Optimisation-based Process Synthesis of Emerging Reaction Pathways for Bio-based Polymers and Monomers:

An Effective Mixed Integer Linear Programming Approach

Qingyuan Kong

Centre for Process Systems Engineering
Department of Chemical Engineering
Imperial College London

A Thesis Submitted to Imperial College London for the Degree of Doctor of Philosophy

September 2018

Declaration

I, Qingyuan Kong, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Signature: Qingyuan Kong

Date: ……07/01/2019……

Abstract

This thesis aims to address the process synthesis problems of emerging reaction pathways for bio-based polymers and their monomers and provide the optimal design of the process flowsheet with simultaneous heat integration. To solve these problems, mixed integer linear programming (MILP) models and solution approaches are developed along with a decomposition method to handle numerical complications arising from the utilisation of a large number of discretised variables.

Regarding the process synthesis problem, an optimisation-based framework is first developed to identify the optimal configuration of a process network that consists of both reaction and separation systems. The problem is formulated as a mixed integer linear programming (MILP) model with the objective to maximise the economic potential. A logic-based formulation using binary variables is designed for the simultaneous synthesis of separation sequences. The solution of the optimisation problem includes the best possible economic performance, identification of active reactions, reaction ordering and separation sequences along with the corresponding flowsheet of the entire process. Secondly, simultaneous heat integration is incorporated into the model without introducing nonlinearity by using a novel variable discretisation approach. The modified model aims to provide additional information of the optimal flowsheet such as the utility cost, the minimum cooling and heating duties required, and energy savings due to heat integration.

The process synthesis model consists of a large number of integer variables and becomes difficult to solve. A decomposition method, which utilises the GAMS grid facility, breaks down the original problem into an equivalent set of subproblems by the partitioning of decision variable domains and allows inter-job communications to exchange the best bound of each subproblem. The results indicate that significant improvement in the computational efficiency is achieved when inter-job communication is integrated with the decomposition method.

Acknowledgements

My four-year PhD study has come to an end, and the journey has become an important part of my life. Throughout my time as a PhD student, I have run into numerous obstacles coming from both the work and my personal life, which I could not have successfully overcome without the support and encouragement I have been offered. Therefore, before starting the thesis, I would like to pay my sincere gratitude towards all that have either helped me with my PhD study or inspired me to become a better person.

First of all, I would like to express my deepest thanks and appreciation to my supervisor Prof. Nilay Shah. During the past few years, Prof. Nilay Shah has given me countless advice on academic research and my career path decision. I sincerely appreciate his patient help towards my ignorance and congenial support during my graduate years. Without his visionary guidance and invaluable advice, it would not be possible for me to finish my PhD study so smoothly, and I could not have had a better supervisor.

I would also take the opportunity to thank the committee of the President's PhD Scholarships and the Department of Chemical Engineering of Imperial College, as this project would have been impossible without their financial support.

Besides my supervisor, I would like to thank my colleagues in the group. Particularly, I would give special thanks to Dr Miao Guo, Dr Xiaonan Wang, Dr Kamal Kuriyan and Dr Hongxing Niu for their insightful comments, hard questions and inspirational ideas.

Last but not the least, I am grateful to my friends, from both the College and outside the College: Yingjian Guo, Geng Cheng, Ning Zou, Panpan Ying, Shu Chen, etc. for their genuine support and encouragement throughout my PhD study.

Table of Contents

1. Chapter I: Introduction 1

1.1. Demand for Polymers 1

1.2. Issues with Fossil Fuels 2

1.3. Bio-based Polymers 4

1.4. Process Synthesis 5

1.5. Mathematical Programming Approaches in Process Synthesis Optimisation 7

1.6. Aims and Scope of This Thesis 7

1.7. Thesis Overview 8

2. Chapter II: Literature Review 10

2.1. Recent Development in Bio-based Polymers 10

2.1.1. Bacterial Polymers 10

2.1.2. Feedstock Replacement for Fossil-based Polymers 11

2.1.3. Emerging Bio-based Polymers 12

2.2. Process Synthesis 12

2.2.1. Reaction Path Synthesis 13

2.2.2. Reaction Network Synthesis 14

2.2.3. Synthesis of Separation Sequences 15

2.2.4. Heat Integration 17

2.3. Synthesis of an Entire Process 19

2.3.1. MILP 19

2.3.2. MINLP 20

2.3.3. Process Alternatives Modelling 22

2.4. Decomposition Methods 23

2.4.1. Lagrangean Relaxation and Decomposition 24

2.4.2. Bilevel Decomposition 25

2.4.3. Benders Decomposition 26

2.4.4. Rolling Horizon Decomposition 27

3. Chapter III: Synthesis of Reaction-Separation Network 28

3.1. Problem Statement 28

3.2. Mathematical Formulation 31

3.2.1. Concepts 33

3.2.2. Objective Function 34

3.2.3. Master Stage Mass Balance 35

3.2.4. Value Constraints 36

3.2.5. Separation Stage Mass Balance 37

3.2.6. Separation Mechanism and Cost 40
3.2.7. Other Key Constraints 43
3.3. Case Studies 46
3.3.1. Case I: D-limonene Reaction Network 46
3.3.2. Case II: Itaconic Acid Reaction Network 50
3.3.3. Group Contribution Method 54
3.4. Concluding Remarks 58
4. Chapter IV: Process Synthesis with Simultaneous Heat Integration 59
4.1. Problem Statement 59
4.2. Mathematical Formulation 60
4.2.1. General Approach 63
4.2.2. Objective Function 64
4.2.3. Reactor Heat Surplus 67
4.2.4. Reboilers and Condensers 68
4.2.5. Master Stage Streams (Mstream) 70
4.2.6. Separation Stage Streams (Dstream) 72
4.3. Case study 73
4.3.1. Case Outline 73
4.3.2. Initial Solutions 78
4.3.3. Column Pressure Modification 80
4.4. Sensitivity Analysis 80
4.4.1. Interval Size 81
4.4.2. Reflux Ratio 82
4.4.3. Reaction Temperature 83
4.5. Concluding Remarks 84
5. Chapter V: Enhancement of Computational Efficiency via the Partitioning of the Feasible Region 86
5.1. Background 86
5.1.1. Discretisation 86
5.1.2. Branch and Bound 88
5.1.3. Feasible Region 89
5.2. Problem Statement 90
5.3. Methodology 91
5.3.1. Partition of the Feasible Region 91
5.3.2. Communication between Jobs 93
5.4. Case study 96
5.4.1. Initial Solutions 98
5.4.2. Variable Partition without Communication 103

5.4.3. Variable Partition with Communication 105
5.4.4. Variable Partition into Smaller-Size Subproblems 107
5.5. Concluding Remarks 109
6. Chapter VI: Conclusions and Future Work 111
6.1. Concluding Remarks 111
6.2. Future Work 113
Appendix A Supplementary Information for Chapter III 115
A.1 Bi-linear Equation & Linearisation 115
A.2 Direct Removal Constraints for the Separation Bypass 116
A.3 Separation Coefficient Estimation 117
Appendix B Supplementary Information for Chapter V 122
B.1 Properties of Reactions and Components involved in the Case Study 122
B.2 Equations for the Polymer Separation Unit 123
B.3 Validation Runs for the 16-Subproblem Case 126
B.4 Validation Runs for the 81-Subproblem Case 127
Appendix C GAMS Implementation 128
Appendix D List of Publications 133
References 134

List of Figures

Figure 1.1: Global polymer production. Data from Ryan (2015) 1
Figure 1.2: Global average spot prices of crude and natural gas from 1997 to 2017. Data of natural gas is from Henry Hub Natural Gas spot price (EIA 2018). Data of Crude is from Europe Brent spot price (EIA 2018). 2
Figure 1.3: Projection of the global production capacities of bioplastics from 2017 to 2022. Data from European Bioplastic (2017). 5
Figure 3.1: A demonstrative superstructure for the design of reaction network 28
Figure 3.2: Conceptual diagram of the process superstructure 30
Figure 3.3. Illustration of the solution procedure for the proposed methology 45
Figure 3.4. Proposed reaction network-reaction. Reactions R3 and R4, R5 and R6, R9 and R10 are all competitive reactions. Optimal reaction network is outlined with the dashed envelope. 48
Figure 3.5: Optimal process flowsheet of the d-limonene case, all numbers are in mol/min 49
Figure 3.6: Proposed reaction network superstructure for the itaconic acid case 50
Figure 3.7: Optimal reaction network of the itaconic acid case 51
Figure 3.8: Optimal process flowsheet of the itaconic acid case, all numbers are in mol/min 53
Figure 3.9: Sensitivity analysis of the effect of changing conversions on the EP and Separation Cost 54
Figure 4.1: Implementation of the new constraints and objective function 60
Figure 4.2: Conceptual diagram of the process superstructure 66
Figure 4.3: Proposed reaction network-reaction. Reactions R3 and R4, R5 and R6, R9 and R10 are all competitive reactions. Optimal reaction network is outlined with the dashed circle. 74
Figure 4.4: Optimal process flowsheet of the case study, all flowrate are in the unit of mol/min (temperature intervals of 10 °C) 77
Figure 4.5: The grand composite curve ($\Delta T = 10$ °C) 78
Figure 4.6: The grand composite curve ($\Delta T = 10$ °C) with column depressurisation 80
Figure 4.7: Grand composite curve for the sensitivity analysis on the interval size with original TRj and $Rf = 1.3$ 81
Figure 4.8: Grand composite curve for the sensitivity analysis on the reflux ratio (Rf) with original TRj and $\Delta T = 10\ ℃$ 82
Figure 4.9: Grand composite curve for the sensitivity analysis reaction temperature (TRj) with $\Delta T = 10\ ℃$ and $Rf = 1.3$ 83
Figure 5.1: Variable discretisation using binary variables (a) 87
Figure 5.2: Variable discretisation using binary variables (b) 87
Figure 5.3: An illustration of the branch and bound process for MILP problems with three binary variables 88
Figure 5.4 a): $t1 - t2$ space; b) Feasible region of the whole problem in $t1 - t2$ space 89
Figure 5.5 a): Objective function and optimal solution using graphical method; b) feasible region partition 90
Figure 5.6 i), ii), iii), iv): Graphical illustration of each sub-problem and their respective optimal solutions 92
Figure 5.7: Full solution enumeration of the example maximisation problem 93
Figure 5.8. Solution procedure of the proposed method 95
Figure 5.9: Polymer synthesis reaction network for the case study 97
Figure 5.10: Flow diagram of the conceptual dissolution – precipitation separation process 98
Figure 5.11: The optimal reaction network for the case study 99
Figure 5.12. The discretised grand composite curve ($\Delta T = 2$ °C) 100
Figure 5.13. Optimal process flowsheet of the case study (temperature intervals of 2 °C) 101
Figure 5.14: Objective value of each subproblem without inter-job communication 107
Figure 5.15: CPU time (s) of each subproblem without inter-job communication 108

Figure 5.16: Comparison of CPU time (s) between cases with and without inter-job communication 108

Figure A 1: Bz/To distillation system schematic 117

Figure B 1: CPU times of Sample runs with inter-job communication (81 subproblems) 127

List of Tables

Table 3.1: Discretised split fraction 38
Table 3.2: Boiling points and estimated prices of components in the limonene case study 47
Table 3.3: Conversion/selectivity of each reaction 47
Table 3.4: Model statistics and solution summary of the limonene case 49
Table 3.5: Boiling points and estimated prices for components of the itaconic acid case 51
Table 3.6: Model statistics and solution summary of the itaconic acid case 52
Table 3.7: Flowrates around reactor I and separation cost VS conversion 54
Table 3.8: Market prices of well-known chemicals 55
Table 3.9: Number of groups j in each known chemical and the estimated price of each group j 56
Table 3.10: Number of groups j in each chemical involved in both case studies 57
Table 3.11: Estimated price ($/mol) for chemicals involved in this chapter 58
Table 4.1: Thermodynamic properties of reactions involved in the case study 76
Table 4.2: Thermodynamic properties of components involved in the case study 76
Table 4.3: Inlet/outlet temperatures of each heat exchanger or enthalpy change unit 79
Table 4.4: Model statistics and solution summary of the case study 79
Table 4.5: Summary of problems sizes and solving time for cases with different interval sizes 82
Table 4.6: Summary of key variables for cases with different interval sizes 82
Table 4.7: Summary of key variables for cases with different reflux ratios 83
Table 4.8: Summary of key variables for cases with different reaction temperatures 84
Table 5.1: Modified variable range of the sub-problems 91
Table 5.2: Inlet/outlet temperatures of each 'stream' 100
Table 5.3: Model statics of the original problems 102
Table 5.4: Comparison of the solution statistics for different values of relative stopping tolerance 102
Table 5.5: Theoretical solution status of example subproblems 103
Table 5.6: Feasible region of each subproblem 104
Table 5.7: Model statistics of each subproblem 105
Table 5.8: Model statistics of each subproblem (with communication) 106

Table A 1: Correction factors summary 119
Table A 2: Important parameters used in the calculation 119
Table A 3: Results summary 120
Table A 4: Annualised cost of the separation column 120
Table A 5: Value of σ for the Bz/To case 120
Table A 6: Separation cost coefficient summary of the three cases 121

Table B 1: Reaction conditions for all the reactions involved in the case study 122
Table B 2: Thermodynamic properties of the components involved in the case study 122
Table B 3: Market prices of components involved in the case study 123
Table B 4: CPU times of sample runs with inter-job communication (16 subproblems) 126

1. Chapter I: Introduction

This chapter provides background information on the global polymer market and issues with fossil-based polymers first followed by the discussion on the potential of bio-based polymers. Secondly, the importance and the rules of process design and synthesis strategies are introduced. Finally, mathematical programming approaches used in process synthesis optimisation are explained.

1.1. Demand for Polymers

Polymers are a family of hundreds of important and ubiquitous materials with a wide variety of properties. They are designed and synthesised to meet the needs of various applications in the world's economy and everyone's daily lives. For example, high-resistance (to acids) polymers such as polypropylene (PP) can be used as automotive parts; polymers with thermal-insulation properties such as polystyrene (PS) are used as insulating pipes to reduce the energy loss encountered during heat transfer; and thermally stable polymers such as high-density polyethylene (HDPE) are widely used in the packaging industry. There are two main categories of polymers based on their feedstock orientation. Conventional synthetic polymers are derived from fossil-based feedstocks, such as petroleum and natural gas, and made by scientists and engineers to fulfil specific needs. On the other hand, natural

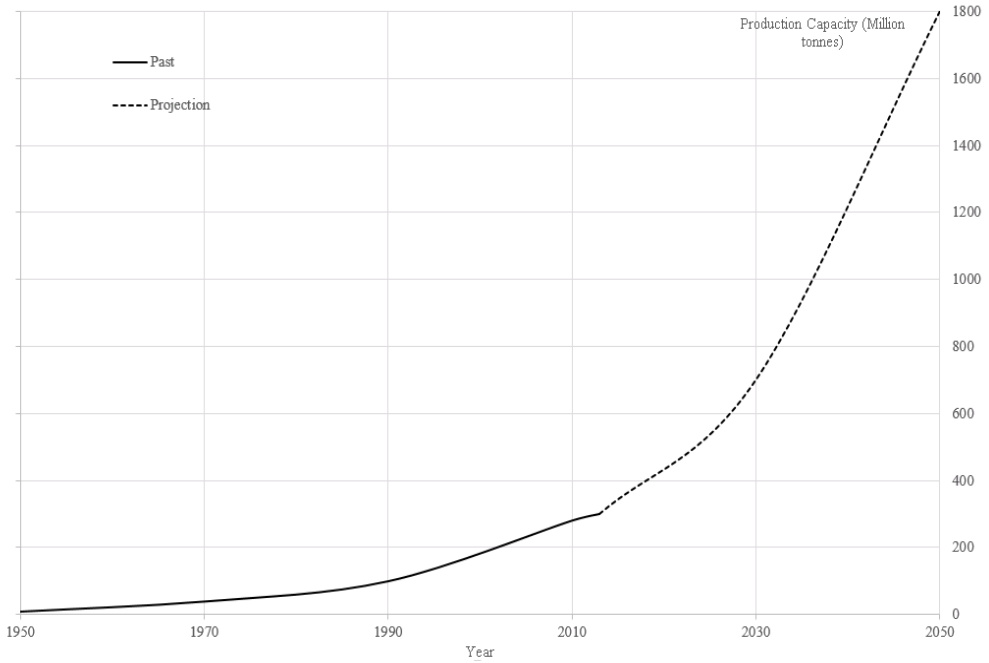


Figure 1.1: Global polymer production. Data from Ryan (2015)

polymers such as rubber, which is harvested as a latex (milky liquid) from several types of trees, and bio-synthetic polymers such as polyhydroxyalkanoates (PHA) are both bio-based polymers. As of 2017, fossil-based polymers still account for almost 99% of the 320 million tonnes of polymers produced, with the remaining 1% being bio-based polymers (European Bioplastic 2017). The polymer production capacity is expected to grow by approximately 4% every year, which is extrapolated using data from

past years (PlasticsEurope 2014). The historical data from the 1950s to 2013 and the projected figures up until 2050 is shown in Figure 1.1.

The exploitation of fossil resources for the production of fuels and other chemicals has been the focal point of discussion for the past few decades. Fundamental problems such as volatility in prices and greenhouse gas emissions pose great uncertainty on the future of fossil fuels. Although only 4 to 6% of all the oil and gas is employed in the production of plastic material, the strong dependence of polymers on fossil resources can cause significant problems for the polymer market. In the next section, the importance of developing bio-based polymers is demonstrated by evaluating the issues surrounding the use of fossil fuel.

1.2. Issues with Fossil Fuels

It is anticipated that the total amount of fossil fuels (coal, natural gas, and oil) on earth that can be extracted using currently available technologies can support current and future needs for at least another hundred years (National Research Council 2005). However, the fast expansion of fossil-based industries is causing many problems to the society and environment. Here, three key reasons why fossil-based polymer production may face great challenges in the future decades are listed.

The first one is the price volatility of fossil fuels caused by factors such as the sudden change in the supply of oil/gas and political instability of some major oil and gas suppliers. Figure 1.2 provides the historic oil spot prices of Europe and gas spot prices of the U.S. since 1997. In the oil graph, the sudden increase in price from 2009 to 2010 is due to the decision made by Organization of the Petroleum Exporting Countries (OPEC) to cut their output to stabilise the oil price (Mouawad 2009). On the other

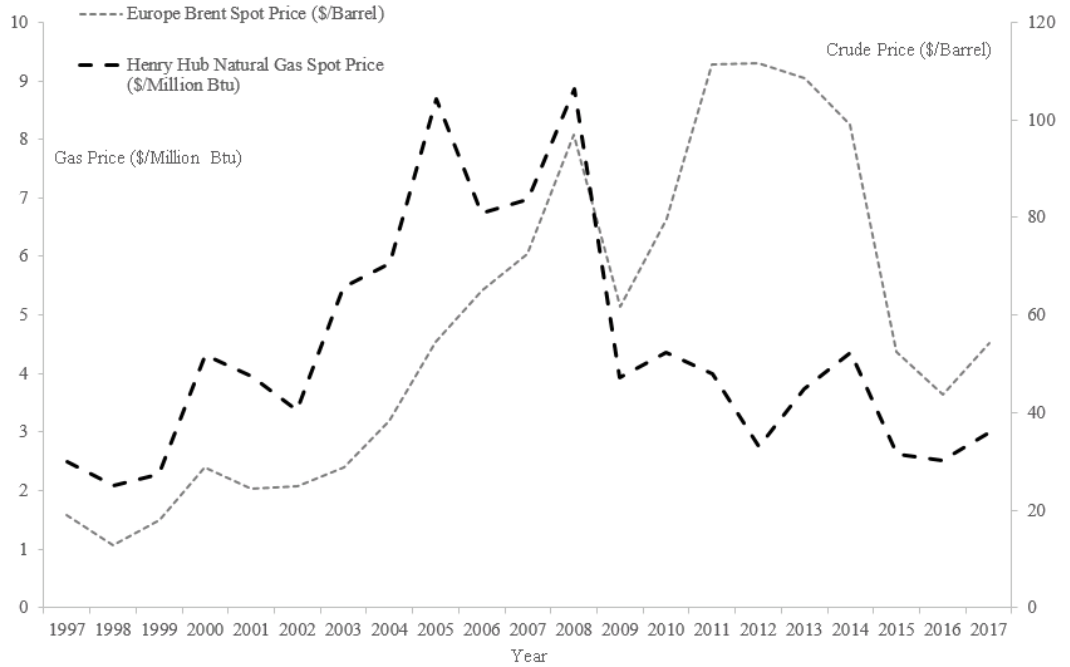


Figure 1.2: Global average spot prices of crude and natural gas from 1997 to 2017. Data of natural gas is from Henry Hub Natural Gas spot price (EIA 2018). Data of Crude is from Europe Brent spot price (EIA 2018).

hand, in the gas price chart, the sudden price drop in-between 2008 and 2009 is caused by the global gas glut induced by weakening gas demand (Oil & Gas Journal 2009). These uncertainties in fossil fuel prices expose polymer manufacturing companies to great financial risks since their performance is highly susceptible to raw material costs.

