```

We now have solved all ODEs to obtain the steady state concentration of every variable as a function of concentrations of App and $\epsilon$ as shown below.

$$A = \frac{k_4\, c_4\, c_3\, App\, k_3}{\epsilon^2\, k_2\, c_2\, k_1\, c_1}$$

$$Ap = \frac{c_3\, App\, k_3}{\epsilon\, k_2\, c_2}$$

$$AK = \frac{k_4\, c_4\, c_3\, App\, k_3\, P_{Total}\, d_b}{k_1\left(c_2\, k_2\left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\epsilon + App\, c_3\, c_4\, d_b\, k_3\right)}$$

$$ApK = \frac{c_2\, c_3\, App\, k_3\, \epsilon\, P_{Total}\, d_b}{c_2\, k_2\left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\epsilon + App\, c_3\, c_4\, d_b\, k_3}$$

$$AppP = \frac{c_3\, App\, P_{Total}\, \epsilon\, k_2\, c_2\, d_b}{c_2\, k_2\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\, \epsilon + App\, c_3\, c_4\, d_b\, k_3}$$

$$ApP = \frac{c_4\, c_3\, App\, k_3\, P_{Total}\, d_b}{c_2\, k_2\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\, \epsilon + App\, c_3\, c_4\, d_b\, k_3}$$

$$P = \frac{P_{Total}\, \epsilon\, k_2\, c_2\, d_b}{c_2\, k_2\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\, \epsilon + App\, c_3\, c_4\, d_b\, k_3}$$

The only two conservation equation that is yet to be solved are ACon_Red and KCon (corresponding to the conservation equation of the substrate in the enzyme limiting regime, and the enzyme kinase). Now as mentioned earlier, in order to have an isola we need to distinct solutions for App, S (and $\epsilon$) where $\frac{dApp}{dS} = 0$.

Thus if we assume there is an isola, then ACon_Red, which has no Signal parameter would take the following form upon differentiation with S

$$\frac{dACon\_Red}{dS} = 0 = \frac{dACon\_Red}{dApp} \cdot \frac{dApp}{dS} + \frac{dACon\_Red}{d\epsilon} \cdot \frac{d\epsilon}{dS}$$

where,

$$\frac{dACon\_red}{dApp} = simplify\left(diff\left(ACon\_Red, App\right)\right) = \frac{k_1\, \epsilon\, \left(\epsilon\, k_2\, c_2 + c_3\, k_3\right)\, c_1 + k_4\, c_4\, c_3\, k_3}{c_1\, c_2\, k_1\, k_2\, \epsilon^2}$$

$$\frac{dACon\_Red}{d\epsilon} = simplify\left(diff\left(ACon\_Red, \epsilon\right)\right) = -\frac{c_3\, App\, k_3\, \left(k_1\, \epsilon\, c_1 + 2\, c_4\, k_4\right)}{\epsilon^3\, k_2\, c_2\, k_1\, c_1}$$

But since $\frac{dApp}{dS} = 0$, and since $\frac{dACon\_Red}{dApp}$ and $\frac{dACon\_Red}{d\epsilon}$ are non-zero positive functions of other parameters and variables, the equation $\frac{dACon\_Red}{dS}$ above necessarily implies that $\frac{d\epsilon}{dS} = 0$

Thus, in order to obtain an isola, there exists a steady state satisfying $\frac{d\epsilon}{dS} = \frac{dApp}{dS} = 0$

Thus now differentiating the other conservation equation yet remaining, KCon, with Signal, we get

$$\frac{dKCon}{dS} = 0 = \frac{dKCon}{dApp} \cdot \frac{dApp}{dS} + \frac{dKCon}{d\epsilon} \cdot \frac{d\epsilon}{dS} + \frac{dKCon}{dS}$$

where,

$$\frac{dKCon}{dApp} = simplify\left(diff\left(KCon, App\right)\right) =$$
$$\left(k_2\, c_2\, \left(-c_2\, d_b\, k_1\, k_2\, \left(S\, u_a + u_b + u_d\right)\, \epsilon^2 + k_3\, \left(\left(S\, k_{CFFW} + d_b + d_d\right)\, \left(S\, u_a + u_b\right)\, c_2 - c_4\, d_b\, \left(S\, u_a + u_b + u_d\right)\right)\, k_1\, \epsilon + k_3\, k_4\, c_4\, \left(S\, k_{CFFW} + d_b + d_d\right)\, \left(S\, u_a + u_b\right)\right)\, d_b\, \epsilon\, c_3\, P_{Total}\right) \Big/ \left(\left(c_2\, k_2\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\, \epsilon + App\, c_3\, c_4\, d_b\, k_3\right)^2 k_1\, \left(S\, u_a + u_b\right)\right)$$

$$\frac{dKCon}{d\epsilon} = simplify\left(diff\left(KCon, \epsilon\right)\right) =$$