The second issue caused by the rapid exploitation of fossil fuels is the increase in the greenhouse gas emissions from burning the fuels, the direct effect of which is the deterioration of global climate. It is demonstrated that the amount of global CO_2 emissions is strongly related to the consumption of fossil fuels (IPCC 2014) and CO_2 is a proven greenhouse gas, which is responsible for the significant temperature change over the past few years (Matthews et al. 2009). As a result, policymakers have generally agreed that the average global temperature rise caused by greenhouse gas emissions should not exceed 2 °C above the average global temperature of pre-industrial times (McGlade and Ekins 2015). However, the unabated use of current fossil fuel reserves is incompatible with the 2 °C policy. To satisfy this limit, the future use of fossil resources must be reduced and hence the reduction in the production of fossil-based polymers.

Last but not the least, burning fossil fuels does not only release CO_2 and affect global climate, but it also contributes to air pollution by releasing particulates such as ozone and sulphate, sulphur dioxide gas, and other compounds that can be harmful to human health. For example, it is reported that exposure to sulphur dioxide, nitrogen dioxide, and fine particulate matter from traffic pollution, increases the risk of older adults being hospitalised for pneumonia (Neupane et al. 2010). Also, according to an Environmental Health Perspectives study (Thurston George et al. 2016), long-term exposure to fine particle matter ($PM_{2.5}$) released from fossil fuel combustion can increase the mortality of Ischemic Heart Disease.

In summary, with various limitations on the use of fossil fuels for the production of energy and chemical commodities, it is necessary to consider alternative feedstocks for the production of sustainable polymers from renewable natural resources for decreasing the current dependence on fossil resources. Biomass and its derivatives are well-established raw materials suitable for the production of fuels and commodity chemicals (Huber et al. 2006, Zhou et al. 2008). These materials are naturally generated from available atmospheric CO_2, water and sunlight through biological photosynthesis. As a result, biomass is considered to be the only sustainable alternative carbon source on earth and the perfect equivalent to petroleum for the production of fuels and fine chemicals with potentially net zero carbon emission (Ragauskas et al. 2006, Zhou et al. 2011). In the next section, the advantages of producing polymers using bio-based feedstocks are first discussed followed by its current market status and future outlook.

1.3. Bio-based Polymers

Bio-based polymers are a family of polymers that are made partly or wholly from monomers derived from biomass. The synthesis of bio-based polymers using first generation biomass focuses on deriving highly functional monomers from agricultural food sources such as corn, potatoes, and other carbohydrate feedstocks. One example is the current production of bio-based polyethylene (PE) using bioethanol in Brazil, which relies heavily on sugar crops (Morschbacker 2009). However, there are critical concerns regarding the sustainability of such production schemes, as the consumption of such crops can lead to competition with food production (Cherubini 2010). On the other hand, lignocellulosic feedstocks, which are second-generation biomass, are shown to hold enormous potential for sustainable production of chemicals and fuels (Isikgor and Becer 2015). First of all, lignocellulosic biomass is the most abundant and carbon-neutral renewable source on earth (Somerville et al. 2010); and using it as an alternative to crude oils for the production of biofuels, biomolecules and biomaterials can decrease CO_2 emissions and atmospheric pollutions (Sun et al. 2011, Weiss et al. 2012, Broeren et al. 2016). Secondly, they have a significant edge over the first generation biomass as they are non-edible and do not affect the food supplies (Sun and Cheng 2002). In addition, the cost of lignocellulosic feedstocks is considerably lower than that of food-related feedstocks and crude oil (Huber et al. 2006) thanks to abundantly available agricultural and agro-industrial lignocellulosic wastes, which are accumulated every year in large quantities.

Even with the abovementioned advantages, bio-based polymers still only hold a tiny fraction of the total global polymer market. As mentioned in section 1.1, only about 1% of the commercial polymers are produced using bio-based feedstocks. The holdback is mainly caused by the difficulty in converting biomass to produce value-added chemicals at high selectivities and yields at an economical cost that is comparable to the fossil-based process (Alonso et al. 2013). Nevertheless, the high incentive to switch from fossil-based to bio-based polymers motivates researchers to develop advanced technologies such as the use of ionic liquids and catalytic conversion to refine biomass to fine chemicals and fuels (Holm and Lassi 2011, Zhou et al. 2011).

As shown in Figure 1.3, the global bio-based polymer production capacity is projected to increase from around 2.05 million tonnes in 2017 to approximately 2.44 million tonnes in 2022 (European Bioplastic 2017). To facilitate the industrialisation of innovative bio-based polymers as well as the fast expansion of the global market for bio-based polymers, a systematic way of evaluating and comparing different polymer production routes must be devised. Therefore, process synthesis, which is a well-established concept in the field of process systems engineering for the design and operation of processes, is

employed. In section 1.4, the brief background, future application trends and the structural approach of process synthesis are introduced.

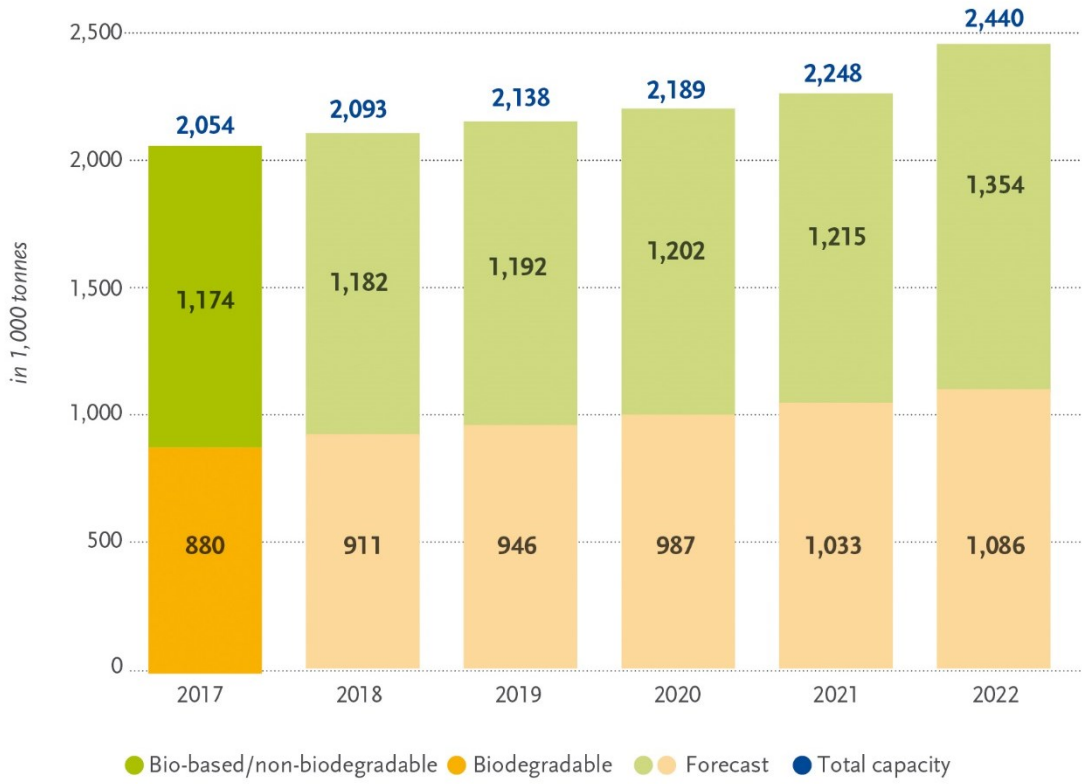


Figure 1.3: Projection of the global production capacities of bioplastics from 2017 to 2022. Data from European Bioplastic (2017).

1.4. Process Synthesis

Process synthesis, as Westerberg (2004) puts it, is the way engineers design and optimise the structure and operating conditions for a new chemical manufacturing process. It involves the determination of the optimal interconnection of all processing units as well as the optimal design of each unit within a process system (Nishida et al. 1981).

The development of process synthesis methodologies has come a long way for the past few decades, and it has experienced significant evolution in dealing with both different research issues and application domains. The predominant areas of focus are, without doubt, the design of chemical reaction networks (Siirola and Rudd 1971), separation systems (Naka et al. 1982), heat exchanger networks (Linnhoff and Flower 1978) and whole process flowsheets (Douglas 1988). However, its general concept is being applied in a wide range of other applications such as environmental impact minimisation (Stefanis et al. 2001) and supply chain design (Beamon 1998). Beyond the conventional definition of process synthesis, emerging process design strategies such as process intensification using advanced technologies (Phan and Harvey 2012), system design at the molecular scale and product design are gaining popularity among both academia and industries (Li and Kraslawski 2004).

Recent developments in biorefinery industries have led to the commercialised production of biopolymers (Maša Šprajcar 2012), biofuels (Essentials 2007) and various biochemicals (Contractors 2016). Companies such as BASF are also switching their focus to more sustainable bio-based products (Hardcastle 2015). A biorefinery system is the equivalent of a refinery system in the bioprocessing industry. It utilises biological, physical and thermochemical conversion technologies to produce biofuels, biochemicals and power more efficiently (Fernando et al. 2006). With the fast development of the biorefinery industry, there are both opportunities and challenges that arise from application transitions of process synthesis from conventional chemical processes to biorefinery processes (Yuan et al. 2013).

The general decision-making and design strategy of process synthesis includes four steps. The first one is defining the clear goal of the process to be synthesised, which could be one or more objectives such as to produce a target product at a specific throughput rate and at the same time the cost of building and operating such a process is minimised. The second is the construction of the design space, i.e. the implementation of necessary constraints on the synthesis process. Once the problem is properly defined, and the boundaries of the problem are set, the third step is to utilise suitable methodologies for the determination of the optimal process flowsheet.

There are two commonly used approaches in the literature that are suitable for the synthesis of the process flowsheet: heuristics-based approaches that use specific process knowledge, experience and physical insights; and superstructure optimisation based approaches where mathematical programming models are formulated and solved. One famous example of heuristics-based approach is the hierarchical decomposition-based process design procedure of Douglas (1988), which decomposes a very large and complex chemical process flowsheeting problem into several smaller, hierarchical sub-problems. Each sub-problem becomes much easier to handle and is solved using related heuristics. Although heuristic-based approaches are suitable for solving large-scale problems, they rarely consider the interactions among various sub-systems, which is proven to be an essential aspect for the optimal design of process systems (Duran and Grossmann 1986). Optimisation-based approaches, on the other hand, are capable of handling implicit trade-offs and interactions among subsystems to make the best decisions for the design of the entire process. This particular type of modelling approach generally consists of three steps (Grossmann 1985). The first step is to postulate a set of superstructures containing a finite number of processing alternatives together with interconnections. Based on the size and the purpose of the model (Chen and Grossmann 2017), each alternative can be evaluated or represented using different approaches including the use of process simulators (ASPEN, etc), surrogate models (Caballero and Grossmann 2008) and short-cut methods (Voll and Marquardt 2012). In the second step, a mathematical programming model (the detailed introduction is provided in the next section) involving sets of discrete and continuous decision variables, constraints and a clearly-defined objective is formulated. The last step is the solution procedure of the formulated optimisation problem to determine the optimal process

flowsheet using one or multiple evaluation metrics such as the economic profit, the energy consumption, and environmental impact.

Having obtained the optimal structure of the entire flowsheet, the fourth and final step of process synthesis is the detailed design of individual operating and interconnecting units that meet all the conditions of the optimal process flowsheet.

1.5. Mathematical Programming Approaches in Process Synthesis Optimisation

As mentioned in the previous section, mathematical programming approaches are one of the key elements in optimisation-based process synthesis. A mathematical programming (or optimisation) problem involves minimising or maximising an objective function subject to a set of equality and inequality constraints, which can be represented as follows:

$$\begin{aligned} &\min f(x,y) \\ s.t. \quad &g(x,y) \leq 0 \\ &h(x,y) = 0 \\ &x \in X, y \in Y \end{aligned}$$

where x and y are the decision variables; $f(x,y)$ is the objective function, $g(x,y)$ and $h(x,y)$ are the inequality and equality constraints, respectively. A typical process synthesis problem involves decisions such as the selection of equipment or the on/off status of a specific valve, which are binary. To handle this special feature, integer variables that can only take the value of 0 or 1 are introduced as "binary variables". This particular type of formulation fall under mixed integer programming (MIP), which can be further differentiated based on the nature of the objective function and the constraints:

- Mixed integer linear programming (MILP): x is a continuous variable and y is a binary (integer) variable; the objective, $f(x,y)$ and the constraints, $g(x,y)$ and $h(x,y)$ are all linear equations,
- Mixed integer nonlinear programming (MINLP): x is continuous variable and y is a binary (integer) variable; the objective, $f(x,y)$ and/or the constraints, $g(x,y)$ and $h(x,y)$ are nonlinear equations.

After having introduced relevant background information and knowledge involved in this thesis, the next section briefly introduces the aims and scope of this thesis.

1.6. Aims and Scope of This Thesis

Over the past few decades, the efficient use of biomass has been recognised as one of the challenging problems among the energy/chemical industries. As a result, much effort has been devoted to the development of process synthesis methods, which utilises well-established process technologies to convert biomass to biofuels and bioenergy (literature are reviewed in section 2.3). On the other hand, the fast development of different synthetic routes towards the production of bio-based polymers is

starting to generate discussions in the polymer industry regarding the use of alternative feedstocks for a wide range of fossil-based polymers. However, the limitation of newly developed bio-based polymers is that their synthetic routes are typically at laboratory-scale and have limited information compared with well-developed reaction pathways, which means existing process synthesis frameworks that require extensive inputs, such as simulation-based and surrogate models, may not be applicable. Therefore, a tailor-made systematic assessment tool is necessary for the screening of all potential reactions and their products to identify the optimal process flowsheet in the very first phase of the process design.

This thesis aims to bridge the gap between the fast development of emerging bio-based polymers and the systematic assessment methods for those synthetic routes by introducing a novel short-curt framework for the process synthesis of bio-based polymers and their monomers. The implications of these routes are investigated using optimisation-based mathematical programming approaches, mainly mixed integer linear programming models. However, due to limited data availability on those reaction pathways, detailed process analysis such as process simulations, investigation on reaction kinetics and azeotropic distillation are not conducted in this thesis. The topics covered in this thesis include: i) a short-curt process synthesis framework considering an economic objective to generate not only the optimal reaction pathways but also the corresponding process configurations (e.g. reactor/distillation sequences) for a given reaction superstructure including the synthesis of bio-based chemicals/polymers; ii) an extended model of (i) for the process synthesis with simultaneous heat integration utilising discretised temperature variables; and iii) a decomposition method featuring the partition of the feasible region to reduce the solution CPU time caused by a large number of discretised variables of the framework developed in (ii). By addressing these issues, this thesis aims to improve current literature on process synthesis for emerging reaction networks as well as make novel contributions to this field.

1.7. Thesis Overview

The rest of this thesis is organised as follows:

A detailed literature review on the recent development of bio-based polymers, process synthesis and mathematical decomposition methods is presented in Chapter II, which highlights relevant background information for the process synthesis problems addressed in the subsequent chapters.

Chapter III, based on which a journal article has been published (Kong and Shah 2016), presents the developed mixed integer linear programming (MILP) framework for the synthesis of a reaction-separation network. This chapter consists of two main subsections including the mathematical formulation and the case studies. The mathematical formulation section provides the comprehensive list of equations developed. The case studies cover the application of the proposed method on two popular bio-based raw materials (limonene and itaconic acid) and synthesis routes to their derivatives.

Chapter IV introduces an extended framework for the process synthesis with simultaneous heat integration based on the approach introduced in Chapter III. One journal article (Kong and Shah 2017) and one conference proceeding (Kong and Shah 2017) have arisen from the content of this chapter. This chapter considers the enthalpy changes of all the important units such as reactors, reboilers and condensers. The continuous variable, temperature, is reformulated as a discretised variable to maintain the model's linearity. The model is applied to the same limonene case study given in Chapter III and provides insight into the effects of simultaneous heat integration on the economic performance of the optimal solution.

Chapter V presents a decomposition method which reformulates the framework of Chapter IV by breaking the original problem into an equivalent set of subproblems. The method utilises the grid-enabled feature in GAMS (Bussieck et al. 2009) and aims to enhance the solver's performance for the specific type of problems with discretised continuous variable through the implementation of inter-job communication. A case study involving the production of several new polymers is used to illustrate the effectiveness of the decomposition method.

Finally, some conclusions are drawn in Chapter VI making links to various chapters, and it also provides recommendations for future work.

2. Chapter II: Literature Review

After having introduced the relevant background of polymer industries focussing on bio-polymer production, process synthesis optimisation and mathematical programming approaches in Chapter I, this chapter presents relevant literature on recent developments in bio-based polymers and process synthesis, the latter of which is the main focus of this thesis. This chapter also includes a section for decomposition methods for which a developed framework is introduced in Chapter V.

2.1. Recent Development in Bio-based Polymers

In this work, bio-based polymers are classified into three different categories based on their current development status as follows:

- Bacterial polymers, which are existing bio-based polymers produced via biosynthetic routes and can serve diverse biological functions and have material properties suitable for numerous industrial and medical applications;
- Well known polymers, traditionally prepared from fossil-derived monomers but which can now be produced using renewable counterparts;
- Emerging polymers, which are not currently available from fossil resources but can be produced via recently developed lab-scale synthetic routes from bio-based raw materials

In this section, the literature related to the recent development of the polymers mentioned above are reviewed.

2.1.1. Bacterial Polymers

The composition, structure, biosynthesis and functional properties of a vast number of bacterial polymers are extensively studied (Rehm 2010). There are four major classes of bacterial polymers, namely polysaccharides, polyesters, polyamides and inorganic polyanhydrides; and several of them are already produced commercially through medium-to-large-scale fermentations. A detailed review of bacterial polymers is not within the scope of this thesis, and therefore, only one illustrative example is reviewed for each type of bacterial polymers in this section.

Naessens et al. (2005) provide a literature summary covering the production, properties and applications of the bio-polysaccharide dextran. The primary synthesis routes of dextran are via fermentation of sucrose with *Leuconostoc mesenteroides* strains and other lactic acid bacteria or fermentation of maltodextrins with certain *Gluconobacter oxydans* strains. Dextran has many applications including being used as a lubricant in eye drops and a therapeutic agent in restoring blood volume for mass casualties. Polyhydroxyalkanoates (PHAs), a family of biodegradable polyesters with diverse structures, can be synthesised by some native bacterial strains, recombinant bacterial strains, and recombinant eukaryotes using various carbon sources. These biologically produced polyesters have many attractive

properties and are produced for use as bulk commodity plastics, fishing lines, and medical uses (Lu et al. 2009). Candela and Fouet (2006) review the biochemical information and synthesis routes of poly-gamma-glutamate (PGA), which is one of the commercialised polyamides. The main applications of PGA, which are metal chelate, absorbent, cryoprotectant, ageing inhibitors, drug carrier or humectant, are based on the properties of PGA, which include ion trapping and high solubility in aqueous solutions. Polyphosphate, which is an inorganic polyanhydride, can be used to form a membrane-anchored complex with low-molecular-mass plyhydroxybutyrate (PHB) to facilitate the uptake of DNA and various ions. It can also be used in various industrial applications ranging from the replacement of asbestos as a flame retardant to their use as flavour enhancers in food. However, commercial production of bacterial polyphosphates is not yet feasible, as polyphosphates can be easily produced by dehydration of rock phosphate. Nevertheless, the ability of bacteria to accumulate polyphosphate while removing phosphate from the environment could make bacterial polyphosphate production a useful by-product of commercial wastewater treatment (Rehm 2010).

2.1.2. Feedstock Replacement for Fossil-based Polymers

This section provides three example polymers, which are traditionally produced from fossil resources and can now also be produced either entirely or partially from renewable feedstocks via commercially viable synthesis routes.

Polyethylene (PE), which is the most common plastic being used in packaging (plastic bags, plastic films, geomembranes, containers, etc.), is produced through the self-polymerisation of ethylene. The production of bio-based ethylene by dehydration of ethanol has a long history, especially in Brazil, where the building of large-scale bio-ethanol plants is motivated by the economic competitiveness of Brazilian sugarcane-based ethanol and by the low carbon footprint of this production route (Morschbacker 2009). Polyethylene Terephthalate (PET) is a thermoplastic polymer of the polyester family and can be used in fibres for clothing, containers for liquids and foods, thermoforming for manufacturing, and in combination with glass fibre for engineering resins. It is primarily produced via direct esterification of terephthalic acid (TPA) with ethylene glycol. Through the use of bio-based ethylene glycol, the partially bio-based PET has been produced by companies such as Coca-Cola for their bottles (Tullo 2017). On the other hand, the production of TPA still relies heavily on petrochemical feedstocks, but extensive research has demonstrated the feasibility of producing TPA using biofeedstocks (Collias et al. 2014). Polybutylene succinate (PBS), which is another thermoplastic polymer resin of the polyester family, is commercially available since 1993 (Xu and Guo 2010). It is one of the biodegradable polymers and can be used as mulching films, compostable bags, nonwoven sheets & textiles, catering products and foams. PBS is currently being produced via the direct esterification of succinic acid (SA) and 1,4-butanediol (BD), which are commonly obtained from fossil resources and are readily available on the market. However, both SA and BD can be obtained from

renewable sources through fermentation and Companies such as DSM, BASF and Purac are producing bio-based SA at an industrial scale (Chen and Patel 2012).

2.1.3. Emerging Bio-based Polymers

The synthetic pathways and characteristics of some of the emerging new bio-based polymers, which have no alternative synthesis routes from fossil resources, are reviewed in this section.

Hauenstein et al. (2015) propose a synthesis route for the production of completely bio-based, high molecular weight poly (limonene carbonate) (PLimC). The polymer is produced via the ring-opening copolymerisation of cis/trans-limonene oxide and CO_2 in the presence of a β-diiminate zinc complex as the catalyst. Limonene oxide (LO), on the other hand, is synthesised from renewable limonene using a suitable catalyst. The molar mass of the polymer could be controlled by changing the ratio of catalyst and monomer, keeping the dispersity of molar mass low. The thermal and optical properties of the product PLimC, is proven to be very attractive (T_g = 130 °C, transmission 94% and haze 0.75%). β-pinene, which is a naturally-occurring alicyclic monomer that is isolated from turpentine, is reported to undergo cationic polymerisation with the presence of $AlCl_3$ etherates ($AlCl_3OPh_2$ or $AlCl_30.8EtOAc$) at room temperature to produce poly(β -pinene)s (Kukhta et al. 2011). The product polymer has a relatively high molecular weight (M_n = 9000–14 000 g mol^{-1}) with excellent thermal properties (T_g = 82–87 °C). Lowe et al. (2009) utilise the natural product dihydrocarvone, found in caraway oil, to form an epoxylactone on a multi-gram scale. The resulting epoxylactone is used as a multifunctional monomer and cross-linker in ring-opening polymerisation reactions. Homo-polymerisation of the monomer produces low-weight polymers (M_n less than 2.5 kg/mol). Copolymerisation of ε-caprolactone with 0.3 to 50% of the monomer gives flexible cross-linked polymers on a multi-gram scale in a one-step synthesis. The ease of synthesising the copolymers makes them especially attractive, and the shape memory properties of the copolymers and their likelihood of degradability makes them good candidates for biomedical devices.

In summary, as technologies develop, more bio-based polymers with appropriate properties will become available, and the need for process synthesis and screening for corresponding processes will grow simultaneously. Therefore, after having introduced the recent development on different types of bio-based polymers, the next sections review the relevant literature work on process synthesis, which is a key step for the design and screening of newly developed reaction pathways. In section 2.21.4, the synthesis of individual units within a process is discussed and followed by the synthesis of an entire process in section 2.3.

2.2. Process Synthesis

Process synthesis contains two strongly related stages: i) a higher-level decision-making process including the construction of process superstructure, definition of objectives for evaluating the system and systematic optimisation to determine the technologies that attain the objectives; ii) a lower-level

decision-making process consists of the decomposition of the optimal superstructure and the detailed design of those subsystems (Nishida et al. 1981). The decisions made within the two levels are closely interlinked, this is: until the higher-level decisions are made, one can never make the lower-level decisions; on the other hand, one cannot make the optimal decisions at the higher levels without knowing the optimal lower-level decisions.

The decomposition of a large process synthesis problem into sub problems allows complex systems to be analysed and solved quickly and provides a good initial flowsheet. There are extensive research focusing on the subcomponents of process synthesis problems (e.g. reaction path synthesis, heat exchanger network synthesis, separation system synthesis, reactor network synthesis, entire flowsheet synthesis, and control system synthesis) using either engineering heuristics or optimisation-based approaches. The subcategories of process synthesis are reviewed in the next few sub-sections, and control system synthesis, which is not within the scope of this thesis, is not reviewed.

2.2.1. Reaction Path Synthesis

Reaction path synthesis, by definition, is the synthesis, analysis and evaluation of different alternative reaction pathways, which are considered as candidate synthesis routes for the production of desired products. The evaluation of each reaction is based on raw material costs, kinetics, potential separation difficulties, catalyst costs, safety and reliability, as well as other factors that would arise in its industrial implementation. There is extensive literature on the synthesis of reaction paths.

Govind and Powers (1981) present a review on the status of research on the reaction path synthesis problem and compare the key attributes of each study such as the molecular structure representation and the reaction representation notation as well as the programming language used. They also propose a systematic framework for the synthesis of the most suitable reaction path by searching a database of 200 generalised reactions. Each reaction is evaluated based on raw material costs, thermodynamic feasibility, yield estimates and energy costs. Rotstein et al. (1982) propose a strategy for the creation and screening of alternative reaction paths, leading to the desired chemical product. The highlight contribution of the method is that it simplifies the overwhelming process of evaluating the alternative reaction paths based on all the processing characteristics. It is done by utilising only some of the invariant algebraic properties of the chemical reactions in the space of ΔG (free energy change) and T (temperature).

Fornari and Stephanopoulos (1994) modify the preliminary screening method for the synthesis of reaction paths, which is based on thermodynamic feasibilities and can potentially generate an infinite amount of alternative paths, to a more robust framework. It includes additional constraints related to role specification, gross added value and demand/supply requirement, which are formulated using the group contribution method. Example cases are solved showing that the addition of the new constraints results in a significant decrease in the number of alternative reaction paths.

Buxton et al. (1997) present a systematic approach for organic reaction path synthesis in which minimum environmental impact considerations are incorporated in order to exploit the earliest opportunities for waste reduction. The size of the reaction path synthesis problem is reduced by decomposing it into a two-step procedure which includes the co-material design and stoichiometry selection. With the raw materials, stoichiometric co-products and product known, the reaction path synthesis problem is no longer open ended and becomes reasonably easy to solve.

2.2.2. Reaction Network Synthesis

The reactor network synthesis problem involves the determination of the type, size, and interconnections of the reactor units, optimal concentration and temperature profiles, and the heat load requirements of the process. Although only part of the overall synthesis problem, a carefully designed reactor network is crucial to the design of the entire process. An extensive review of reactor network synthesis approaches can be found in the literature with key contributions reviewed in this section.

Jackson (1968) propose a modelling framework to simulate the principal features of many different types of reactor systems, e.g. tubular reactor with bypass or recycle. The framework is used to derive an algebraic structure to evaluate the effect of arbitrarily small changes in the mixing pattern or flow configuration on the reactor performance. The aim is to use the evaluation results as a basis for optimising the performance by successive adjustments in the mixing pattern. In this way, it is possible to find the mixing pattern which is best matched to any particular set of chemical reactions.

Achenie and Biegler (1986) develop a novel nonlinear programming (NLP) model for optimally generating the reactor network based on Jackson's work. The new model is designed to handle networks containing both ideal and non-ideal reactors as well as complex mixing patterns. Heat addition/removal is considered to allow the modelling of non-isothermal reactors. In this work, the type of reactor is determined using the dispersion model which postulates that the dispersion coefficient of each reactor is proportional to the length of the reactor. Literature examples are solved to demonstrate the capability of the approach and the solution contains key decisions including the network structure, the reactor type, and the amount of heat addition. Achenie and Biegler (1990) modify their earlier work by replacing the dispersion reactor model with a more generalised and flexible recycle reactor superstructure (RRS), which can be used for more complicated reaction systems. The superstructure consists of a series of basic modules, and within each basic module, there are a reactor with recycle and a heat exchanger. The utilisation of RRS is proven to significantly reduce redundancy in the reaction network and also allows for different modes of heat addition; i.e. through inter-stage cooling/heating and heat exchange of reactor units with the surroundings.

Lakshmanan and Biegler (1996) utilise attainable-region properties derived from geometric targeting techniques to design a concise reactor module for reactor network synthesis. The reactor module is made up of a continuous stirred tank reactor and a plug flow reactor (for two-dimensional targeting) or

a differential sidestream reactor (for higher dimensions). These modules are used to synthesise the optimal reactor network target with respect to a specific objective function, and the problem is formulated as a compact mixed integer nonlinear program. Several example problems are solved to illustrate the feasibility of the proposed algorithm and to demonstrate its applications to simultaneous reactor network and process synthesis.

Schweiger and Floudas (1999) present a general optimisation-based framework for the synthesis of optimal chemical reactor networks. The possible design alternatives are represented via a process superstructure which includes continuous stirred-tank reactors and cross-flow reactors along with mixers and splitters that connect the units. The superstructure is mathematically modelled using differential and algebraic constraints, and the resulting problem is formulated as an optimal control problem. The generality of the overall methodology and effectiveness of the optimisation strategy are illustrated through the demonstration of numerous examples and improved solutions are obtained when compared with previous literature. Ashley and Linke (2004) propose a knowledge-based optimisation framework for the synthesis of chemical reactor networks that combines knowledge acquisition tools with stochastic optimisation methods to robustly and quickly address complex reactor design problems. The method consists of three stages: stage 1-analysis of reaction network pathways; stage 2-development and formulation of rules that are used within the optimal search; and stage 3-stochastic optimisation using meta-heuristics in the form of tabu search. Several illustrative examples are solved using the method, and it is observed that tabu search strategy has a significant edge over other random move selections in terms of improvement rates and solution speed.

2.2.3. Synthesis of Separation Sequences

In addition, synthesis of the distillation sequence is also a necessary procedure when designing a chemical process (Nishida et al. 1981). Generally, large fractionating columns are considered and modelled as a sequence of simple columns with two key components between which the separation takes place (Hengstebeck 1961). Such sequences are conventionally categorised into two main classes: direct and indirect distillation. In a direct sequence, the lightest component is removed as the distillate, with the rest of the mixtures travelling to the next column through the reboiler. On the contrary, the heaviest component is separated first through the bottom of each column in an indirect sequence. Conventional crude distillation usually is an indirect sequence due to its energy saving edge over a direct sequence (Ji and Bagajewicz 2002). However, the conventional sequences may not represent the optimal separation sequence, which can only be obtained by means of optimisation. The optimisation problem for the synthesis of separation sequences is to find the optimal column sequence to separate a multicomponent mixture into several, often pure component, products. This topic is considered by a number of papers in the literature.

Andrecovich and Westerberg (1985) present a method for solving the problem of distillation sequence synthesis with heat integration for sharp split separation. The method is formulated as a mixed integer linear programming (MILP) model which solves within a few seconds of CPU time. A key feature of the proposed model is that it allows the generation of the superstructure for both conventional and multi-effect distillation sequences and therefore permits solutions to include multi-effect sequences if they are economically preferred. Wehe and Westerberg (1987) propose an algorithmic approach for the synthesis of a sequence of simple, sharp split distillation columns with both bypass and mixers to minimise costs. The implementation of bypass and mixers allows the model to handle separation systems that produce multi-component products, i.e. it is not necessary to separate a feed mixture completely into pure components. Most of the equations used are linear except the ones for stream splitters. The nonlinearity is dealt with by introducing a decomposition method that generates a nonlinear submodel for each structurally different flowsheet, which is relaxed to a linear model and readily solved.

Caballero and Grossmann (2004) develop a novel superstructure for the design of sequences of distillation columns that ranges from conventional to fully thermally coupled systems and goes through all the alternative intermediate possibilities. The problem is formulated as a two-stage generalised disjunctive programming problem and is solved as an MINLP. The first stage obtains the best sequence of separation tasks, in which each separation task is considered as a pseudo-column. In the second stage, the sequence of actual columns is extracted among all the thermodynamically equivalent alternatives. A few demonstrative examples are solved using the proposed method, and the results prove the model to be very robust and reliable with small computational time. Ivakpour and Kasiri (2009) propose a framework for the synthesis of distillation column sequences in non-sharp separations with bypass and possible multicomponent products. The algorithm is constructed based on the desired product specifications in the form of maximum and minimum acceptable values of flow rate and composition. The separation matrix is introduced to generate alternative distillation column sequences, which are strictly confined within the user's definition. Results of the studied case demonstrate that the implementation of the proposed method can reduce the separation costs significantly by taking new complex distillation column configurations into consideration.

Regarding azeotropes, Thong et al. (2004) introduce a two-stage procedure for the synthesis of sequences of distillation columns for separating homogeneous multicomponent azeotropic mixtures. The first stage consists of identifying classes of splitting patterns, which may be characterised as feasible or potentially feasible. Sequences are generated based on the combination of these classes of splits to satisfy a given separation requirement. In the second stage, a set of logic-based rules generates a recycle superstructure for every identified sequence, and a range of flow rates are considered for each possible recycle connection. Once the recycle options have been defined, the feasibility of a proposed

flowsheet is assessed rigorously. The feasibility assessment, in turn, forms the basis for systematic flowsheet generation and optimisation.

2.2.4. Heat Integration

Another component of process synthesis, which has a significant impact on the overall performance of a process is the heat recovery network. The task of a heat recovery network is to exchange the available heat of all process streams through heat integration to reduce the consumption of heating and cooling utilities. Since the cost of utilities is usually the dominant item, there is a great incentive to design heat recovery networks that integrate process streams efficiently. The main difficulty in the heat integration problem is its inherent combinatorial nature since usually there is an enormous number of possible networks. There is a large number of methods in the literature for tackling the synthesis problem of heat integration, and several key contributions are listed in this section.

Papoulias and Grossmann (1983) propose a transshipment framework for the synthesis of heat recovery networks. The framework consists of several key functional models. The first model solves a linear programming problem to determine the heat recovery networks with minimum utility cost. The second then utilises integer programming to select the optimal configuration with the minimum number of heat exchanger units out of all the solution networks from the first model. The key feature of the model is its flexibility, which allows the model to be easily incorporated within a mixed-integer programming approach for synthesising chemical processing systems. Several numerical examples are presented which show that the proposed models are computationally very efficient.

Floudas and Ciric (1989) develop two decomposition strategies for overcoming the inherent uncertainties in heat exchanger network synthesis that are associated with the selection of matches and network optimisation tasks. The combination of these tasks can provide the solution to the general problem of selecting the optimal process stream matches and heat exchanger network satisfying the minimum utility cost criterion. Lewin et al. (1998) present a novel framework for automated HEN synthesis based on a genetic algorithm. An efficient, physically meaningful HEN structure representation is first determined by the genetic algorithm; and the heat loads of units are fixed by the Simplex algorithm to meet maximum energy recovery which is used by the genetic algorithm to rate fitness. Through the investigation of several case studies, the proposed method shows to provide the efficient synthesis of HENs.

Yerramsetty and Murty (2008) demonstrate the application of differential evolution, which is one of the evolutionary algorithms that promises simple, fast and robust optimisation, for the synthesis of heat exchanger networks. The solution algorithm employs a simultaneous method of approach to optimise the heat exchanger network. The variables optimised during the synthesis are the string variables associated with the network structure, the heat duties and heat capacity flowrates of the split streams for the heat exchangers in the network. The proposed model is applied to several case studies available

in the literature, and the results of these studies show that the DE-based model is both robust and accurate in identifying the global optima. Peng and Cui (2015) present an efficient simultaneous synthesis method that provides the optimal networks in a two-level procedure. Two probability models are proposed to generate candidate structures by random perturbations in the upper level. The minimum total annualised cost of each candidate structure is solved in the lower level and then sent to the upper level where different structures are evaluated by a simulated annealing mechanism. This two-level method is applied to solve four benchmark cases of heat exchanger network synthesis, and the computational results show that satisfactory heat exchanger network designs can be guaranteed with reasonable cooling schedules.

Heat integration is instrumental to process synthesis. As a result, researchers have invested time and effort into investigating the effect of simultaneous heat integration while performing the optimisation of a process flowsheet. Duran and Grossmann (1986) present a procedure for solving nonlinear optimisation and synthesis problems of chemical processes simultaneously with the minimum utility target for heat recovery networks. A set of constraints that are based on a pinch point location method are formulated for embedding the minimum utility target within the process optimisation. The unique characteristic of this approach is that variable flow rates and temperatures of the streams are handled within a nonlinear optimisation framework. The procedure is straightforward to apply and can be used in the optimisation of fixed flow sheet structures, as well as in the simultaneous structural and parameter optimisation methods for process synthesis. An example problem is presented to illustrate the substantial economic savings that can be obtained with this procedure.

Yee et al. (1990) integrate an NLP superstructure model for heat integration that can simultaneously target for area and energy cost while properly accounting for the differences in heat transfer coefficients between the streams (Yee et al. 1990) and a mixed integer nonlinear programming (MINLP) model which can generate heat exchanger networks where utility cost, exchanger areas and selection of matches are optimised simultaneously (Yee and Grossmann 1990). The integrated model is embedded into a given process flowsheet or superstructure, and the combined model is optimised simultaneously, where flowrates and temperatures of the potential heat integrated streams are treated as variables. Two examples are solved showing that the trade-offs between capital and utility costs can be properly modelled through the use of the proposed simultaneous approach. The simultaneous solution is also compared with the sequential solution, which further confirms that heat integration must be considered at the stage of synthesising or optimising the process.

Quirante et al. (2017) introduce a disjunctive formulation for the simultaneous optimisation and heat integration of systems with variable inlet and outlet temperature in process streams as well as the possibility of selecting and using different utilities. Disjunctive formulation is implemented to model the Pinch Location Method, which is shown to have equal or lower relaxation gap than the best

alternative reformulation, thus reducing computational time. In addition, the model is further modified and extended for the optimisation of heat integration with unclassified streams and area estimation (Quirante et al. 2018).

After having given detailed information about the literature studies on the synthesis of individual processes, the next section presents research work on the synthesis of entire processes combining all components necessary in a process.

2.3. Synthesis of an Entire Process

The synthesis of the entire process by simply connecting the best individual units generated from the component process synthesis yields feasible and usually good solutions. However, it does not consider the interaction between different decision layers such as the effect of heat integration on separation sequencing. Therefore, the proper integration of all process interactions, constraints and objectives is needed to generate the best possible flowsheet. In that sense, optimisation-based process synthesis is regarded as a powerful and sophisticated tool, which can assist researchers to rigorously search through a proposed design space for the optimal process configuration. This section introduces the literature work on the two most iconic versions of process synthesis optimisation models: i) mixed integer linear programming (MILP) in 2.3.1 and ii) mixed integer nonlinear programming (MINLP) in 2.3.2, followed by a review on different approaches employed to simulate process alternatives in 2.3.3.