$$-\left(\left(-\epsilon^2\, c_2\, k_1\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b + d_d\right)\, \left(S\, u_a + u_b + u_d\right)\, k_2^2 + \left(-2\, d_b\, k_1\, \left(S\, u_a + u_b + u_d\right)\, \epsilon + k_4\, \left(\left(App\, c_3 + 1\right)\, d_b + S\, k_{CFFW} + d_d\right)\, \left(S\, u_a + u_b\right)\right)\, App\, k_3\, c_4\, c_3\, k_2 - k_3^2\, App^2\, c_3^2\, c_4\, d_b\, k_1\, \left(S\, u_a + u_b\right)\right)\, c_2\, d_b\, P_{Total}\right) \Big/ \left(\left(c_2\, k_2\, \left(App\, c_3\, d_b + S\, k_{CFFW} + d_b\right.\right.\right.$$

$$+ d_d)\,\epsilon + App\,c_3\,c_4\,d_b\,k_3)^2\,k_1\,(S\,u_a + u_b))$$

But since $\frac{dApp}{dS} = \frac{d\epsilon}{dS} = 0$, and since the denominators of $\frac{dKCon}{dApp}$ and $\frac{dKCon}{d\epsilon}$ are non-zero positive functions of other parameters and variables, the equation $\frac{dKCon}{dS}$ above necessarily implies that $\frac{dKCon}{dS} = 0$

However, $\frac{dKCon}{dS}$ is never equal to zero for feasible (positive and real) values of the variables and parameters as seen below.

$$\frac{dKCon}{dS} = simplify(diff(KCon, S)) =$$
$$-\left(\left(\left(S^2\,{u_a}^2 + 2\,S\,(u_b + u_d)\,u_a + u_b\,(u_b + u_d)\right)\,k_{CFFW} + u_a\,u_d\,(App\,c_3\,d_b + d_b + d_d)\right)\,k_2\,c_2\,k_1\,\epsilon^2 + App\,k_3\,\left(k_{CFFW}\,(S\,u_a + u_b)^2\,c_2 + u_a\,u_d\,c_4\,d_b\right)\,k_1\,c_3\,\epsilon + k_4\,c_4\,c_3\,App\,k_3\,k_{CFFW}\,(S\,u_a + u_b)^2\right)\,k_2\,c_2\,d_b\,\epsilon\,P_{Total}\Big)\Big/\left((c_2\,k_2\,(App\,c_3\,d_b + S\,k_{CFFW} + d_b + d_d)\,\epsilon + App\,c_3\,c_4\,d_b\,k_3)^2\,k_1\,(S\,u_a + u_b)^2\right)$$

**Hence, we have a contradiction with our earlier assumption that there exists an isola. Thus, an isola isn't possible in the cFFW model operating in the enzyme limiting regime.**

Limiting enzyme regime thus constrains the coherent feedfoward network from satisfying the necessary condition for an isola to exist (which is also a necessary condition for biphasic behavior in the dose response curve). However moving away from this enzyme regime, the full network allows for both biphasic and multiphasic behavior to be seen in the dose response curve -- see fig 6B in the main text (isolas however have not been observed to the best of our knowledge).

## Semi-analytical Approaches

In this section we detail a couple of illustrative examples of how semi-analytical approaches can be brought to bear in dissecting systems behavior, while dealing with the complexity of multisite modification level and network level regulation.

Throughout this *Maple* workbook, we have derived various analytical expressions which are simplified expressions representing the steady state involving concentrations of variables as functions of each other and parameters (themselves significantly reduced in complexity from the original system). These are used to
1. Predict the presence and features of a qualitative behavior, such as biphasic dose-responses or multi-stationarity (in parameter space)
2. Reveal regions in parameter space where the system is incapable of exhibiting these qualitative behaviors.

In a number of instances it might be relevant to ascertain the role or effect of one single parameter on the presence, absence or features of a behavior. This may in some instances be easy to ascertain when the resulting analytical expression (representing the simplified description of the system at steady state) is tractable with respect to the single parameter of concern. However in most cases, the expressions appear in a complicated manner as coupled polynomials. In such instances, specification of a number of parameters can enable us to simplify the polynomials and allow us to obtain clear insights on the influence of certain key parameters of focus on the possibility, impossibility, presence and features of behaviors of interest.

We call this approach semi-analytical, where the insights aren't strictly binary as they are in the analytical approach, but nevertheless reveal sharp insights in huge swaths of parameter space, going beyond computational insight which is relevant to a specific parameter set and its close vicinity.

We show two examples of such processes here using the the conditions obtained for incoherent feed-forward and the positive feedback from App with the ordered DSP.

The two examples shown here are only illustrative and not exhaustive. Similar approaches can be utilized in other models, and more broadly this line of thinking can be used to answer a number of questions about systems behavior when the resulting system is highly complicated for traditional analytical approaches. To aid the readers, key equations (often the steady state simplified form of networks) are highlighted in the proofs (highlighted in yellow) included in this document to indicate their propensity to be adopted towards these semi-analytical approaches (for steady state analysis).

# Incoherent feedforward network: Understanding the role of feedforward strength in enabling and disabling multi-stationarity

**Recapitulation**
We earlier obtained the following two equations (refer to sub-section 2 of 3.2.1) and their associated conditions to determine if the ordered double site phosphorylation with incoherent feedfoward regulation through kinase and phophatase activation is guaranteed to exhibit either multi-stationarity or is incapable of multi-stationarity. The equations (rewritten below) involve a combination of network parameters and feedforward strength.

1. $k_1 k_3 - k_2 k_4$ :

2. $\left(\left(S u_a + u_\beta\right)\left(k_{IFFW} S + d_d + d_\beta\right)\left(k_1 k_3 \left(k_{ub4} + k_4\right) k_{b3} - k_{b4} k_2 k_4 \left(k_{ub3} + k_3\right)\right) k_{b2} + k_{b3} k_{b4} k_1 k_3 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub2} + k_2\right)\right) k_{b1} - k_{b2} k_{b3} k_{b4} k_2 k_4 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub1} + k_1\right)$ :

Here,

1. If eq 1. < 0 then multistability is guaranteed for some total concentration of substrate and enzymes.
2. If eq 1. > 0 and eq 2. > 0 then the system is guaranteed to be incapable of multi-stationarity (irrespective of other values of other parameters).
3. If eq 1. > 0, and eq 2. < 0, then the system can exhibit multistability for some total concentration of substrate and enzymes but it is not guaranteed to do so.

**RESULT**
The current structure of these equations (more particularly equation 2) makes it nonintuitive and difficult to understand the effect of key parameters in the behavior of the system. The equations involve feedforward strength and signal, which appear in a nonlinear fashion.

Using semi-analytical approaches, we show two strands of insights here,

1. By identifying structures of the kinetic parameters within these equations and their effect, we reveal key biological insights relevant to the biochemistry of the system that enables or prevents multi-stationarity along with feedforward strength.
2. When some parameters are indeed known, we show that it is possible to obtain regimes of feedforward strength or signal where the system is guaranteed to be incapable of multi-stationarity.

We point out out that these results are intended to serve as representative examples of how semi-analytical tools might be brought to bear. Similar (and more elaborate) insights can be ascertained by focusing on the analytical reduction of the steady state pursued in the relevant proofs, and these results and approaches can also be extended beyond the immediate systems explored in this manuscript.

**Result 1**: We can evaluate biologically relevant analytical expressions from eq 2. that can provide further insight into the network structures enabling (and disabling) multi-stationarity in the network.

*restart* :

We begin our work with the assumption that eq 1. > 0, and thus we currently operate in a regime of enzyme activity where the possibility of multi-stationarity is dependent on the sign of eq 2. (rewritten

below as the expression T).