2.3.1. MILP

Mixed integer linear programming models are well studied and have the advantages of being easy to solve and guaranteed global optimality. There are a number of MILP models in the literature that deal with different aspects of the optimisation of process synthesis, which include synthesis of chemical and biorefinery processes, process synthesis with simultaneous heat integration, short-cut methods and multi-objective optimisation frameworks. Several examples are discussed in this section.

Grossmann (1985) provide a common mathematical framework in the form of MILP for solving a variety of synthesis problems including total processing systems. The work shows how algorithmic methods can play a useful role in process synthesis when they are coupled together with the heuristic and thermodynamic approaches. Results of example problems demonstrate not only that the solutions can be obtained with a reasonable computational expense, but that very often these solutions are rather economical.

Bao et al. (2011) propose an MILP framework for the synthesis and screening of integrated biorefineries. A conceptual superstructure that consists of feedstocks, target products and various processing technologies is introduced to demonstrate the solution process. Three case studies for the production of gasoline from biomass feedstocks are solved with different objective functions, i.e. maximum yield and minimum payback time. The solutions of the case study illustrate the ability of the proposed approach

to generate a wide variety of pathways that achieve the same target but vary significantly in their building blocks and interconnections. Kim et al. (2013) develop an MILP model for the identification and evaluation of biomass-to-fuel production strategies. The formulation is based on a technology superstructure that includes well-known conversion/separation technologies, feedstocks, intermediates, and final products. The model can be adapted to perform strategy evaluations based on different objective criteria such as minimum production cost, minimum energy consumption and maximum profit. An illustrative case study is solved, and the results demonstrate that the model is a suitable tool for the systematic comparison of alternative biomass-to-fuel conversion strategies, the identification of technology gaps and cost drivers in existing strategies as well as synergies between distinct strategies.

Voll and Marquardt (2012) adopt a novel MILP formulation developed by Hechinger et al. (2010), reaction network flux analysis (RNFA), to systemically identify and rank a large number of alternative reaction pathways based on limited data. The generic model enables a fast calculation of reaction network scenarios and subsequently identify the most promising route based on a variety of economic, energetic or ecological objectives. The case study demonstrates how the knowledge gained from a comprehensive mass balance helps to identify the most promising reaction pathways and the most severe bottlenecks in order to guide further research activities in a rational way. Ng (2010) presents an optimisation-based MILP model that can determine the maximum biofuel production and revenue targets of an integrated biorefinery. The approach is based on automated targeting pinch technique, which is commonly used for mass-exchange networks. A biomass-to-biofuel case study is examined using the proposed framework, and the solution provides the optimal network design of an integrated biorefinery for each of those objectives.

Andiappan et al. (2014) introduce a linear multi-objective optimisation approach for the synthesis of a sustainable integrated biorefinery. The metrics of different reaction pathways such as economic performance, environmental impact and energy consumption are compared and evaluated simultaneously to produce the optimal biorefinery configuration. In this work the environmental impact of a given reaction pathway in an integrated biorefinery is quantified by using a new assessment method known as Incremental Environmental Burden Assessment (IEBA). The multi-objective formulation is embedded as bilevel optimisation, and the results for the case study prove the model's capability of handling more than one objective.

2.3.2. MINLP

Linear models are much easier to solve compared with non-linear ones with similar complexities, but they have limitations when it comes to complex constraints and non-linear objective functions. On the other hand, non-linear models require more computational effort but are more robust and adaptable to complex constraints. There are numerous optimisation-based non-linear models dedicated to the design and synthesis of optimal process configurations.

Li et al. (2000) propose a two-level MINLP optimisation model for the synthesis of chemical reaction pathways with an objective to maximise the economic potential of the process. At the first level, given the specified raw materials, the desired product, and the probable by-products, the optimal reaction network is synthesised regardless of thermodynamic and kinetic feasibilities of reactions. At the second level, a series of thermodynamically feasible intermediate reactions are used to replace those unachievable reactions in the first level. Ponce-Ortega et al. (2012) introduce a set of disjunctive programming models for the identification of optimal biorefinery configurations for a given criterion (e.g., the economics, environmental, or safety, etc.). The example problem for the selection of the best economic route to convert lignocellulosic biomass to bioalcohols is solved twice using the same model but with different cost equations. The first case is a mixed integer linear programming model, whereas in the second case, more rigorous nonlinear cost correlations are implemented. The results show that even with the same problem, the two models produce two entirely different pathways.

Baliban et al. (2011) present a thermochemical based process superstructure and its MINLP model for the decision making in converting biomass (switchgrass), coal, and natural gas to liquid (CBGTL) transportation fuels. The model incorporates predetermined discretised temperature values for the simultaneous heat and power integration utilising heat engines to recover electricity from the process waste heat. The nonlinear constraints are introduced to account for stream splitting fractions in mass balances and complex thermodynamic equations, such as vapour liquid equilibrium equation. Four detailed case studies are solved using the proposed method with an objective to minimise plant cost. Good quality local optimum solutions are obtained by fixing several binary variables.

Tang et al. (2013) develop a robust MINLP to determine the optimal biorefinery network structure from an integrated biorefinery pathway map which embeds and interconnects the latest processing technologies. The model is designed so that it can handle the predefined sets of uncertainties, such as biofuel market demands and raw biomass feedstock supply variations. The introduction of nonlinearity allows the model to calculate and display the flowrate ratio of one reactant to another in the same processing technology. A case study involving the production of biofuels from different biomass feedstocks is solved with an objective to maximise the economic performance. A sensitivity analysis is carried out, which emphasises the importance of the availability of raw biomass feedstocks and the market prices of biofuels to the conceptual design of an integrated biorefinery.

Kelloway and Daoutidis (2014) propose an optimisation-based multi-objective framework for the synthesis of the optimal biorefinery configuration, which produces a multitude of chemicals and fuels from multiple feedstocks. The problem is formulated as an MINLP model with energy balances, the splitter mass balance and one of the objective functions being nonlinear. Monte Carlo type sampling is applied to the product prices to generate scenarios for the sensitivity analysis. The results show that

with reasonably high prices of fuels, the combined production of fuels and chemicals is most likely to be profitable and flexible.

2.3.3. Process Alternatives Modelling

Most of the above works do not consider the dynamic or detailed operation of the units within each process alternative and assess each unit using simple mass and energy balances. However, there exists extensive literature focusing on using rigorous implicit simulation models to describe process alternatives. These models typically require a good knowledge of the process, in the form of either well investigated first principle models or rich flowsheet data.

One of the popular approaches is the process synthesis optimisation based on implicit simulation models such as ASPEN PLUS, SuperPro Designer, etc. Diwekar et al. (1992) propose an MINLP process synthesiser that iterates between a) the MILP master problem that determines the process topology and b) the NLP optimiser until specific stopping criteria is reached. The model contains implicit constraints generated from process simulators, which allows the optimisation of process topology as well as the continuous optimisation of the selected flowsheet. A case study is solved to demonstrate that superstructures of synthesis problems with changes in the topology of the flowsheet can be handled effectively with this approach. Caballero et al. (2007) introduce a framework addressing the design and optimisation of chemical processes using process simulators that include detailed models. Constraints of the framework are divided into implicit ones, which include all the equations in the process simulators with an input-output black box structure, and other third-party equations (i.e. sizing and costing correlations) and explicit constraints in the form of equalities or inequalities. The framework is formulated as a generalised disjunctive programming problem and reformulated and solved as an MINLP. The work promotes the used of process simulators for well-known processes in rigorous optimisation problems, with all the advantages of using rigorous models instead of the shortcut models. Brunet et al. (2012) present a systematic method for the optimal development of bioprocesses that relies on the integration of simulation packages and optimisation tools. The design problem is formulated as a mixed-integer dynamic optimisation (MIDO) problem and is solved iteratively between a primal dynamic optimisation problem and a tailored MILP problem. The primal problem focuses on the optimisation of continuous decision variables such as the operating conditions, reaction kinetics and equipment sizes using the solutions from process simulators, whereas the master levels entail the solution of a tailored mixed-integer linear programming (MILP) model that decides on the values of the integer variables. The research shows that that the rich theory available for synthesising standard chemical process flowsheets can be modified and applied to the same extent to their biochemical counterparts. The combined power of simulation and optimisation is also widely used in other applications such as the design of adsorption systems (Brunet et al. 2012), the design of thermodynamic cycles to minimise the LCA impact (Brunet et al. 2012) and the design of distillation sequences (Leboreiro and Acevedo 2004).

Another emerging trend in the field of process systems engineering is the growing use of surrogate models for optimisation-based process synthesis (Chen and Grossmann 2017). In order to account for all complex interactions between engineering decisions when optimising process structure and operating conditions, the resulting optimisation model is typically large-scale and non-convex MINLP. To address this issue, high-quality surrogate models derived from artificial neural networks can be used to replace the complex first-principle unit models. This type of models is advantageous over both the simulation-based and the short-cut approaches utilised in process synthesis optimisation. As discussed by Henao and Maravelias (2011), it reduces the size of the optimisation model compared with simulation-based optimisation but without any compromises in solution quality. Caballero and Grossmann (2008) present a methodology to transform the black box modules in modular flowsheet optimisation to meta-models based on a kriging interpolation. Several examples are solved to demonstrate that the algorithm presented is robust and reliable when some of the most difficult to converge and usually noisy unit operations (distillation columns and reactors) are substituted by a kriging model. Fahmi and Cremaschi (2012) develop a surrogate model utilising artificial neural networks (ANNs), which is used to replace the unit operation, thermodynamics and mixing models in a superstructure optimisation model to synthesise the optimal biodiesel production plant. The optimal solution obtained using the model differs less than one percent from the result obtained by modelling the units in a process simulator. Boukouvala and Ierapetritou (2013) apply surrogate models developed using kriging response surface modelling combined with black-box feasibility analysis on the optimisation of a pharmaceutical tablet manufacturing process. The model shows great potential in handling stochastic-constrained processes which are normally computationally expensive for simulation-based optimisation.

Another important aspect of optimisation-based frameworks is the problem size and its resulting computational time. Real-world problems often involve a large number of variables and equations and can be difficult for the solvers to handle, and the significant increase in solution time is becoming a challenging issue. Various attempts have been made to deal with this issue, and among which, decomposition methods are one of many that show the most promising prospects. Therefore, in the next section, literature related to different optimisation decomposition methods are reviewed.

2.4. Decomposition Methods

Mixed integer optimisation problems often can be difficult to solve due to various reasons, e.g.

- Some may involve one or more ‘hard’ constraints that make the problem intractable,
- Some large-scale mixed integer problems involve decision making for both the design and the operational planning problem, which can be difficult for conventional solvers to solve,

- Stochastic programming optimisation problems, which accounts for the uncertainty in parameters using scenario-based approaches, usually result in large-size problems that are difficult to solve,
- Large-scale optimisation problems over a large number of time periods require expensive computational efforts.

This section introduces literature related to decomposition algorithms that can deal with the four scenarios mentioned above and accelerate the solution processes.

2.4.1. Lagrangean Relaxation and Decomposition

Lagrangean relaxation is a well-studied method that is suitable for dealing with difficult mixed integer optimisation problems that consist of some particular 'hard' constraints. The basic idea behind Lagrangean relaxation is that the problem is dualised by transferring the 'hard' constraints into the objective function in the form of a penalty function. The reformulation results in a model that is easier to solve and the optimal solution of the relaxed problem provides an upper bound to the original problem (maximisation problem).

Lagrangean decomposition is a special case of Lagrangean relaxation, particularly applicable where the original model consists of two or more structured constraint sets. It utilises the unique structure of these constraints, which relates a variable set (e.g. y_n) to one and only one variable (i.e. $y_1 = x, \dots, y_n = x$). Identical copies are first made of the variable that occurs in all the constraints, and an equality constraint is added to set the original variable equal to the copies (i.e. $x = z_1 = \cdots = z_n$). The new equality constraints are treated as the 'hard' constraints and are dualised, and the model decomposes into sub-problems that can be solved independently of each other thus reducing the computational effort required. There are many studies that focus on the application of Lagrangean decomposition in the literature.

Rana and Vickson (1991) propose a mixed integer nonlinear programming (MINLP) model to determine the optimal stopping sequence, cargo allocation for each ship and the number of trips each ship makes in a predetermined period of time for a container-ship routing problem. The complex model is decomposed into sub-problems, one for each ship, using Lagrangean decomposition. The results show the strength of Lagrangean decomposition method in obtaining good feasible solutions quickly. Equi et al. (1997) develop a solution heuristics to reduce the gap between the optimal solution of the original problem and the relaxed solution obtained through Lagrangean decomposition. The method's capability is demonstrated through a case study of a large-scale combined transportation and scheduling problem. They conclude that for each sub-problem, the solution heuristics can help to obtain better optimal solutions than the relaxed ones. van den Heever et al. (2001) present a multi-period MINLP model for the long-term design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. The objective function takes into account complex fiscal rules such as tariff, tax and royalty

thus increasing the computational effort required. The solution time is significantly reduced by the application of Lagrangean decomposition and an iterative solution scheme.

Terrazas-Moreno et al. (2011) compare the differences between the temporal and the spatial Lagrangean decomposition. The comparison is made using a large-scale multi-site, multi-period, and multi-product planning problem, which is formulated as an MILP model. Their results indicate that the temporal decomposition performs better than the spatial decomposition regarding the gap between the final solve and the optimal solution.

2.4.2. Bilevel Decomposition

There are many optimisation problems that focus on the design and planning of large-scale multi-period operations. Often, those problems involve two levels of decision-making processes where the design decisions are made at the first level, which affects the second level operational decisions. Optimal solutions can be obtained by solving the two problems together, but the process can be rather slow. As a result, the bilevel decomposition method, which reformulates the original problem into a two-level problem and then solves them iteratively, is a suitable approach to reduce the computational effort for solving such problems. The higher level design problem excludes the operational binary variables in each time period. For the decisions made at the first level, a lower level planning problem consists of operational optimisation for each period is solved. The gap between solutions obtained from both levels is tighten using an iterative solution procedure combined with integer cut technique.

Iyer and Grossmann (1998) propose a MILP formulation to optimise the synthesis and operational planning of utility systems for multi-period operation with varying demands. The synthesis problem and the operational problem are separated and solved using bilevel decomposition. Three examples are solved and all yield better solutions in competitive computational times compared with the solution from the full space branch and bound method. Laukkanen and Fogelholm (2011) develop a mixed integer non-linear programming (MINLP) model for the synthesis of heat exchanger networks (HENS). The model aims is to simultaneously optimise the heat transfer area, the number of heat exchanger units and the utility costs of the HENS. In their model, the problem is decomposed into four sub-models using not only bilevel decomposition but also the streams grouping and streams aggregation techniques. The bilevel decomposition is implemented to separate the decision-making processes for the optimisation of stream matches and the optimisation of HENS with fixed stream matches. The method's capability is demonstrated using three existing examples from literature, and the results show that the decomposition method can provide comparable solutions to HENS problems with reduced computational effort.

You et al. (2011) compare the solution performance differences between Lagrangean decomposition and bilevel decomposition on solving a large-scale multi-period MILP model. The problem aims to simultaneously obtain the optimal planning of production, capacity and product distribution for a

multisite system including a number of production and market sites. The computational experiments show that bilevel decomposition requires smaller CPU time and provides better feasible solutions compared to the full space solution and Lagrangean decomposition.

2.4.3. Benders Decomposition

Benders decomposition (Benders 1962) is a solution technique in mathematical programming and is often used for large linear programming problems that are uniquely structured, e.g. stochastic programming with uncertainties being treated as scenarios. The solution procedures of Benders decomposition are as follows:

- First, the original problem is divided into a master problem, which only contains the detrimental constraints, and a set of subproblems consists of the variables with uncertainties,
- Second, the master problem is solved and provides the solution to the first set of variables, which is an upper bound for a maximisation problem,
- Third, the subproblem is solved by fixing the values of the variables obtained from the master problem to give a lower bound,
- Finally, the solution of the subproblem is added as the Benders Cut to the master problem (i.e. solutions with worse objective than the subproblem are removed), and the step 2-4 are repeated and a specific target is reached between the lower and upper bound.

Magnanti et al. (1986) study the application and acceleration of Benders decomposition using generic network design problems. Their work shows that the reformulation of several lower bounding inequalities from the literature is suitable to be used as Benders cuts. They also introduce stronger Benders cuts to improve the performance of the decomposition algorithm. The computational results show that the stronger cuts improve considerably upon the usual cuts. Santoso et al. (2005) apply Benders decomposition to a stochastic programming model for solving the design problems of supply chain network and improve the convergence behaviour of the generic Benders decomposition using a number of acceleration techniques such as the implementation of trust region and Knapsack inequalities. The solution reveals that the proposed algorithm can quickly compute high-quality solutions to large-scale stochastic supply chain design problems with a vast number of scenarios. You and Grossmann (2013) develop a multi-cut version of the Benders decomposition method for solving two-stage stochastic linear programming problems. The basic concept is that as many Benders cuts as the number of scenarios are added to the master problem in each iteration. Two examples are solved using the proposed algorithm, and the computational studies show that significant improvement is achieved in CPU time when using the multi-cut version of the Benders decomposition.

2.4.4. Rolling Horizon Decomposition

Rolling horizon (Bassett et al. 1996) is another decomposition technique, which is based on the aggregation of time periods. It can be used to tackle large-scale multi-site and multi-products MILP problems that aim to optimise operational planning over a large number of time periods. The idea of rolling horizon method is that the detailed problem is solved only over a small portion of the time periods at each iteration with the variables of the solved time periods fixed and the rest of time periods aggregated.

Dimitriadis et al. (1997) propose two rolling horizon strategies to accelerate the solution process of the Resource Task Network (RTN) for the scheduling of multipurpose plants (Pantelides 1994). The two strategies differ in the relative positions of the detailed and aggregated time blocks (i.e. backward and forward rolling horizon). The example results show that the solutions obtained using the two algorithms are excellent approximations of the detailed solution but with far less computational time. Kostin et al. (2011) apply the rolling horizon technique to reduce the computational burden, which often occurs in large-scale strategic supply chain models. The capability of the algorithm is demonstrated using a sugarcane industry supply chain design problem in Argentina. The "rolling horizon" algorithm provides near-optimal solutions (i.e., with less than 3% of optimality gap) in a fraction of the time spent by CPLEX. Beraldi et al. (2008) develop a tailor-made compact rolling horizon heuristic for the identical parallel machine lot-sizing and scheduling problem with sequence-dependent set-up costs. The formulation is based on the hypotheses of identical machines. The heuristic solution proves to be of high quality compared with the lower bound provided by a truncated branch-and-bound solution process.

Having introduced the literature work on optimisation-based process synthesis frameworks and relevant decomposition methods, Chapter III introduces a mixed integer linear programming (MILP) model along with a conceptual superstructure for the preliminary process synthesis of optimal chemical reaction networks integrated with distillation sequencing. Two different reaction networks are solved to demonstrate the effectiveness and applicability of the developed approach.

3. Chapter III: Synthesis of Reaction-Separation Network

In this chapter, the first part of the conceptual optimisation framework (Framework I) concerning mainly the mass balance constraints is introduced. It is designed to not only synthesise the optimal reaction pathways for a given reaction network but also to produce the corresponding process configurations such as reactor/distillation sequences. The content of this chapter is organised as follows. Section 3.1 outlines the problem to be solved. Section 3.2 presents the detailed mathematical formulation of the MILP framework maximising the economic (single) objective based on a process superstructure (Figure 3.2). Finally, in section 3.3, two case studies are used to demonstrate the functionality of our framework.

3.1. Problem Statement

There is a wide range of decisions to be made when synthesising new reaction sequences given a set of raw chemicals and reaction schemes. An illustrative reaction network is presented in Figure 3.1. The network under consideration includes the following three elements: 1) raw chemicals which are passed to a set of phase I reactions; 2) the intermediates produced which are fed to phase II reactions; 3) the final products.