$$T := \left(\left(S\,u_a + u_\beta\right)\left(k_{IFFW}\,S + d_d + d_\beta\right)\left(k_1\,k_3\left(k_{ub4} + k_4\right)k_{b3} - k_{b4}\,k_2\,k_4\left(k_{ub3} + k_3\right)\right)k_{b2} + k_{b3}\,k_{b4}\,k_1\,k_3\left(S\,u_a + u_d + u_\beta\right)\left(k_{IFFW}\,S + d_\beta\right)\left(k_{ub2} + k_2\right)\right)k_{b1} - k_{b2}\,k_{b3}\,k_{b4}\,k_2\,k_4\left(S\,u_a + u_d + u_\beta\right)\left(k_{IFFW}\,S + d_\beta\right)\left(k_{ub1} + k_1\right)$$

$$T := \left(\left(S\,u_a + u_\beta\right)\left(k_{IFFW}\,S + d_d + d_\beta\right)\left(k_1\,k_3\left(k_{ub4} + k_4\right)k_{b3} - k_{b4}\,k_2\,k_4\left(k_{ub3} + k_3\right)\right)k_{b2} + k_{b3}\,k_{b4}\,k_1\,k_3\left(S\,u_a + u_d + u_\beta\right)\left(k_{IFFW}\,S + d_\beta\right)\left(k_{ub2} + k_2\right)\right)k_{b1} - k_{b2}\,k_{b3}\,k_{b4}\,k_2\,k_4\left(S\,u_a + u_d + u_\beta\right)\left(k_{IFFW}\,S + d_\beta\right)\left(k_{ub1} + k_1\right) \quad \textbf{(1.1)}$$

We can simplify this analytical expression by using the following rearrangement of variables and parameters.

$$k_{b1} := c_1\cdot\left(k_{ub1} + k_1\right) : k_{b2} := c_2\cdot\left(k_{ub2} + k_2\right) : k_{b3} := c_3\cdot\left(k_{ub3} + k_3\right) : k_{b4} := c_4\cdot\left(k_{ub4} + k_4\right) :$$

This results in T taking the following simplified form.

$$T := -\left(k_{ub1} + k_1\right)\left(k_{ub4} + k_4\right)\cdot\left(k_{ub3} + k_3\right)\left(k_{ub2} + k_2\right)\cdot\left(\left(S\,k_{IFFW} + d_d + d_\beta\right)\left(S\,u_a + u_\beta\right)\left(-c_3\,k_1\,k_3 + c_4\,k_2\,k_4\right)c_2\,c_1 + c_3\,c_4\left(-c_1\,k_1\,k_3 + c_2\,k_2\,k_4\right)\left(S\,k_{IFFW} + d_\beta\right)\left(S\,u_a + u_d + u_\beta\right)\right) :$$

If we ignore the first four multiplicative terms for the expression we can further simplify this equation into (this does not affect the sign of the overall expression)

$$T := -\left(\left(S\,k_{IFFW} + d_d + d_\beta\right)\left(S\,u_a + u_\beta\right)\left(-c_3\,k_1\,k_3 + c_4\,k_2\,k_4\right)c_2\,c_1 + c_3\,c_4\left(-c_1\,k_1\,k_3 + c_2\,k_2\,k_4\right)\left(S\,k_{IFFW} + d_\beta\right)\left(S\,u_a + u_d + u_\beta\right)\right) :$$

**By analyzing the structure of this equation we can observe that, if $-k_3\,c_1\,k_1 + k_4\,c_2\,k_2$ and $-k_1\,k_3\,c_3 + c_4\,k_2\,k_4$ are positive, then independent of the feedforward strength, the system is incapable of exhibiting multi-stationarity. It is imperative to notice that these key combinations belong to that of the basic/isolated multisite module.**

A similar type of insight can be arrived at by writing the equation above as a function of feedforward strength as shown below.

$$collect\left(T, k_{IFFW}\right)$$

$$\left(-S\left(S\,u_a + u_\beta\right)\left(-c_3\,k_1\,k_3 + c_4\,k_2\,k_4\right)c_2\,c_1 - c_3\,c_4\left(-c_1\,k_1\,k_3 + c_2\,k_2\,k_4\right)S\left(S\,u_a + u_d + u_\beta\right)\right)k_{IFFW} - \left(d_d + d_\beta\right)\left(S\,u_a + u_\beta\right)\left(-c_3\,k_1\,k_3 + c_4\,k_2\,k_4\right)c_2\,c_1 - c_3\,c_4\left(-c_1\,k_1\,k_3 + c_2\,k_2\,k_4\right)d_\beta\left(S\,u_a + u_d + u_\beta\right) \quad \textbf{(1.2)}$$

**We can now observe that if the coefficient of $k_{IFFW}$ is negative, this guarantees that increasing the feedforward strength indefinitely will result in this entire expression turning negative - thus bringing it out of the region in parameter space where the system has only one single steady state. On the other hand, this expression if positive will similarly guarantee that at increased feedforward strength we are guaranteed to be incapable of multi-stationarity.**

**Result 2:** Identify explicitly regimes of feedforward strength when the system can exhibit multi-stationarity and where it cannot.

Let's assume that the following parameter values are known (except the feedforward strength $k_{IFFW}$, and

the activating signal value)

```
restart : with(plots) : with(ColorTools) :

k1 := 1 : k2 := 1 : k3 := 3 : k4 := 2 :
kb1 := 1 : kb2 := 1 : kb3 := 1 : kb4 := 3 :
kub1 := 1 : kub2 := 1 : kub3 := 1 : kub4 := 1 :
ua := 1 : ud := 4 : uβ := 0.01 : dβ := 0.01 :
da := 2 : dd := 1 :
```

With this information, if we re-evaluate eq 1. and eq 2. from earlier, we obtain.

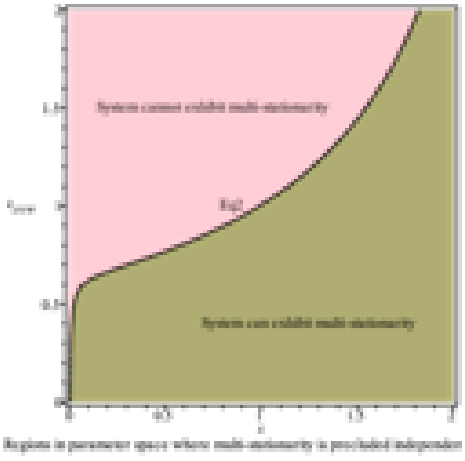
$$Eq1 := k_1 k_3 - k_2 k_4 = 1$$

$$Eq2 := \left(\left(S u_a + u_\beta\right)\left(k_{IFFW} S + d_d + d_\beta\right)\left(k_1 k_3 \left(k_{ub4} + k_4\right) k_{b3} - k_{b4} k_2 k_4 \left(k_{ub3} + k_3\right)\right) k_{b2} + k_{b3} k_{b4} k_1 k_3 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub2} + k_2\right)\right) k_{b1} - k_{b2} k_{b3} k_{b4} k_2 k_4 \left(S u_a + u_d + u_\beta\right)\left(k_{IFFW} S + d_\beta\right)\left(k_{ub1} + k_1\right)$$

$$Eq2 := -15\,(S + 0.01)\left(k_{IFFW} S + 1.01\right) + 6\,(S + 4.01)\left(k_{IFFW} S + 0.01\right) \qquad \textbf{(1.3)}$$

This simplified version allows us to ascertain the ranges of signal and $k_{IFFW}$ when the system is capable of exhibiting multi-stationarity and when it has only one single steady state.

```
C := contourplot(-15 (S + 0.01) (kIFFW S + 1.01) + 6 (S + 4.01) (kIFFW S + 0.01), S = 0 ..2,
    kIFFW = 0 ..2, contours = [0], caption
    = typeset("Regions in parameter space where multi-stationarity is precluded independently of total
    concentrations of substrates or enzymes"), filledregions = true, coloring = [khaki, pink], axes = box)
    :
t1 := textplot([0.75, 1.5,'typeset'("System cannot exhibit multi-stationarity")]) :
t2 := textplot([1.25, 0.4,'typeset'("System can exhibit multi-stationarity")]) :
t3 := textplot([0.85, 1,'typeset'("Eq2")]) :
display({C, t1, t2, t3},'view' = [0 ..2, 0 ..2], size = [0.5, 500])
```



The y-axis plots the feedforward strength $k_{IFFW}$, while the x-axis plots the signal strength S. The curve

lineates the parameter space into one where multi-stationarity is precluded is observed (the region above the curve) and one where multi-stationarity is possible (below).

**This type of analysis allows us to ascertain and identify key parameters and combinations that dictate the possibilities and impossibilities of qualitative behavior exhibited by the whole system (independent of key parameters such as total concentrations of enzymes or substrate).**

# Positive feedback from App: Understanding the role of feedback strength in increasing multi-stationarity

**Recaptulation**
As we have seen earlier in the document, positive feedback from App can result in increased multi-stationarity in the ordered double site modification network. This is the case in the full model where positive feedback results in tristability (when the isolated network without feedback is only capable of atmost bistability). This was similarly shown to be the case when the system has some key complexes removed and where positive feedback from App in such a network was shown to be capable of inducing multi-stationarity (when the isolated network without such feedback was incapable of multi-stationarity).

While our earlier analysis (both computational and analytical) established the presence of increased multi-stationarity, how this relied on the feedback strength was not studied.

**RESULT:**
We establish the following results using the App positive feedback model (for when the feedback is directly effected and the network operates in the enzyme limiting regime).