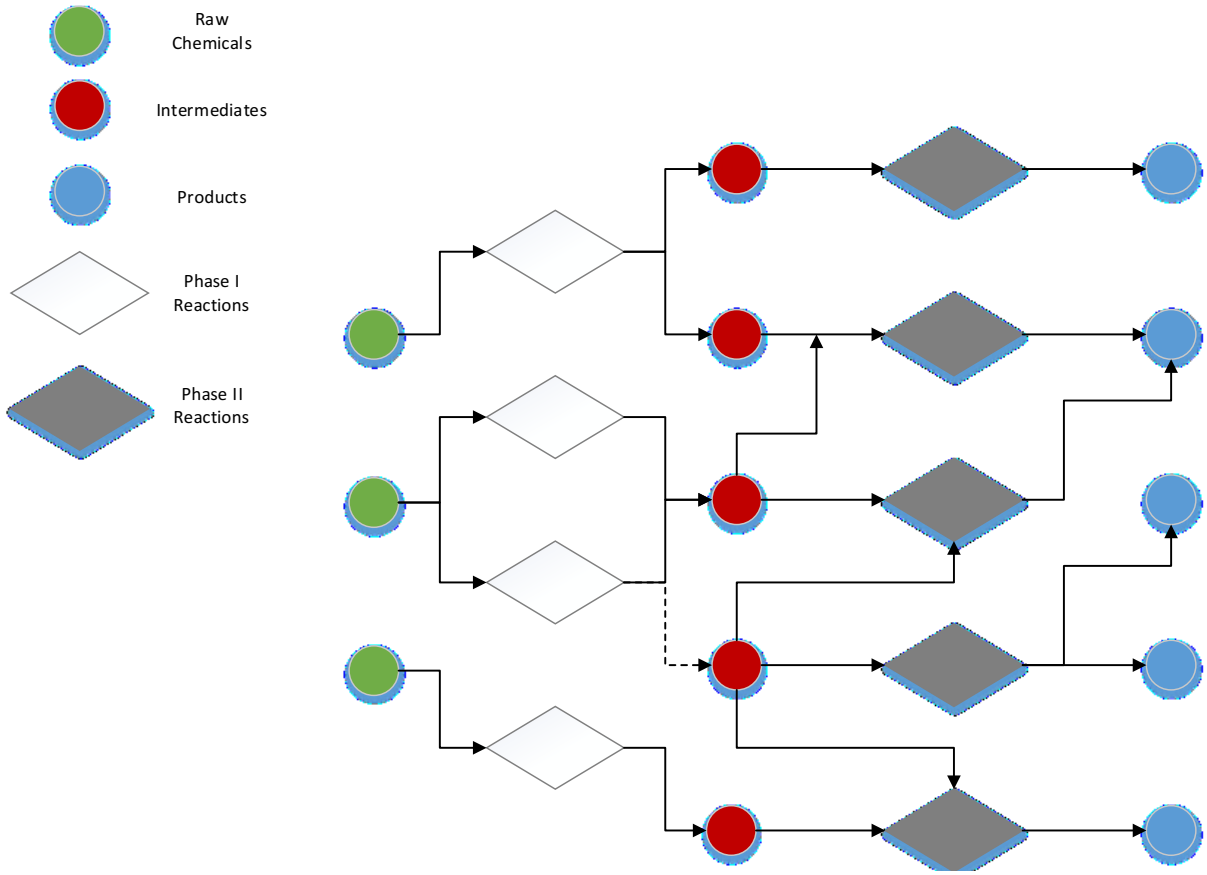


Figure 3.1: A demonstrative superstructure for the design of reaction network

The problem can be described as follows:

Given are:

- a set of chemical feedstocks can be produced from biomass with limited annual supply,
- a set of reactions capable of converting those feedstocks to value-added products,
- characteristics of relevant reactions (conversion, selectivity and conditions),
- the market prices of each reactant and product (can be estimated if unknown),
- the bubble point (@ 1 atm) of each component without complex thermodynamic interactions,
- aggregated operational cost correlations (utility, labour and other combined),
- capital cost standards (mainly distillation columns),

To determine the optimal:

- reaction scheme and the raw material required as well as the end product,
- number of reactors and separation columns required respectively,
- flows of materials (e.g. recycles, intermediates flows and feeds),
- process flowsheet including reaction and separation sequences,

to maximise economic potential of the process.

The model introduced in this work has the following assumptions and justifications:

1. The entire process is in steady-state operation, which is typically the case for well-developed processes.
2. The majority of the cost is assumed to be operational cost, as is the case in most of the modern commodity chemical plant – operational cost is always the primary concern.
3. The operational cost consists of primarily the raw material cost, whereas other operational costs such as labour and maintenance costs are considered indifferent to the optimisation (e.g. manufacturing different chemicals might still need similar numbers of employees).
4. The capital cost differences generated when selecting reactors for each reaction are not considered, as the cost of reactors, if depreciated over time, cannot outweigh the cost of raw materials.
5. Distillation is selected as the separation technique as it is the most widely used and well-studied separation technology.
6. An aggregate coefficient is used to relate all costs involved in distillation including operational and capital costs, to the flowrates and the boiling points of the substances being separated.
7. Sharp splits with no recycle are assumed for the distillation process.
8. The separation of water is not considered in detail in this problem as water can form complex azeotrope with organic compounds, which is not within the scope of this thesis.

The mathematical formulation of the developed model is presented in the following section.

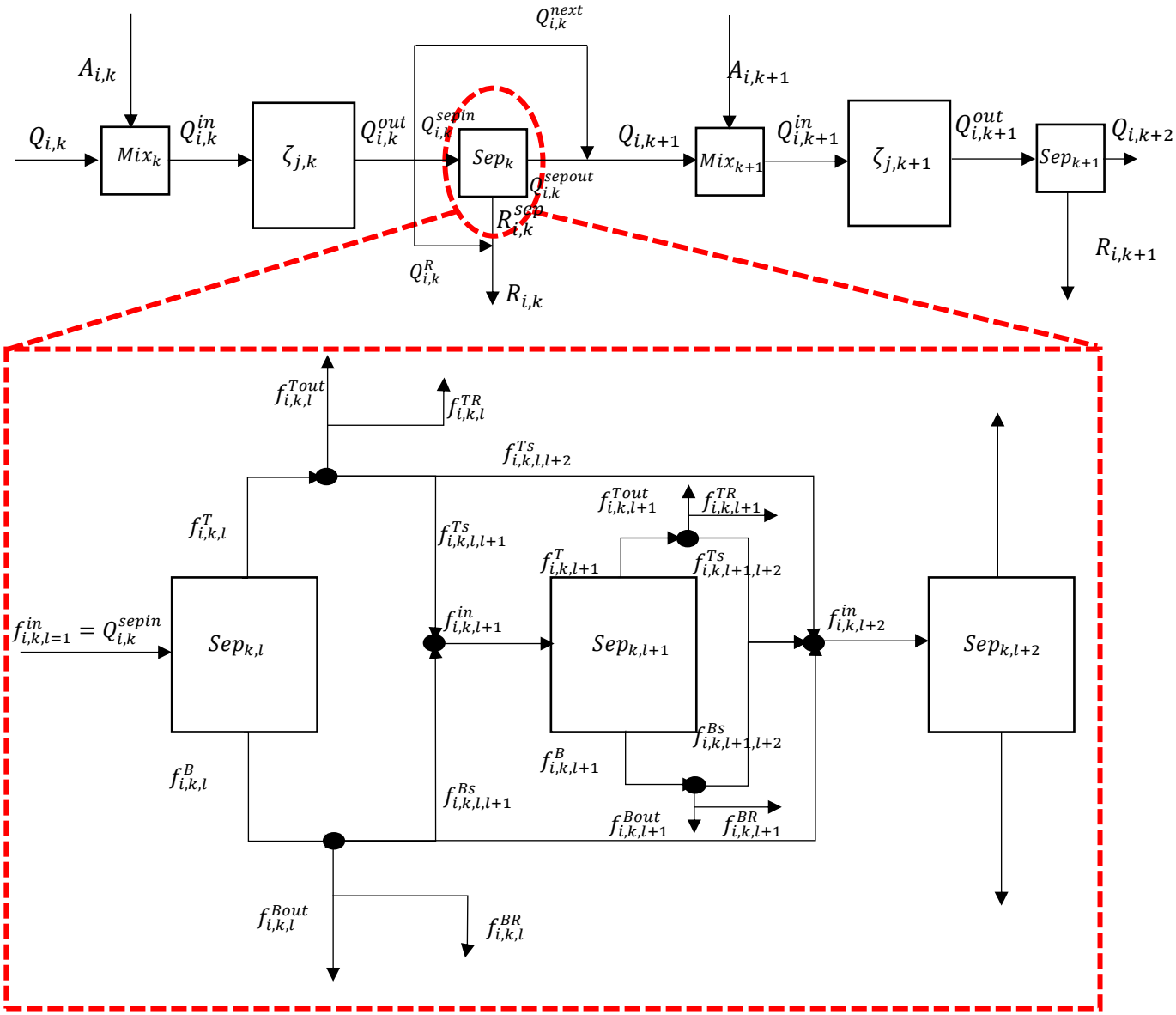


Figure 3.2: Conceptual diagram of the process superstructure

3.2. Mathematical Formulation

The proposed model for the process synthesis optimisation is described in this section. The problem is formulated as a mixed integer linear programming (MILP) model with the following notation:

Nomenclature

Sets	Descriptions
i	Set of components
j	Set of reactions
k	Set of master stages
l	Set of separation stages
p	Set of discretised split fraction points

Binary Variables

b_k^{next}	$= 1$ if all the flow from reaction stage k goes directly to the next stage; 0 otherwise
b_k^R	$= 1$ if all the flow from reaction stage k goes to direct removal; 0 otherwise
b_k^{sepin}	$= 1$ if all the flow from reaction stage k goes to the separation; 0 otherwise
$EX_{i,k,l}^T$	$= 1$ if top removal of separation stage l of master stage k contains only pure component i; 0 otherwise
$EX_{i,k,l}^B$	$= 1$ if bottom removal of separation stage l of master stage k contains only pure component i; 0 otherwise
$EXQ_{i,k}^R$	$= 1$ if stream $Q_{i,k}^R$ contains only pure component i; 0 otherwise
$SA_{k,l}$	$= 1$ if separation stage l in master stage k is active; 0 otherwise
$w_{i,k,l}$	$= 1$ if component i is the light key at separation stage l of master stage k; 0 otherwise
$x_{k,l,l'}^T$	$= 1$ if the stream from stage l to stage l' is active at the top of stage l in master stage k; 0 otherwise
$x_{k,l,l'}^B$	$= 1$ if the stream from stage l to stage l' is active at the bottom of stage l in master stage k; 0 otherwise
$y_{j,k}$	$= 1$ if reaction j occurs in master reaction stage k; 0 otherwise
$Yb_{i,k,l}$	$= 1$ if component i is chosen as the heavy key at separation stage l and in master stage k; 0 otherwise
$\lambda_{k,l,p}^{TR}$	$= 1$ if value of point p is chosen for top removal split fraction at separation stage l of master stage k; 0 otherwise

$\lambda_{k,l,p}^{Tout}$	$= 1$ if value of point p is chosen for top out split fraction at separation stage l of master stage k; 0 otherwise
$\lambda_{k,l,l'p}^{Ts}$	$= 1$ if value of point p is chosen for top split fraction for stream goes from separation stage l to l' at master stage k; 0 otherwise
$\lambda_{k,l,p}^{BR}$	$= 1$ if value of point p is chosen for bottom removal split fraction at separation stage l of master stage k; 0 otherwise
$\lambda_{k,l,p}^{Bout}$	$= 1$ if value of point p is chosen for bottom out split fraction at separation stage l of master stage k; 0 otherwise
$\lambda_{k,l,l'p}^{Bs}$	$= 1$ if value of point p is chosen for bottom split fraction for stream goes from separation stage l to l' at master stage k; 0 otherwise
$NQ_{i,k}^{R}$	$= 1$ if component i exists in flow $Q_{i,k}^{R}$; 0 otherwise
$N_{i,k,l}^{in}$	$= 1$ if component i exists at the inlet of separation stage l of master stage k; 0 otherwise
$N_{i,k,l}^{T}$	$= 1$ if component i exists at the top stream of separation stage l of master stage k; 0 otherwise
$N_{i,k,l}^{B}$	$= 1$ if component i exists at the bottom stream of separation stage l of master stage k; 0 otherwise
$NQ_{i,k}^{in}$	$= 1$ if component i exists at the inlet stream of reaction stage k; 0 otherwise

Continuous Variables

$A_{i,k}$	Flowrate of component i added to the master mixing stage k
$Cost_{k,l}$	Separation cost of each stage l of master stage k
$CostT_k$	Total separation cost of master stage k
$DX_{i',i'',k,l}$	Dummy Variable for linearisation of separation cost equation
EP	Economic potential of the process
$f_{i,k,l}^{in}$	Flowrate of component i entering separation stage l
$f_{i,k,l}^{T}$	Flowrate of component i leaving the top of separation stage l
$f_{i,k,l,l'}^{Ts}$	Flowrate of component i leaving the top of separation stage l and going to stage l', where $l < l'$
$f_{i,k,l}^{Tout}$	Flowrate of component i leaving the top of separation stage l and going straight to the next reaction stage
$f_{i,k,l}^{TR}$	Flowrate of component i leaving the top of separation stage l and going straight to removal as product

$f_{i,k,l}^{B}$	Flowrate of component i leaving the bottom of separation stage l
$f_{i,k,l,l'}^{Bs}$	Flowrate of component i leaving the bottom of separation stage l and going to stage l', where $l < l'$
$f_{i,k,l}^{Bout}$	Flowrate of component i leaving the bottom of separation stage l and going straight to the next reaction stage
$f_{i,k,l}^{BR}$	Flowrate of component i leaving the bottom of separation stage l and going straight to removal as product
$Q_{i,k}^{in}$	Flowrate of component i entering master reaction stage k
$Q_{i,k}^{out}$	Flowrate of component i leaving master reaction stage k
$Q_{i,k}^{sepin}$	Flowrate of component i entering separation at stage k
$Q_{i,k}^{sepout}$	Flowrate of component i leaving separation stages of k and goes directly to the mixing point at stage $k+1$
$Q_{i,k}^{R}$	Flowrate of component i goes to direct removal at stage k
$Q_{i,k}^{next}$	Flowrate of component i goes directly to the mixing point at stage $k+1$
$Q_{i,k}$	Flowrate of component i entering mixing point at stage k
$R_{i,k}$	Flowrate of component i removed from the master separation stage k
$R_{i,k}^{sep}$	Total amount of i out from separation going straight to the removal
$U_{i',k,l}$	Dummy Variable for linearisation of separation cost equation
$VQ_{i,k}^{R}$	Economic value of component i in flow $Q_{i,k}^{R}$
$V_{i,k,l}^{TR}$	Value of top removal of component i at separation stage l of master stage k
$V_{i,k,l}^{BR}$	Value of bottom removal of component i at separation stage l of master stage k
$VRtotal_{i,k}$	Total value of component i removed at master stage k
Z	Objective Function
$\zeta_{j,k}$	Extent of reaction j in master reaction stage k

3.2.1. Concepts

The process superstructure is shown in Figure 3.2. Here, the concepts of master process stages (denoted as k) and their sub separation stages (l) are introduced. Within each master stage k, a mixing point (mix_k) is connected to a reactor ($\zeta_{j,k}$), which is subsequently linked with a separation unit (Sep_k) consisting of a certain amount of distillation columns. The key features of the flow mechanism are given as follows: i) the reactor inlet stream ($Q_{i,k}^{in}$) of each reaction stage comprises the mixture of fresh feed/recycle streams ($A_{i,k}$) and any essential materials produced from previous stages ($Q_{i,k}$); ii) the reactor outlet stream ($Q_{i,k}^{out}$), if contains only one component, bypasses the separation system and is

subject to direct removal ($Q_{i,k}^{R}$); iii) or if it contains whatever is required in the following reactions, it is fed directly to the next reactor ($Q_{i,k}^{next}$); iv) otherwise the mixture passes through the separation system ($Q_{i,k}^{sepin} = f_{i,k,l=1}^{in}$); v) as shown in the dashed envelope, being a conventional distillation column, each separation stage contains one feed ($f_{i,k,l}^{in}$), one distillate ($f_{i,k,l}^{T}$) and one bottom stream ($f_{i,k,l}^{B}$); vi) similar to the reactor, a three-way split mechanism is installed for both the distillate and bottom stream, which includes the removal stream ($f_{i,k,l}^{TR}$ and $f_{i,k,l}^{BR}$), the feed stream for the next separation stage ($f_{i,k,l,l'}^{Ts}$ and $f_{i,k,l,l'}^{Bs}$) and the feed stream to the next reactor ($f_{i,k,l}^{Tout}$ and $f_{i,k,l}^{Bout}$). Note that in each separation unit, sufficient columns are provided initially and the optimisation will determine how many of them are applicable in the optimal scenario. Master stage $k+1$ is coupled with k only if there are inter-stage streams; otherwise, those two stages can be treated as two individual operational units.

The objective function is subject to various constraints, which are categorised into four main types, namely the master stage mass balance, separation stage mass balance, separation mechanism constraints and other constraints. The detailed mathematical formulation of the model is given in the following sections.

3.2.2. Objective Function

In this work, the primary objective is to rank different process networks by their economic potentials (i.e. annual profit). The model accounts for the raw materials cost, which was reported to almost always be the most significant contributor (typically 70 to 80%) to the total cost of production (Julián-Durán et al. 2014). Meanwhile, distillation cost (capital and utility) is also modelled in the objective function to rank different column sequences. The objective function (total economic potential of the process) is written as follows:

$$EP = WH\left(\sum_{i}\sum_{k}\left(VRtotal_{i,k} + VQ_{i,k}^{R} - A_{i,k}(mc_i + AC)\right) - \sum_{i} Q_{i,k=1} mc_i\right) - \sum_{k} CostT_k \quad (3.1)$$

$$Z = \min EP \quad (3.2)$$

where WH is the annual operational hour of the plant, $VRtotal_{i,k}$ is the product value of component i in master stage k, AC is the cost of adding component into mixing stage k (the significance of this parameter is that it avoids overflow of reactant into reactors thus the value of AC is set to be very small but still effective in the model), $Q_{i,k=1}$ is the flowrate of component i entering the first mixing stage and $CostT_k$ is the annualised cost of the separation unit at stage k given as follows:

$$CostT_k = \sum_{l} Cost_{l,k}, \forall k \quad (3.3)$$

3.2.3. Master Stage Mass Balance

The material flow of each stage starts with mixing, i.e. different components i are first mixed at the mixing point before entering the reactor of master stage k, and the reactor inlet ($Q_{i,k}^{in}$) is simply:

$$Q_{i,k}^{in} = Q_{i,k} + A_{i,k}, \forall\, i,k \tag{3.4}$$

The reactor outlet of each component ($Q_{i,k}^{out}$) is equal to the amount of that component in the reactor inlet stream less its consumption or plus its generation in the reactor:

$$Q_{i,k}^{out} = Q_{i,k}^{in} + \sum_{j} v_{i,j}\, \zeta_{j,k}, \forall\, i,k \tag{3.5}$$

where $v_{i,j}$ is the stoichiometric coefficient of component i in reaction j.

The amount of each component i fed to reactor k is set to be higher than the actual amount of that reactant consumed by all reactions j in that stage to avoid model infeasibility, i.e.

$$Q_{i,k}^{in} + \sum_{j} \frac{v_{i,j}\zeta_{j,k}}{x_j} \geq 0, \forall k, i = R_j \tag{3.6}$$

where x_j is the conversion of reaction j.