1. By simplifying (analytically) the expressions from before denoting the steady state of the system, we explain the shortcomings of analytical approaches towards establishing insights on the role of feedback strength in enabling multi-stationarity.
2. We illustrate how semi-analytical approaches might be brought to bear to illustrate role of feedback strength, when some or most other parameters are known.

We point out out that these results are intended to serve as representative examples of how semi-analytical tools might be brought to bear. Similar (and more elaborate) insights can be ascertained by focusing on the analytical reduction of the steady state pursued in the relevant proofs, and these results and approaches can also be extended beyond the immediate systems explored in this manuscript.

**Result 1:** We show the shortcomings of using analytical approaches to discern role of single parameter (feedback strength) in enabling multi-stationarity in the App PFB (direct) system operating in the enzyme limiting regime.

$restart : with(plots) :$

After successive algebraic reduction of the system at steady state, the steady state description of the App PFB (direct) model can be represent as solutions of two coupled polynomials involving App (maximally modified substrate) and the ratio of the enzymes $\epsilon$. These are reproduced here below as equations 2.1 and 2.2 (stored in *Maple* as expressions T and G respectively). Note: The remaining variables can be ascertained as a function of the concentrations of App and '$\epsilon$' and the other parameter values (see file 2.2.1. for more details).

$$T := -c_1 c_2 k_1 k_2 \left(App - A_{Total}\right) \epsilon^2 - c_3 App k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App k_3$$

$$T := -c_1 c_2 k_1 k_2 \left(App - A_{Total}\right) \epsilon^2 - c_3 App k_3 \epsilon k_1 c_1 - k_4 c_4 c_3 App k_3 \quad \textbf{(2.1)}$$

$$G := \left(-\left(c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right) \epsilon + K_{Total} c_4 k_3\right) c_3 k_{PFB} App^2 + \left(k_{PFB} P_{Total} c_2 k_2 \epsilon^2 - c_2 \left(\left(k_2 K_{Total}\right.\right.\right.$$

$$
\begin{aligned}
&- k_3 P_{Total}\right) \left(S u_a + u_\beta\right) c_3 + k_{PFB} K_{Total} k_2\right) \epsilon - K_{Total} c_3 c_4 k_3 \left(S u_a + u_\beta\right)\Big) App \\
&+ c_2 \epsilon k_2 \left(P_{Total} \left(S u_a + u_\beta + u_d\right) \epsilon - K_{Total} \left(S u_a + u_\beta\right)\right)\Big) k_1 + k_4 c_4 c_3 App\, k_3 P_{Total} \left(App\, k_{PFB} \right. \\
&\left. + S u_a + u_\beta\right)
\end{aligned}
$$

$$
\begin{aligned}
G := &\left(-\left(c_2 \left(k_2 K_{Total} - k_3 P_{Total}\right) \epsilon + K_{Total} c_4 k_3\right) c_3 k_{PFB} App^2 + \left(k_{PFB} P_{Total} c_2 k_2 \epsilon^2\right.\right. \\
&- c_2 \left(\left(k_2 K_{Total} - k_3 P_{Total}\right) \left(S u_a + u_\beta\right) c_3 + k_{PFB} K_{Total} k_2\right) \epsilon - K_{Total} c_3 c_4 k_3 \left(S u_a + u_\beta\right)\Big) \\
&App + c_2 \epsilon k_2 \left(P_{Total} \left(S u_a + u_d + u_\beta\right) \epsilon - K_{Total} \left(S u_a + u_\beta\right)\right)\Big) k_1 \\
&+ k_4 c_4 c_3 App\, k_3 P_{Total} \left(App\, k_{PFB} + S u_a + u_\beta\right)
\end{aligned}
\tag{2.2}
$$

These expressions are valid when the system operates in the enzyme limiting regime. As we can notice - these are higher order coupled polynomials in App and '$\epsilon$', thus implying that a simple reduction of them is not possible. We can however solve eq 2.1 for the steady state concentration of App in terms of that of '$\epsilon$' whose re-substitution into eq 2.2 results in it taking the form shown below. This simplified expression is stored as H.