The extent of reaction, $\zeta_{j,k}$, is a non-negative variable and is bounded by a maximum value ζ^{max}, which is determined by either the user or the maximum capacity of the reactor:

$$0 \leq \zeta_{j,k} \leq y_{j,k}\zeta^{max}, \forall j,k \tag{3.7}$$

In each reactor of master stage k, at most one reaction can occur unless simultaneous reactions exist, which is modelled using the binary variable $y_{j,k}$ as follows:

$$\sum_{j \neq j'} y_{j,k} \leq 1\,, \forall k, j\&j' \; are \; simultaneous \tag{3.8}$$

$$y_{j,k} = y_{j',k}, \forall k, j\&j' \; are \; simultaneous \tag{3.9}$$

One simplification made in this method is that instead of having the same reaction in different reactors, large reactors are available to ensure the extent of each reaction is accounted for in one reactor, i.e. all reactions are set to be active at most once in the entire process flowsheet:

$$\sum_{k} y_{j,k} \leq 1\,, \forall j \tag{3.10}$$

In the case where there are simultaneous reactions (i.e. multiple reactions occur under the same conditions), to ensure the model fully depicts the reaction mechanism, the extents of those reactions are directly linked to the inlet flowrate of the limiting reactants (LR_j), as well as the conversion and selectivity of that reaction:

$$\zeta_{j,k} \geq y_{j,k} \frac{x_j \alpha_j Q_{i,k}^{in}}{-v_{i,j}}, \forall j,k,i = LR_j, v_{i,j} < 0 \tag{3.11}$$

The resulting equation is bi-linear (i.e. involving the product of one continuous variable and one binary variable) and is linearised as follows:

$$\zeta_{j,k} \geq \frac{x_j \alpha_j D1_{i,j,k}}{-v_{i,j}}, \forall j,k,i = LR_j, v_{i,j} < 0 \tag{3.12}$$

$$D1_{i,j,k} \leq Q^{max} y_{j,k}, \forall j,k,i = LR_j \tag{3.13}$$

$$Q_{i,k}^{in} + Q^{max}\left(y_{j,k} - 1\right) - D1_{i,j,k} \leq 0, \forall j,k,i = LR_j \tag{3.14}$$

$$D1_{i,j,k} - Q_{i,k}^{in} \leq 0, \forall j,k,i = LR_j \tag{3.15}$$

where α_j is the selectivity of reaction j. The default value of selectivity for reactions that have no competing reactions is set to be 1 and lies between 0-1 otherwise. $D1_{i,j,k}$ is a dummy variable introduced for the purpose of linearisation.

As mentioned above, a bypass system is installed at the reaction outlet and the corresponding mass balance equation is simply:

$$Q_{i,k}^{out} = Q_{i,k}^{sepin} + Q_{i,k}^{R} + Q_{i,k}^{next}, \forall\, i,k \tag{3.16}$$

In reality, each reactor outlet stream can potentially split into the three bypass streams by any fraction number (as long as they add up to 1) to accommodate various process needs. However, continuous spilt fraction as a variable introduces nonlinearity to the model potentially making the problem non-convex and difficult to solve. This is contrary to the original purpose of developing the model, which is to fast screen and rank reaction pathways with limited data. Therefore, in this model, the split fractions are assumed to be binary and are subject to the following constraint:

$$b_k^{sepin} + b_k^{R} + b_k^{next} \leq 1, \forall k \tag{3.17}$$

which ensures that at most one of the streams can be active and for the other inactive streams, flowrates are 0, i.e.

$$Q_{i,k}^{sepin} \leq b_k^{sepin} Q^{max}, \forall i,k \tag{3.18}$$

$$Q_{i,k}^{R} \leq b_k^{R} Q^{max}, \forall i,k \tag{3.19}$$

$$Q_{i,k}^{next} \leq b_k^{next} Q^{max}, \forall i,k \tag{3.20}$$

where Q^{max} is an arbitrary number defined as the upper limit for all flowrates.

3.2.4. Value Constraints

However, without any value difference between a pure component and a mixture, the bypass system is bound to remove the reactor outlet directly as the product to avoid incurring any separation cost even if

it is indeed a mixture. Therefore, the exclusiveness binary variable ($EXQ_{i,k}^{R}$) is introduced here to distinguish between pure-component and mixture streams. It is formulated as follows:

$$EXQ_{i,k}^{R} \leq NQ_{i,k}^{R}, \forall i,k \tag{3.21}$$

$$EXQ_{i,k}^{R} \leq 1 - NQ_{i',k}^{R} \; \forall i' \neq i,k \tag{3.22}$$

$$NQ_{i,k}^{R} \leq \; Q_{i,k}^{R} \forall i,k \tag{3.23}$$

$$NQ_{i,k}^{R} M \geq Q_{i,k}^{R}, \forall i,k \tag{3.24}$$

The binary indicator ($NQ_{i,k}^{R}$) is used to indicate the existence of each component i in the direct removal stream in each stage k, i.e. $NQ_{i,k}^{R} = 1$ if component i is in stream $Q_{i,k}^{R}$, and $NQ_{i,k}^{R} = 0$ otherwise. Equation (3.23) and (3.24) guarantee that the indicator is related to the flowrate, given M is an appropriately large number. While constraints (3.21) and (3.22) provide the bound for $EXQ_{i,k}^{R}$ if components i' other than i do not exist in the stream, it is the following two constraints together with the objective function ensure the value of the removal outlet ($VQ_{i,k}^{R}$) is equal to its actual value if it consists of only pure component (mc_i is the market price of component i):

$$VQ_{i,k}^{R} \leq mc_i Q_{i,k}^{R}, \forall i,k \tag{3.25}$$

$$VQ_{i,k}^{R} \leq mc_i Q^{max} EXQ_{i,k}^{R}, \forall i,k \tag{3.26}$$

The constraints above allow the system to freely remove any pure components as products, whereas mixtures are further separated to produce more valuable pure components. In the case where the reactor outlet mixture contains the reactant required in the next reactor, it will enter that reactor directly without any separation due to the separation cost term embedded in the objective function.

3.2.5. Separation Stage Mass Balance

As shown in Figure 3.2, each master separation unit comprises a number of sub separation stages denoted as l. The mass balance at the inlet of each separation unit is written as follows:

$$f_{i,k,l=1}^{in} = Q_{i,k}^{sepin}, \forall i,k \tag{3.27}$$

In this work, distillation is selected to be the default separation technology. In a sharp distillation column, the feed ($f_{i,k,l}^{in}$) is separated to form the distillate ($f_{i,k,l}^{T}$) and the bottom ($f_{i,k,l}^{B}$) product (from here and on, both streams are denoted as "product streams"). The mass balance around each distillation column can be written as follows:

$$f_{i,k,l}^{in} = f_{i,k,l}^{T} + f_{i,k,l}^{B}, \forall i,k,l \tag{3.28}$$

Since there are no recycles for the distillation columns, the product streams have three potential end destinations in the same way as the reactor outlets: the inlets to the subsequent separation stages ($f_{i,k,l,l'}^{Ts} / f_{i,k,l,l'}^{Bs}$), the direct removal streams ($f_{i,k,l}^{TR} / f_{i,k,l}^{BR}$) or the input to the next reaction stage

($f_{i,k,l}^{Tout}/f_{i,k,l}^{Bout}$). Therefore, the mass balance equation at the split point for each product stream in separation stage l of master stage k is modelled as follows:

$$f_{i,k,l}^{T} = \sum_{l'>l} (f_{i,k,l,l'}^{Ts}) + f_{i,k,l}^{Tout} + f_{i,k,l}^{TR}, \forall i,k,l \tag{3.29}$$

$$f_{i,k,l}^{B} = \sum_{l'>l} (f_{i,k,l,l'}^{Bs}) + f_{i,k,l}^{Bout} + f_{i,k,l}^{BR}, \forall i,k,l \tag{3.30}$$

Instead of the binary system used for the reactor bypass, a new discretised split fraction system is introduced. This adjustment improves the robustness of the model which can now address more complicated scenarios such as: (i) a small part of a non-pure mixture is removed directly if separating that part is not necessary and it is economically more efficient to remove the mixture as waste directly; (ii) part of the pure component i can be removed as a product if the remaining amount of i entering the next master stage is sufficient for the subsequent reaction.

To formulate the modification, 11 discretised split fractions dsf_p (from 0 to 1) are firstly assigned to 11 points of p, as shown in Table 3.1 (Note here that in this work, 11 points are enough to test the functionality of the method while maintaining the CPU time of solving the model within an acceptable range and the number of points can be increased to improve the accuracy of the optimum results if required).

Table 3.1: Discretised split fraction

p	1	2	3	4	5	6	7	8	9	10	11
dsf_p	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

Second, sets of binary variables λ_p are used at each product stream split indicating which point is selected (i.e. $\lambda_p = 1$ for $p = 3$ if $dsf_p = 0.2$). For each split stream ($f_{i,k,l}^{TR}/f_{i,k,l}^{BR}$ and $f_{i,k,l}^{Tout}/f_{i,k,l}^{Bout}$), the number of active fraction points is limited to 1 to avoid degeneracy in optimisation, i.e.

$$\sum_{p} \lambda_{k,l,p}^{TR} \leq 1, \forall k,l \tag{3.31}$$

$$\sum_{p} \lambda_{k,l,p}^{Tout} \leq 1, \forall k,l \tag{3.32}$$

$$\sum_{p} \lambda_{k,l,p}^{BR} \leq 1, \forall k,l \tag{3.33}$$

$$\sum_{p} \lambda_{k,l,p}^{Bout} \leq 1, \forall k,l \tag{3.34}$$

On the other hand, if the product stream is a mixture and requires further separation, it can be directed to subsequent stage l', with a unique pathway selection from l to l', i.e.

$$\sum_{l'>l} x^{T}_{k,l,l'} \leq 1\ , \forall k, l \tag{3.35}$$

$$\sum_{l'>l} x^{B}_{k,l,l'} \leq 1, \forall k, l \tag{3.36}$$

where $x^{T}_{k,l,l'}/x^{B}_{k,l,l'}$ are binary variables and are equal to 1 if the pathway from separation stage l to l' is selected in master stage k.

The split that allows streams to go from stage l to l' can only occur if the pathway is established, i.e.

$$\sum_{p} \lambda^{Ts}_{k,l,l',p} \leq x^{T}_{k,l,l'}, \forall k, l, l' > l \tag{3.37}$$

$$\sum_{p} \lambda^{Bs}_{k,l,l',p} \leq x^{B}_{k,l,l'}, \forall k, l, l' > l \tag{3.38}$$

Third, the sum of those split fractions at each product stream split must be equal to 1 following the rule of mass conservation, which is given as follows:

$$\sum_{p} (\lambda^{TR}_{k,l,p} + \lambda^{Tout}_{k,l,p} + \sum_{l'>l} \lambda^{Ts}_{k,l,l',p}) dsf_p = 1, \forall k, l \tag{3.39}$$

$$\sum_{p} (\lambda^{BR}_{k,l,p} + \lambda^{Bout}_{k,l,p} + \sum_{l'>l} \lambda^{Bs}_{k,l,l',p}) dsf_p = 1, \forall k, l \tag{3.40}$$

Finally, the overall mass balance constraints for each stream at each split point are:

$$f^{TR}_{i,k,l} = \sum_{p} \lambda^{TR}_{k,l,p}\, dsf_p f^{T}_{i,k,l}, \forall k, i, l \tag{3.41}$$

$$f^{Tout}_{i,k,l} = \sum_{p} \lambda^{Tout}_{k,l,p}\, dsf_p\, f^{T}_{i,k,l}, \forall k, i, l \tag{3.42}$$

$$f^{BR}_{i,k,l} = \sum_{p} \lambda^{BR}_{k,l,p}\, dsf_p f^{B}_{i,k,l}, \forall k, i, l \tag{3.43}$$

$$f^{Bout}_{i,k,l} = \sum_{p} \lambda^{Bout}_{k,l,p}\, dsf_p\, f^{B}_{i,k,l}, \forall k, i, l \tag{3.44}$$

$$f^{Ts}_{i,k,l,l'} \leq \sum_{p} \lambda^{Ts}_{k,l,l',p}\, dsf_p\, f^{max}, \forall k, i, l, l' > l \tag{3.45}$$

$$f^{Bs}_{i,k,l,l'} \leq \sum_{p} \lambda^{Bs}_{k,l,l',p}\, dsf_p\, f^{max}, \forall k, i, l, l' > l \tag{3.46}$$

Equations (3.41)-(3.44) are bi-linear and they can be linearised (details are included in Appendix A. 1), whereas equations (3.45) and (3.46) relax the inter-stage flowrates from l to l' so that the degree of freedom does not contradict with previous mass balance and the fraction summation Equations ((3.29) - (3.30) & (3.39) - (3.40)).

The inlet stream ($f_{i,k,l'}^{in}$) of each separation stage $l'(l' > l)$ can be formulated as the sum of inter-stage flows originating from all stages l, which have l' as the destination column, i.e.

$$f_{i,k,l'}^{in} = \sum_{l<l'} (f_{i,k,l,l'}^{Ts} + f_{i,k,l,l'}^{Bs}), \forall i, k, l' \tag{3.47}$$

It is also subject to the presence of the distillation column, i.e. no feed for columns that are not in operation:

$$\sum_{i} f_{i,k,l}^{in} \leq M \times SA_{k,l}\,, \forall k, l \tag{3.48}$$

where $SA_{k,l}$ is a binary variable and is equal to 1 if the column is in operation, 0 otherwise. The column sequencing is ordered by imposing the following constraint (i.e. it is not possible for column $l+1$ to be active while column l is inactive):

$$SA_{k,l} \geq SA_{k,l+1}, \forall k, l \tag{3.49}$$

The total amount of component i that is directly removed ($R_{i,k}^{sep}$) from the separation unit at stage k:

$$R_{i,k}^{sep} = \sum_{l} (f_{i,k,l}^{TR} + f_{i,k,l}^{BR}), \forall i, k \tag{3.50}$$

and the part that leaves the separation unit for the next reaction stage ($Q_{i,k}^{sepout}$):

$$Q_{i,k}^{sepout} = \sum_{l} (f_{i,k,l}^{Tout} + f_{i,k,l}^{Bout}), \forall i, k \tag{3.51}$$

Here, another point to note is that the removal of a mixture in the separation stage must be avoided unless the values of the respective component are less than the cost of separating the mixture. Therefore, similar constraints to those mentioned in section 3.2.4 are also applied to the separation removal $R_{i,k}^{sep}$(given in Appendix A.2).

The post-separation mass balance equations of total removal at stage k ($R_{i,k}$) and the total flowrate of component i leaving stage k for stage $k+1$ ($Q_{i,k+1}$) can be written as follows:

$$Q_{i,k}^{R} + R_{i,k}^{sep} = R_{i,k}, \forall\, i, k \tag{3.52}$$

$$Q_{i,k}^{next} + Q_{i,k}^{sepout} = Q_{i,k+1}, \forall i, k \tag{3.53}$$

3.2.6. Separation Mechanism and Cost

As mentioned above, distillation is used in this work as the separation technique. This is because that distillation is most widely used and most important technique for separating fluid mixtures in industrial settings, and approximately 95% of all liquid separations involves distillation (Pohorecki et al. 2010). The design and analysis of distillation columns has been the interest of chemical engineers for decades. Equilibrium based methods were developed to obtain important design variables such as the number of

required stages (Fenske 1932, Gilliland 1940, Underwood 1949), feed stage location (McCabe and Thiele 1925) and composition of the feed and outflows (Thiele and Geddes 1933) for the characterisation of multicomponent distillation columns. However, for complex multicomponent systems, these iterative methods can have difficulties in calculating the required results. Julka and Doherty (1990) proposed a simple non-iterative design procedure for multicomponent distillation columns, which was based on the geometry of the composition profiles in the state space, to determine the number of theoretical trays in the column the optimum, feed-plate location, and the exact product distribution (Julka and Doherty 1993). Non-equilibrium stage model, on the other hand, removes the requirement for increasingly complex modifications to the equilibrium stage model and was shown to be more effective and accurate (Krishnamurthy and Taylor 1985). Doherty and Perkins presented a dynamic model for multicomponent simple distillation using residue curve maps (Doherty and Perkins 1978), which was capable of characterising the distillation of ideal and non-ideal solutions as well as azeotropic ternary mixtures (Doherty and Perkins 1978, Doherty and Perkins 1979). However, residue curve maps enable one to gain insight into infinite-reflux column behaviour only (Tapp et al. 2004), whereas column profile maps (CPMs) obtained by using the difference point equation represent the finite-reflux case. This allows creative designs for the synthesis of more complex non-ideal distillation (Holland et al. 2004, Holland et al. 2010, Beneke et al. 2011).

However, these studies require extensive knowledge about the distillation process to be conducted, which is a luxury for the type of reaction pathways this thesis aims to optimise. Therefore in this section, an innovative distillation mechanism, which simplifies the above-mentioned techniques and mimics sharp and ideal separation, is introduced. First, in each column l, it is assumed that components are separated only between the light and heavy keys (LK and HK). In addition, the components that are more volatile than the HK (i.e. boiling point less than that of the HK or $i \leq i'$) end up as the distillate while the rest ($i > i'$) leave the column as the bottom product:

$$f_{i,k,l}^{T} \geq f_{i,k,l}^{in} - M\left(1 - w_{i',k,l}\right), \forall k, l, i \leq i' \tag{3.54}$$

$$f_{i,k,l}^{T} \leq M(1 - w_{i',k,l})\ \forall k, l, i > i' \tag{3.55}$$

Second, binary variable $w_{i,k,l}$ is introduced to identify the LK, which is defined so that it is equal to 1 if component i is selected to be the light key in separation stage l. Given the sharp split assumption, each component can be the LK in any column for at most once (e.g. in an A,B,C,D system with ascending boiling points, if B is the LK for the first column, then B will be the HK for the A-B mixture before being removed from the separation system):

$$\sum_{l} w_{i,k,l} \leq 1, \forall i, k \tag{3.56}$$

For distillation columns that are in operation (i.e. $SA_{k,l} = 1$), exactly one LK must be assigned for that column, and no LK is necessary for non-selected columns, i.e.

$$\sum_i w_{i,k,l} = SA_{k,l}, \forall k, l \tag{3.57}$$

Furthermore, the LK can only be chosen from the components that are present in the inlet stream of that separation stage l and is formulated as follows:

$$w_{i,k,l} \leq N_{i,k,l}^{in}, \forall i, k, l \tag{3.58}$$

where $N_{i,k,l}^{in}$ is the binary variable indicating the presence of component i in the inlet stream of separation stage l and is related to the flowrate via the following constraints:

$$N_{i,k,l}^{in} \leq f_{i,k,l}^{in}, \forall k, i, l \tag{3.59}$$

$$N_{i,k,l}^{in} M \geq f_{i,k,l}^{in}, \forall k, i, l \tag{3.60}$$

Similarly, in each column, at most one existing component can be selected as the heavy key and that component must be present in the inlet stream of that separation stage as well. This is modelled through the heavy key selection binary variable $Yb_{i,k,l}$ as follows:

$$\sum_i Yb_{i,k,l} \leq 1, \forall k, l \tag{3.61}$$

$$Yb_{i,k,l} \leq N_{i,k,l}^{in}, \forall i, k, l \tag{3.62}$$

In this work, a heavy key is the component with the next higher boiling point than that of the light key, i.e. if component i is determined by the system to be the light key in the column l and component $i+1$ is present in that stage, then it must be the heavy key:

$$\left(1 - w_{i,k,l}\right) + \left(1 - N_{i+1,k,l}^{in}\right) \geq 1 - Yb_{i+1,k,l}, \forall i, k, l \tag{3.63}$$

Note here that the components are ordered by increasing boiling points (i.e. $Bp_{i+1} > Bp_i$).

However, if component $i+1$ is not present and possibly even component $i+2$ is not part of the inlet, the following equation is then used to define the correct heavy key in this scenario:

$$\left(1 - w_{i,k,l}\right) + \sum_{i \leq i'} Yb_{i+1,k,l} + \left(1 - N_{i'+1,k,l}^{in}\right) \geq 1 - Yb_{i'+1,k,l}, \forall i, k, l, i' \geq i \tag{3.64}$$

This condition essentially ensures that if component i is the light key; all components ordering from $i+1$ onwards to i' are not the heavy key and component $i'+1$ must be the heavy key of the column if it exists.