$App := solve(T, App)$ :
$H := simplify(numer(G))$

$$
\begin{aligned}
H := &\left(P_{Total} c_1^2 c_2^2 k_1^2 k_2^2 \left(S u_a + k_{PFB} A_{Total} + u_d + u_\beta\right) \epsilon^5 - k_1^2 \left(\left(-\left(A_{Total} \left(S u_a + k_{PFB} A_{Total}\right.\right.\right.\right. \\
&\left. + u_\beta\right) c_2 + 2 S u_a + k_{PFB} A_{Total} + 2 u_d + 2 u_\beta\Big) P_{Total} k_3 + A_{Total} K_{Total} c_2 k_2 \left(S u_a + k_{PFB} A_{Total}\right. \\
&\left. + u_\beta\right)\Big) c_3 + K_{Total} c_2 k_2 \left(S u_a + k_{PFB} A_{Total} + u_\beta\right)\Big) c_2 k_2 c_1^2 \epsilon^4 - k_1 k_3 c_1 c_3 \left(\left(\left(\right.\right.\right. \\
&-\left(A_{Total} \left(S u_a + u_\beta\right) c_2 + S u_a + u_d + u_\beta\right) P_{Total} k_3 + A_{Total} K_{Total} c_2 k_2 \left(S u_a + u_\beta\right)\Big) c_3 \\
&+ K_{Total} c_2 k_2 \left(A_{Total} \left(S u_a + k_{PFB} A_{Total} + u_\beta\right) c_4 + 2 S u_a + k_{PFB} A_{Total} + 2 u_\beta\right)\Big) k_1 \\
&- A_{Total} P_{Total} c_2 c_4 k_2 k_4 \left(S u_a + k_{PFB} A_{Total} + u_\beta\right)\Big) c_1 - 2 c_2 c_4 \left(S u_a + \frac{k_{PFB} A_{Total}}{2} + u_d \right. \\
&\left. + u_\beta\right) k_2 P_{Total} k_4\Big) \epsilon^3 - k_1 k_3 c_1 \left(k_3 \left(S u_a + u_\beta\right) \left(K_{Total} \left(c_4 A_{Total} + 1\right) k_1 \right.\right. \\
&\left. - A_{Total} P_{Total} c_4 k_4\right) c_3 c_1 + c_4 \left(\left(-\left(A_{Total} \left(S u_a + u_\beta\right) c_2 + 2 S u_a + 2 u_d + 2 u_\beta\right) P_{Total} k_3 \right.\right. \\
&\left. + A_{Total} K_{Total} c_2 k_2 \left(S u_a + u_\beta\right)\Big) c_3 + 2 K_{Total} c_2 \left(S u_a + \frac{k_{PFB} A_{Total}}{2} + u_\beta\right) k_2\Big) k_4\Big) c_3 \epsilon^2 \\
&- k_3^2 c_4 k_4 \left(\left(S u_a + u_\beta\right) \left(K_{Total} \left(c_4 A_{Total} + 2\right) k_1 - A_{Total} P_{Total} c_4 k_4\right) c_1 - P_{Total} c_4 k_4 \left(S u_a \right.\right. \\
&\left. + u_d + u_\beta\right)\Big) c_3^2 \epsilon - K_{Total} c_3^2 c_4^2 k_3^2 k_4^2 \left(S u_a + u_\beta\right)\Big) \epsilon k_1 c_2 k_2
\end{aligned}
\tag{2.3}
$$

Now observe that the resulting polynomial is a sixth-degree polynomial in $\epsilon$, that is capable of at most five positive real roots for the variable. If five such feasible roots exist, it implies the system exhibits increased multi-stationarity (enabling tristability) for the given parameter choice. We have earlier shown that this system does admit to such multi-stationarity and have provided parameter sets where it does so. However if we wish to ascertain how a single parameter such as feedback strength may enable or diminish tristability it is not possible, as the parameter appears in multiple coefficients within this polynomial and is attached to non-trivial combinations of other parameters.

This is thus the shortcoming of the analytical approach where effect of any single parameter is often

mired within the complexity of the network and difficult to independently discern.

**Result 2:** Identify explicitly how feedback strength enables the system to exhibit multi-stationarity.

This is where semi-analytical approaches and an awareness of some parameter values can clear up the landscape and reveal cleanly the role of feedback in enabling multi-stationarity. For example, if we assume that the following parameter values are known (arbitrarily chosen).

$$k_1 := 1.5 : \; k_2 := 25 : k_3 := 70 : k_4 := 12 :$$
$$k_{b1} := 0.25 : k_{b2} := 50 : \; k_{b3} := 100 : k_{b4} := 30 :$$
$$k_{ub1} := 1 : k_{ub2} := 1 : k_{ub3} := 1 : k_{ub4} := 1 :$$
$$u_a := 0.4 : u_d := 2.25 : u_\beta := 0 :$$
$$A_{Total} := 140 : K_{Total} := 6 : P_{Total} := 2 :$$
$$c_1 := \frac{k_{b1}}{k_1 + k_{ub1}} : c_2 := \frac{k_{b2}}{k_2 + k_{ub2}} : c_3 := \frac{k_{b3}}{k_3 + k_{ub3}} : c_4 := \frac{k_{b4}}{k_4 + k_{ub4}} :$$

Then, H reduces to become the following expressions in '$\epsilon$' and feedfoward strength.