As Douglas mentioned in his book, the distillation capital cost is proportional to the column height, which can be found to be inversely proportional to the relative volatility between the HK and LK and the operational (utility) cost is proportional to the vapour flowrate inside the column which is in a way proportional to the column feed flowrate (Douglas 1988). Therefore, the separation cost of each stage l, $Cost_{l,k}$ is assumed to be proportional to the inlet flowrate of that stage and inversely proportional to the boiling point difference between the light and heavy key, i.e.

$$Cost_{k,l} = \sigma \sum_{i'} \frac{\sum_i f_{i,k,l}^{in}}{\sum_{i''} BP_{i''} Yb_{i'',k,l} - BP_{i'}} w_{i',k,l}, \forall k, l, i'' > i' \tag{3.65}$$

where σ is the separation cost coefficient, the estimation of which is based on the overall distillation annual cost for well-known chemical separations such as benzene/toluene and methanol/ethanol systems (detailed calculation of separation cost coefficient is given in Appendix A. 3). The linearization of the above equation is presented as follows:

$$Cost_{k,l} = \sigma \sum_{i'} U_{i',k,l}, \forall k, l \tag{3.66}$$

$$D3_{i',k,l} \leq f^{max} w_{i',k,l}, \forall i', k, l \tag{3.67}$$

$$\sum_i f_{i,k,l}^{in} + f^{max}\left(w_{i',k,l} - 1\right) - D3_{i',k,l} \leq 0, \forall i', k, l \tag{3.68}$$

$$D3_{i',k,l} - \sum_i f_{i,k,l}^{in} \leq 0, \forall i', k, l \tag{3.69}$$

$$D3_{i',k,l} = \left(\sum_{i''} BP_{i''} DX_{i',i'',k,l} - BP_{i'} U_{i',k,l}\right), \forall k, l, i'' > i' \tag{3.70}$$

$$DX_{i',i'',k,l} \leq M\, Yb_{i'',k,l}, \forall k, l, i'' > i' \tag{3.71}$$

$$U_{i',k,l} + M\left(Yb_{i'',k,l} - 1\right) - DX_{i',i'',k,l} \leq 0, \forall k, l, i'' > i' \tag{3.72}$$

$$DX_{i',i'',k,l} - U_{i',k,l} \leq 0, \forall k, l, i'' > i' \tag{3.73}$$

where f^{max} is an arbitrary number used for the maximum flowrate in separation stage and $D3_{i',k,l}$, $DX_{i',i'',k,l}$ and $U_{i',k,l}$ are positive dummy variables.

3.2.7. Other Key Constraints

Apart from the abovementioned reaction/separation relevant constraints, a few other constraints are required to allow the algorithm to be fully functional.

The first one is that the inlet of each reaction stage must not contain component i if it is not required as a reactant to avoid accidental reactor contamination:

$$NQ_{i,k}^{in} \leq \sum_{i=RC_j} y_{j,k}, \forall k, i \in non-inert \tag{3.74}$$

where $NQ_{i,k}^{in}$ is the binary variable indicating the existence of i in the inlet stream of the reactor at master stage k, and it is defined as follows:

$$NQ_{i,k}^{in} \leq Q_{i,k}^{in}, \forall i, k \tag{3.75}$$

$$NQ_{i,k}^{in} M \geq Q_{i,k}^{in}, \forall i, k \tag{3.76}$$

Recycle allows the removed non-feedstocks i being added to any other master stages throughout the process, i.e.

$$\sum_k A_{i,k} \leq \sum_k R_{i,k}, \forall i \in non - feedstock \tag{3.77}$$

The total amounts of feedstocks are limited annually for the process and therefore sets a boundary for the system, i.e.

$$Q_{i,k=1} + \sum_k (A_{i,k} - R_{i,k}) \leq {FS_i}^{max}, \forall i \in feedstock \tag{3.78}$$

where ${FS_i}^{max}$ is the maximum supply allowance of feedstock i over a specific period of time.

Note here, equation sets (3.23)-(3.24), (3.59)-(3.60) and (3.75)-(3.76) are only applicable for flowrates that are greater than 1. Therefore the unit of the flowrate is adjusted to avoid low flowrate values, and the results generated by the model is subject to validation to ensure that no flowrates are within the range of 0-1.

Overall, the objective function (3.2) and those inequality/equality constraints from Equations (3.4) to (3.73) form a mixed integer linear program (MILP) which can be solved globally to determine the optimal reaction network along with its corresponding separation sequencing strategy.

Two case studies are presented in the next section to illustrate the functionality of the proposed approach. Both are solved using the optimisation software GAMS (version 24.2.3). The solution procedure of the model is illustrated in Figure 3.3.

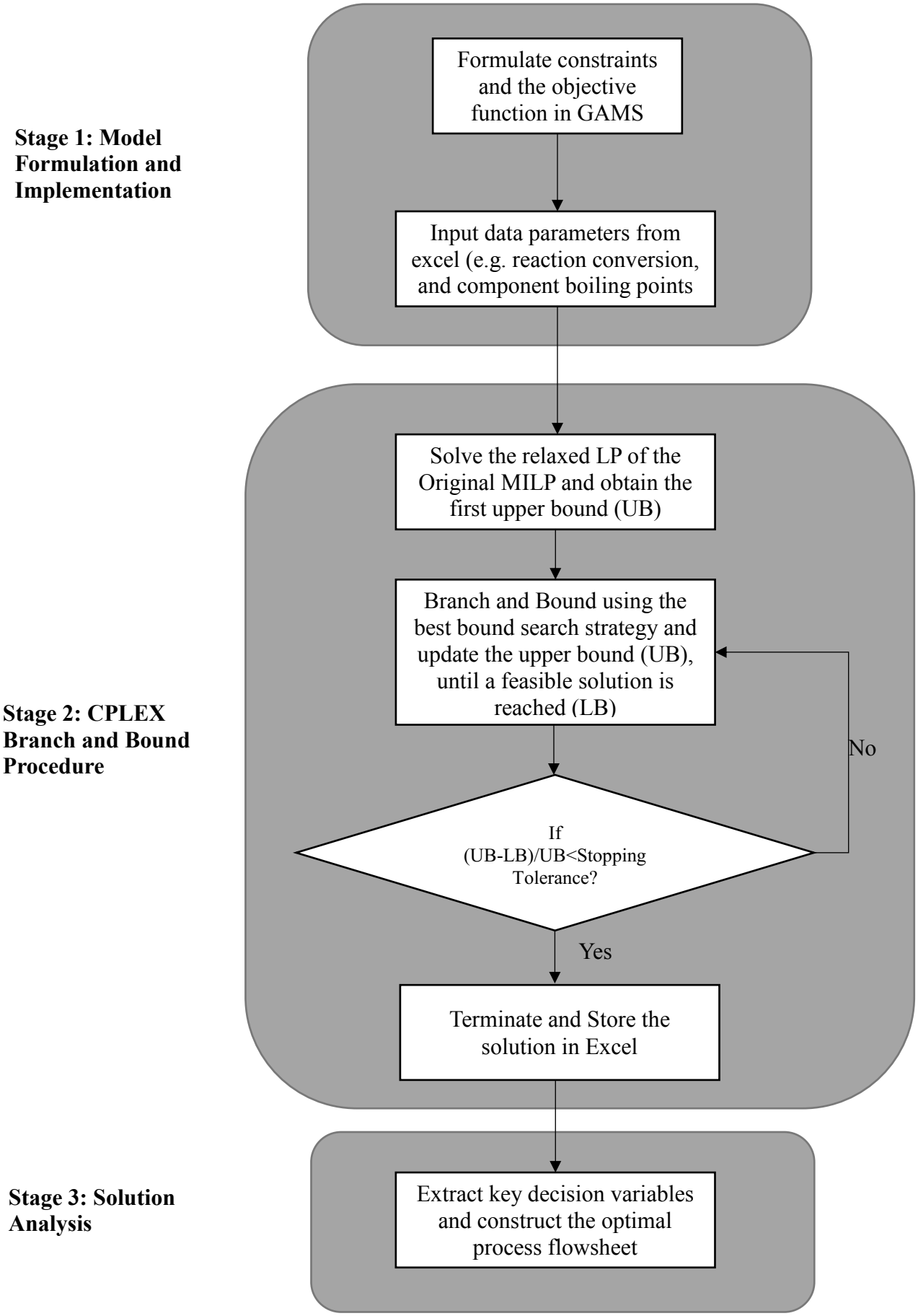


Figure 3.3. Illustration of the solution procedure for the proposed methology

3.3. Case Studies

This work is a part of a 5-year collaborative project funded by the Engineering and Physical Sciences Research Council (EPSRC), the objective of which is to identify bio-polymer production processes that can be commercially competitive and have low environmental impacts. Therefore, bio-derived chemicals and reaction networks starting from those chemicals that are of common interest to the project team are the primary focus. As a result, two bio-based components (which can be synthesised or extracted from biomass feedstocks), d-limonene and itaconic acid, which have high potential to form innovative monomers are selected as the starting materials for the two case studies. The bio-production routes for those two chemicals are not included in the optimisation model since they are fully commercialised processes and will not affect the overall ranking of the process network (i.e. both limonene and itaconic acid are being produced via industrialised technologies, and the optimisation of these biomass conversion technologies is not within the scope of this work).

3.3.1. Case I: D-limonene Reaction Network

The worldwide capacity of d-limonene production is estimated to be greater than 70,000 t in 2013 (Ciriminna et al. 2014) and is assumed to reach approximately 150,000 t in 2019. Therefore, in this case, the hypothetical limonene plant is designed to have an annual processing capacity of 15,000~20,000 t of limonene (roughly 1/10 of the world limonene production) with an annual operational time of 8500 hours.

The proposed reaction network is shown in Figure 3.4. The conversion, selectivity and reaction conditions of each reaction are summarised in Table 3.3. The boiling point of each component is obtained from various databases and can be found in Table 3.2. On the other hand, market prices of some of the uncommon chemicals are not easy to acquire (e.g. limonene dicarbonate and 5,6-epoxycarvone); and therefore a scientific estimation method – group contribution method (Jankowski et al. 2008), which is presented in section 3.3.3, is implemented to estimate approximate bulk prices for those compounds. Note that this method acts only as an approximate pricing tool to obtain rough estimations instead of speculative values. The results may not be precise, but it is sufficient to provide the model with reference data to demonstrate the functionality of the method. In the future, this method can be avoided if proper data is provided.

The optimal reaction network generated by the model is shown in the dashed envelope in Figure 3.4. The simultaneous reactions 3 and 4 starting from d-limonene are preferred over the others. The purified limonene oxide from R3 is a by-product and can be sold directly to the market at its bulk price, whereas carvone produced from R4 is an intermediate reactant for R11. The reaction network ends with R11 producing carvacrol as the main product at a conversion of 80%.

Table 3.2: Boiling points and estimated prices of components in the limonene case study

Component	Molecular Formula	Symbol	Boiling Point (°C)	Estimated Price ($/mol)
D-Limonene	$C_{10}H_{16}$	A	176	0.421
α-Terpineol	$C_{10}H_{18}O$	B	219	0.540
P-cymene	$C_{10}H_{14}$	C	177	0.537
Limonene oxide	$C_{10}H_{16}O$	D	198	0.511
Carvone	$C_{10}H_{14}O$	E	231	0.438
Limonene dioxide	$C_{10}H_{16}O_2$	F	256	0.631
Limonene dicarbonate	$C_{12}H_{16}O_6$	G	350	0.784
Carvacrol	$C_{10}H_{14}O$	H	238	0.703
5,6-epoxycarvone	$C_{10}H_{14}O_2$	I	255	0.651
Dihydrocarvone	$C_{10}H_{16}O$	J	220	0.483
Carvomenthol	$C_{10}H_{20}O$	K	222	0.614

Table 3.3: Conversion/selectivity of each reaction

Reaction No.	Conversion (mol%)	Selectivity (mol%)	Reaction conditions	References
R1	93	-	25 °C and 1 atm	(Adams et al. 2003)
R2	100	-	170 °C and 1 atm	(Martín-Luengo et al. 2008)
R3	100	59.5	75 °C and 1 atm	(Li et al. 2015)
R4	100	40.5	75 °C and 1 atm	
R5	98	74.5	40 °C and 1 atm	(Wiemann Lars et al. 2012)
R6	98	25.5	40 °C and 1 atm	
R7	100	-	140 °C and 30 bar	(Bahr et al. 2012)
R8	32	-	50 °C and 1 atm	(Uguina et al. 2009)
R9	90	62.6	100 °C	(Demidova et al. 2015)
R10	90	37.4	100 °C	
R11	80	-	-	(Kjonaas and Mattingly 2005)

The corresponding optimal process flowsheet is shown in Figure 3.5. In the first master stage, fresh d-limonene (A) is fed to the first reactor where reactions 3 and 4 take places. The product stream of reactor I, which is a mixture of limonene oxide (D) and carvone (E), enters the sub separation unit. The separation system has one column (column I-I) separating between limonene oxide and carvone. Limonene oxide is the LK and distillate due to lower boiling point and leaves from the top of the distillation column. On the other hand, carvone as the HK, is removed through the reboiler and is pumped into the mixer in the next master stage directly.

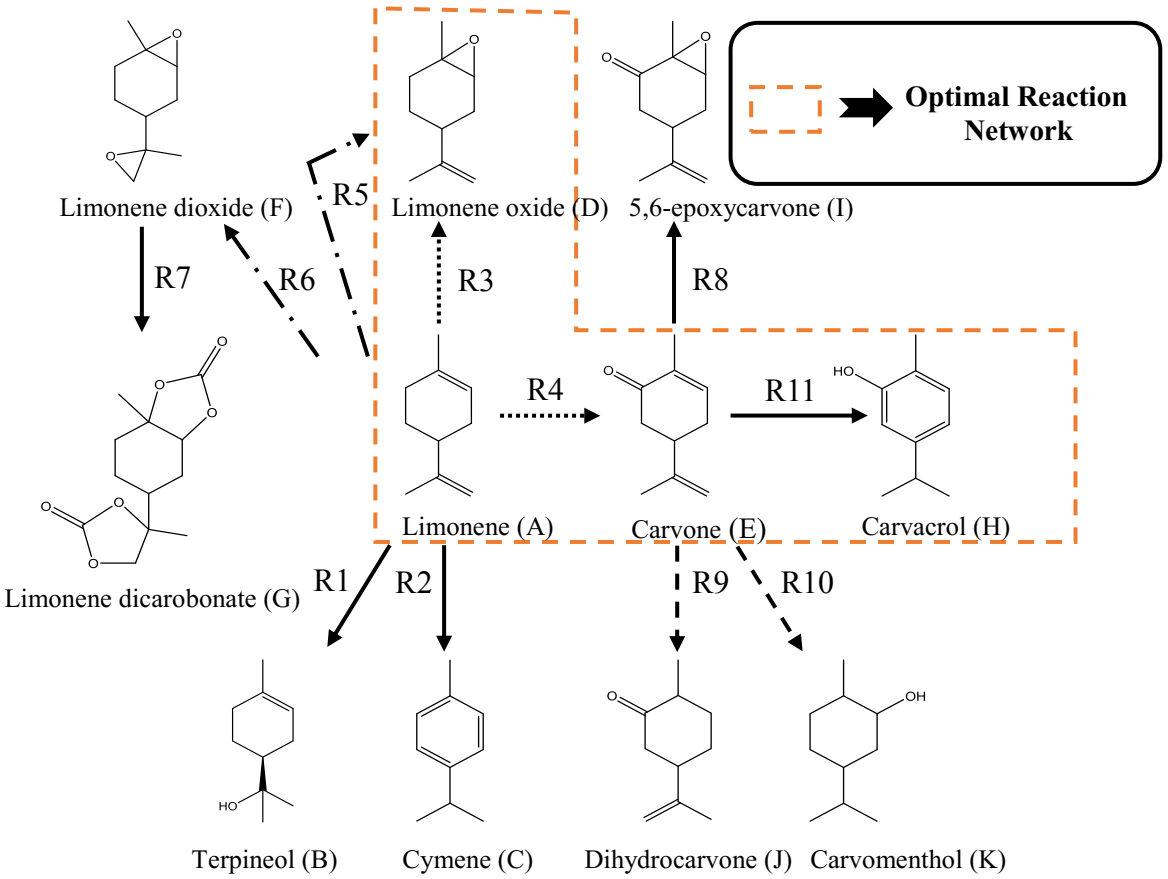


Figure 3.4. Proposed reaction network-reaction. Reactions R3 and R4, R5 and R6, R9 and R10 are all competitive reactions. Optimal reaction network is outlined with the dashed envelope.

In the second master stage, the feed to the reactor is pure carvone, and part of it is from the previous stage with the rest coming from the recycle. Reaction R11 takes place within reator II and produces a two-component mixture, namely the unreacted carvone and carvacrol (H). The mixture is then subject to separation in distillation column II-I. Pure carvone is recycled back to the mixing point while carvacrol is removed and stored as the primary product.

The optimal annual economic potential of the process is \$20.8 million with an estimated annualised separation cost of \$0.56 million. The production rates of limonene oxide and carvacrol are 171.3 and 116.5 mol/min respectively with the consumption rate of d-limonene at 287.8 mol/min. The model statistics and solution summary are given in Table 3.4.

Table 3.4: Model statistics and solution summary of the limonene case

Single Variables	27,953	**MIP Iterations**	227,000
Discrete Variables	5,695	**MIP Nodes Explored**	1,754
Relative Tolerance	0.1	**CPU time (s)**	511.1
Best Possible	22.0×10^6	**Final Solve**	20.8×10^6
Absolute Gap	1.2×10^6	**Relative Gap**	0.059

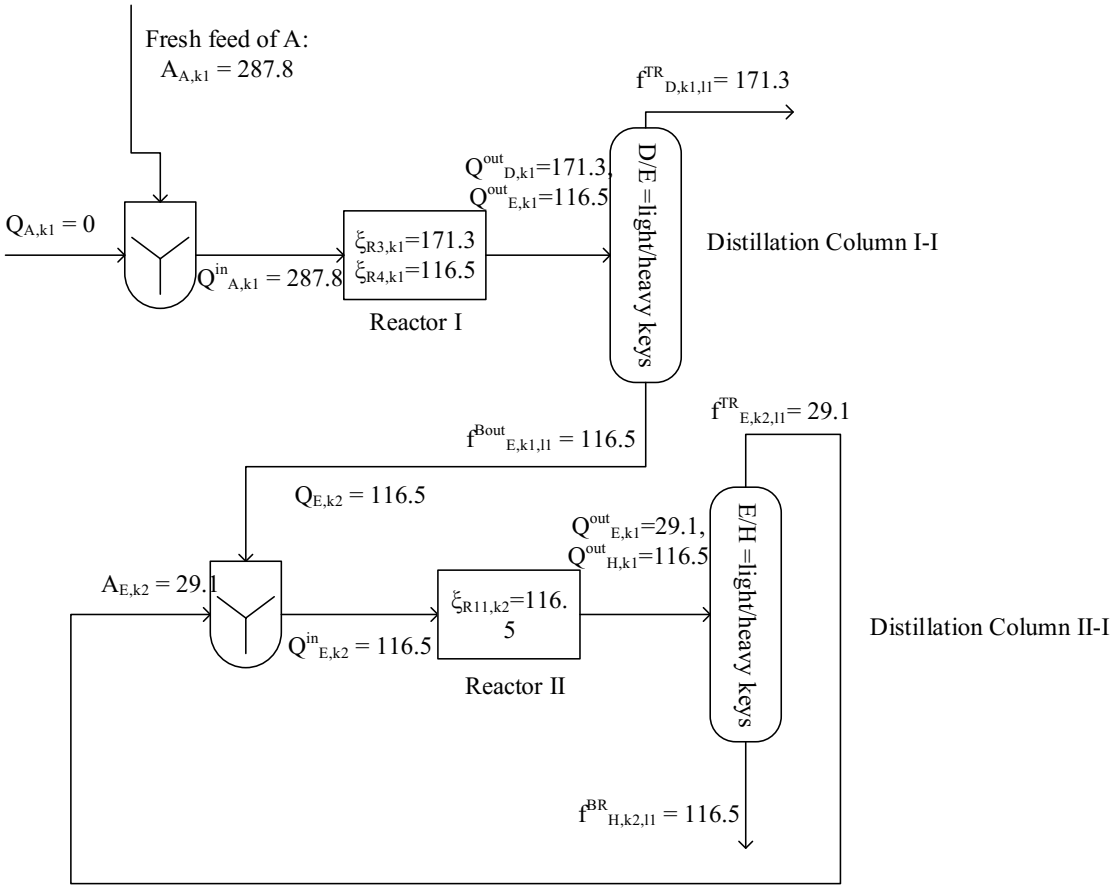


Figure 3.5: Optimal process flowsheet of the d-limonene case, all numbers are in mol/min