$$H := simplify\left(\frac{H}{\epsilon}\right)$$

$$H := \left(3000.362709\, S + 1.050126948 \times 10^6\, k_{PFB} + 16877.04024\right) \epsilon^5 + \left(-115019.5375\, S - 4.241033820 \times 10^7\, k_{PFB} + 69219.63547\right) \epsilon^4 + \left(1.194410632 \times 10^7\, S + 1.284998413 \times 10^7 + 3.869839086 \times 10^9\, k_{PFB}\right) \epsilon^3 + \left(-2.660859352 \times 10^7\, S + 2.620596903 \times 10^7 - 1.192707565 \times 10^9\, k_{PFB}\right) \epsilon^2 + \left(4.178978530 \times 10^9\, S + 2.419012527 \times 10^9\right) \epsilon - 1.290140014 \times 10^9\, S \qquad \textbf{(2.4)}$$

Now by a simple examination we can ascertain that the above simplified form, is a sixth order polynomial with one root guaranteed to be 0 (due to the factorization).

For the system to exhibit increased multi-stationarity (enabling tristability), the system should accommodate five positive solutions for '$\epsilon$'. Thus in order for this to be possible, by Descartes rule of signs, we should require that the signs of the coefficients in the polynomial alternate, that is the coefficient of the fifth, third and first power be positive, while the others be strictly negative.

Note that this isn't a sufficiency condition to guarantee triple steady states in itself, but is a necessary condition.

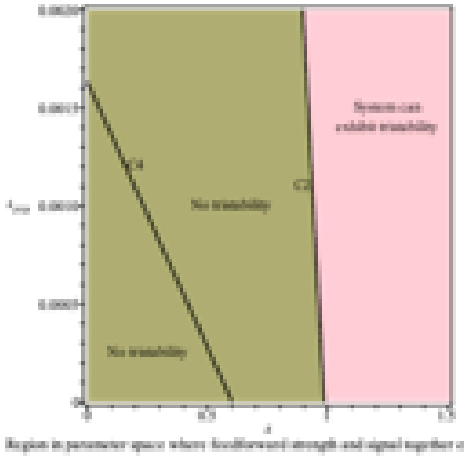
Requiring these for our parameter set defines the region in feedback strength and signal strength that enables increased multi-stationarity (enabling multistabilty).

$$C4 := contourplot\left(coeff(H, \epsilon, 4), S = 0..2, k_{PFB} = 0..0.002, contours = [0], caption = typeset(\text{"Regions in parameter space where feedforward strength enables multi-stationarity"}), filledregions = true, coloring = [khaki, pink], axes = box\right) :$$
$$C2 := contourplot\left(coeff(H, \epsilon, 2), S = 0..2, k_{PFB} = 0..0.002, contours = [0], axes = box, filledregions = true, coloring = [pink, khaki]\right) :$$

$$t1 := textplot([0.25, 0.00025, 'typeset'(\text{"No tristability"})]) :$$
$$t2 := textplot([0.6, 0.001, 'typeset'(\text{"No tristability"})]) :$$

```
t3 := textplot([1.25, 0.0015,'typeset'("System can")]) :
t4 := textplot([1.25, 0.0014,'typeset'("exhibit tristability")]) :
t5 := textplot([0.2, 0.0012,'typeset'("C4")]) :
t6 := textplot([0.9, 0.0011,'typeset'("C2")]) :

display({C4, C2, t1, t2, t3, t4, t5, t6}, 'view' = [0 ..1.5, 0 ..0.002], size = [0.5, 500])
```




C2 and C4 plotted above, represent the coeffecient of the fourth and second exponent in the polynomial. These coeffecients need to be negative for the system to enable increased multi-stationarity. This region where both coeffecients satisfy such a condition is shaded in pink in the above parameter space plot.

The region to the right of the graph represents regimes of feedback strength that enable tristability (please note this isn't a sufficient condition to obtain tristability). Thus, in this manner semi-analytical approaches can shine specific light into how two or more parameters interact with one and other to enable behavior at the systems level.