3.3.2. Case II: Itaconic Acid Reaction Network

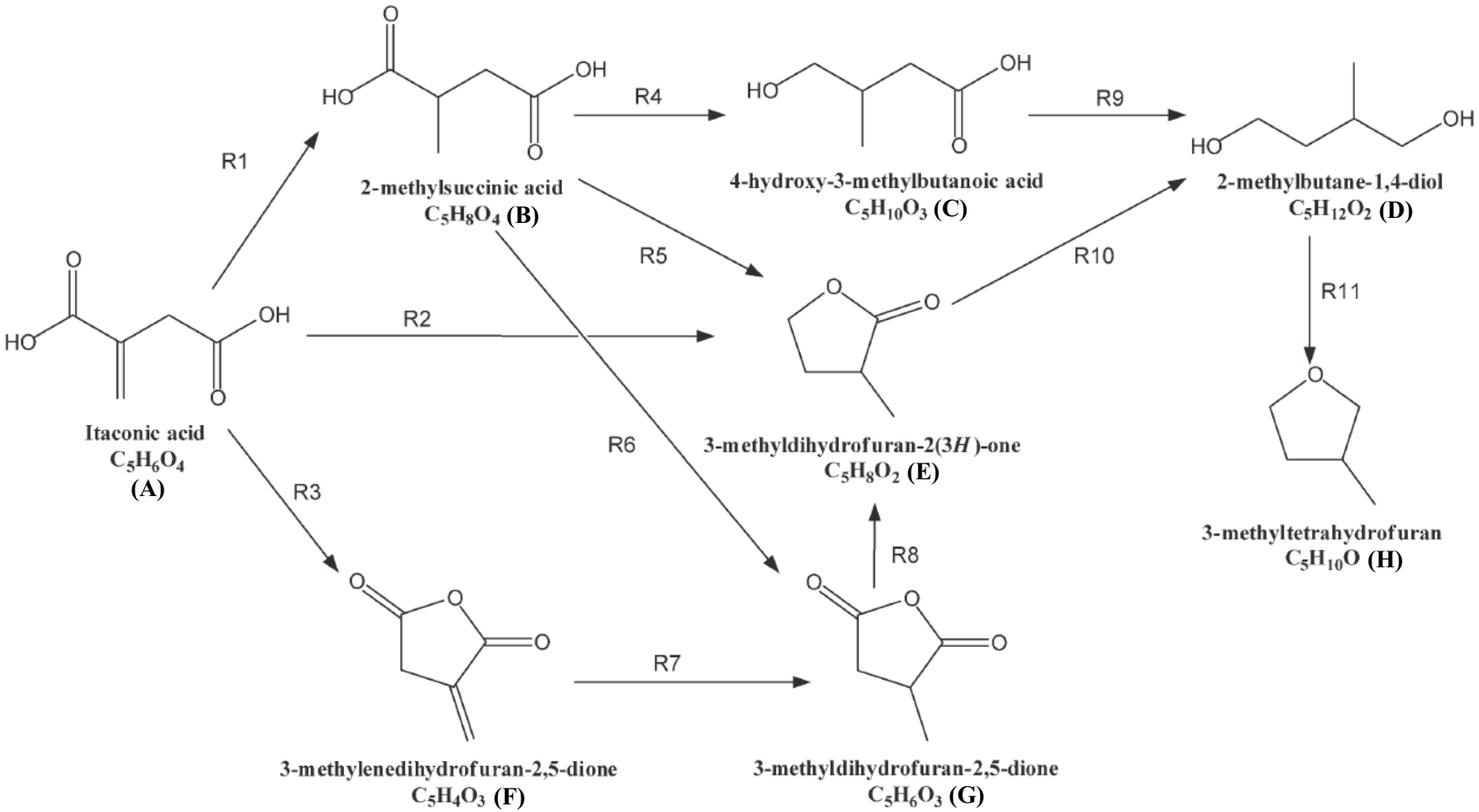


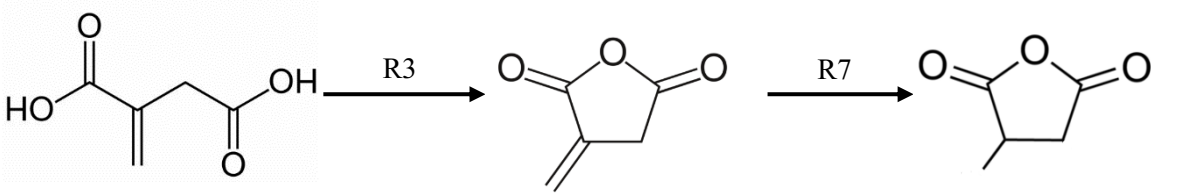
Figure 3.6: Proposed reaction network superstructure for the itaconic acid case

Itaconic acid is commercially produced from the fermentation of glucose mainly by the fungi *Aspergillus terreus* via submerged fungal fermentations (Okabe et al. 2009). The production of itaconic acid remained small at 40,000 t/yr in 2013 due to low demand as a result of few applications and the economic downturn. However, the total manufacturing capacity can be up to 80,000 t/yr, and the market was projected to grow ten stimes by 2020 (Weastra 2013). Therefore, it is reasonable to assume the production capacity can reach 100,000 t/yr in 2018. The same assumptions used in the previous case are also applicable here, and the proposed reaction network superstructure is a sub-network taken from the work conducted by Voll and Marquardt (2012) as shown in Figure 3.6. Boiling points of relevant chemicals are gathered from online databanks, and components market prices, if unknown, are estimated by a group contribution method (Table 3.5). However, due to the lack of literature data for the conversion and selectivity of each reaction mentioned in the network, a constant conversion was used for all reactions, and it is assumed that no simultaneous reactions are present. A sensitivity analysis is thus conducted to determine the effect of changing conversion on the optimal solution.

Table 3.5: Boiling points and estimated prices for components of the itaconic acid case

Component	Molecular Formula	Symbol	Boiling Point (°C)	Estimated Price ($/mol)
Itaconic acid	$C_5H_6O_4$	A	381.4	0.208
2-methylsuccinic acid	$C_5H_8O_4$	B	237	0.221
4-hydroxy-3-methylbutanoic acid	$C_5H_{10}O_3$	C	277	0.208
2-methylbutane-1,4-diol	$C_5H_{12}O_2$	D	212	0.195
3-methyldihydrofuran-2(3H)-one	$C_5H_8O_2$	E	237	0.267
3-methylenedihydrofuran-2,5-dione	$C_5H_4O_3$	F	264	0.259
3-methyldihydrofuran-2,5-dione	$C_5H_6O_3$	G	239	0.284
3-methyltetrahydrofuran	$C_5H_{10}O$	H	89	0.250

The optimal reaction network at a reaction conversion of 90% was generated and is shown in Figure 3.7. Among reactions 1-3, the third one is favoured. It produces itaconic anhydride (F) which is then converted to methyl succinic anhydride (G) via reaction 7. Reaction 8 and subsequent reactions starting from 3-methyldihydrofuran-2(3*H*)-one (E) are deemed to be less favourable due to the low market price of their products and hence are not part of the optimal solution.



Itaconic acid (A) **Itaconic anhydride (F)** **Methyl succinic anhydride (G)**

Figure 3.7: Optimal reaction network of the itaconic acid case

The resulting optimal process flowsheet is shown in Figure 3.8. As can be seen, fresh itaconic acid (A) is fed to the first reactor with a recycle stream and is processed at a conversion of 90% to form itaconic anhydride (F). The product mixture (with unreacted reactant) then enters the distillation column. In the separation unit, the itaconic anhydride stream leaves the top of the distillation stage as distillate and is pumped to the mixing point of the next stage. Purified itaconic acid is recycled and fed to the reactor. Reaction 7 occurs in the 2nd reactor, and the product stream contains both the primary product-methyl succinic anhydride and unreacted itaconic anhydride which is then isolated in the 2nd separation unit and recycled.

The detailed model statistics and solution summary are given in Table 3.6.

Table 3.6: Model statistics and solution summary of the itaconic acid case

Single Variables	20,889	**MIP Iterations**	19,343
Discrete Variables	5,020	**MIP Nodes Explored**	205
Relative Tolerance	0.05	**CPU time (s)**	15.1
Best Possible	11.3×10^6	**Final Solve**	10.9×10^6
Absolute Gap	0.39×10^6	**Relative Gap**	0.036
Separation cost ($mil)	0.289	**Product rate (mol/min)**	301.7

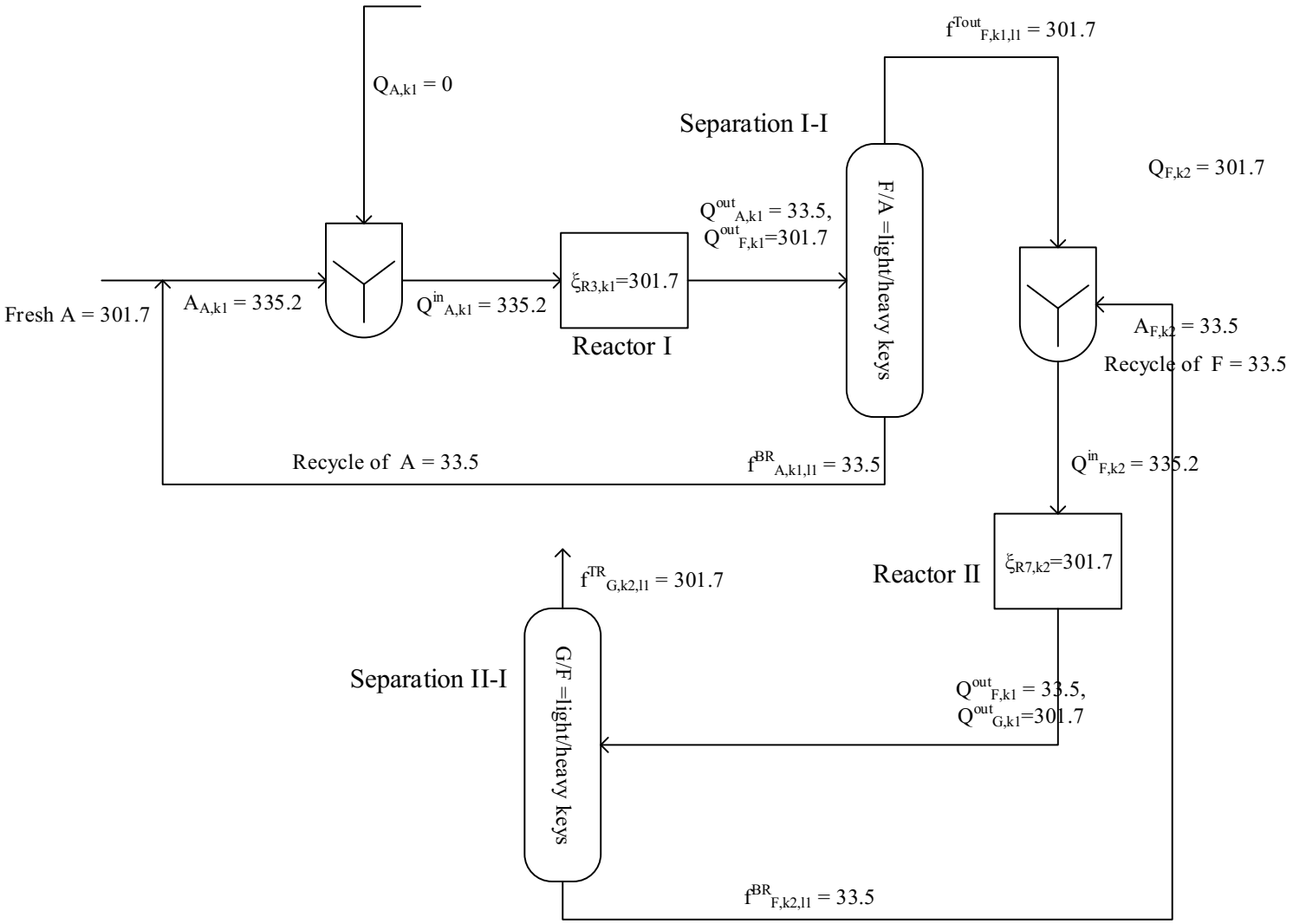


Figure 3.8: Optimal process flowsheet of the itaconic acid case, all numbers are in mol/min

The sensitivity analysis conducted has shown that changing the reaction conversions does not alter the basic structure of the optimal reaction network and its corresponding process flowsheet. However, the reduction in the conversion does lead to an increase in the flowrates of recycle streams, which increases

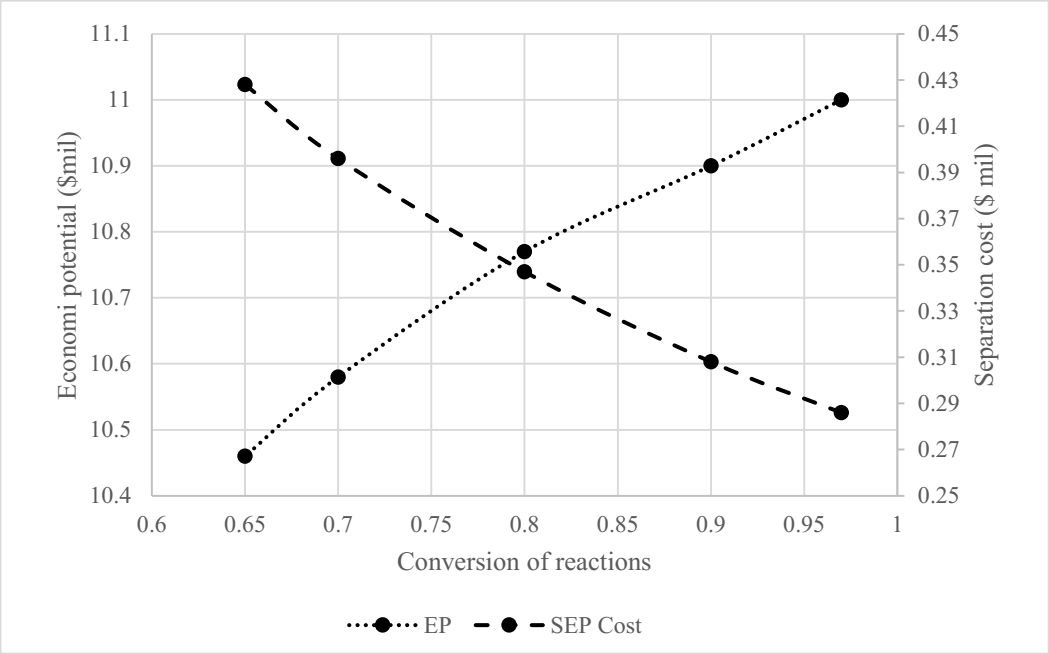


Figure 3.9: Sensitivity analysis of the effect of changing conversions on the EP and Separation Cost

the separation cost as shown in Figure 3.9. This is because that the lower the conversion is, the higher the recycle rate will be to maintain the extent of each reaction at constant leading to higher separation inlets and costs (demonstrated in Table 3.7). As a result, the overall economic potential reduces as conversions of the reactions decrease.

Table 3.7: Flowrates around reactor I and separation cost VS conversion

Conversion	Extent of reaction (mol/min)	Recycle rate (mol/min)	Separation input(mol/min)	Separation cost ($ mil)
0.97	301.7	9.3	311.0	0.289
0.9	301.7	33.5	335.2	0.311
0.8	301.7	75.4	377.1	0.345
0.7	301.7	129.3	431.0	0.395
0.65	301.7	162.5	464.2	0.428

3.3.3. Group Contribution Method

It is assumed that the cost of a molecule is well approximated by the sum of the costs of its constituent groups. A group contribution method can be defined as another optimisation problem: given a set of common compounds i including alkanes, alkenes, alcohols, organic acids and aromatic compounds, the

actual costs of them, $Cost_i^{actual}$ can be acquired from different online sources and thus an estimation of the costs of different groups j, $Cost_j$ can also be made via the least square method shown in equation 3.79.

$$min \sum_i (Cost_i^{actual} - \sum_j Cost_j n_{i,j})^2 \quad 3.79$$

To estimate the market value of each chemical involved in the two case studies, the well-known chemicals selected are components C1 to C17 listed in Table 3.8. All the market prices are bulk prices taken from Alibaba (Alibaba 2015). Molecular groups are classified based on the theory developed by Jankowski et al. (2008). The number of those groups in components C1 to C17 and the estimated cost of different groups j are shown in Table 3.9. Table 3.10 shows the number of various groups that are present in different chemicals involved in both the d-limonene and itaconic acid cases. The estimated market values presented in Table 3.11.

Table 3.8: Market prices of well-known chemicals

Component i	Identification	Market Price ($/mol)	Deviation from GC methods (%)
C1	Cyclohexane	0.126	9.7
C2	Methylcyclohexane	0.245	0.62
C3	Methanol	0.0256	9.24
C4	Cyclohexanol	0.250	0.96
C5	Cyclohexene	0.131	0.32
C6	D-limonene	0.409	3.0
C7	p-cymene	0.537	0.03
C8	cyclohexene oxide	0.147	0.14
C9	alpha-pinene oxide	0.457	0.07
C10	Alpha-terpineol	0.540	0.02
C11	Carvone	0.451	2.81
C12	Benzene	0.109	0.18
C13	Succinic anhydride	0.150	0.07
C14	cyclohexanone	0.118	10.4
C15	propane	0.0529	7.41
C16	n-butane	0.0697	3.23
C17	itaconic acid	0.208	0.08

Table 3.9: Number of groups ***j*** in each known chemical and the estimated price of each group ***j***

Compound *i*/Groups *j*		**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**	**C15**	**C16**	**C17**	**Estimated Price ($/mol)**
$-CH_3$	**G1**	0	1	1	0	0	2	3	0	3	3	2	0	0	0	2	2	0	0.013
$> CH_2$	**G2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	0.023
$> CH_2$ in a ring	**G3**	6	5	0	5	4	3	0	4	2	3	2	0	2	5	0	0	0	0.019
$= CH -$ in a ring	**G4**	0	0	0	0	2	1	0	0	0	1	1	0	0	0	0	0	0	0.028
$= CH -$ in an aromatic ring	**G5**	0	0	0	0	0	0	4	0	0	0	0	6	0	0	0	0	0	0.018
$> C =$ in a ring	**G6**	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0.085
$> C =$ in an aromatic ring	**G7**	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0.170
$> CH -$	**G8**	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.085
$> CH -$ in a ring	**G9**	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0.139
$> CH -$ in fused rings	**G10**	0	0	0	0	0	0	0	2	3	0	0	0	0	0	0	0	0	0.015
$> C =$	**G11**	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0.043
$= CH_2$	**G12**	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0.043
$> C <$	**G13**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0.177
$> C <$ in a ring	**G14**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0.146
$> C <$ in fused rings	**G15**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0.146
$-O -$ in a ring	**G16**	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0.041
$-OH$	**G17**	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	2	0.014
$> C = O$	**G18**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.036
$> C = O$ in a ring	**G19**	0	0	0	0	0	0	0	0	0	0	1	0	2	1	0	0	0	0.036

Table 3.10: Number of groups ***j*** in each chemical involved in both case studies

Components *i′*/Groups	**G1**	**G2**	**G3**	**G4**	**G5**	**G6**	**G7**	**G8**	**G9**	**G10**	**G11**	**G12**	**G13**	**G14**	**G15**	**G16**	**G17**	**G18**	**G19**
Components in the d-limonene case																			
D-Limonene	2	0	3	1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0
Limonene oxide	2	0	3	0	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0
Limonene dioxide	2	0	4	0	0	0	0	0	1	1	0	0	0	1	1	2	0	0	0
α-terpineol	3	0	3	1	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0
p-cymene	3	0	0	0	4	0	2	1	0	0	0	0	0	0	0	0	0	0	0
carvone	2	0	2	1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	1
Limonene dicarbonate	2	0	4	0	0	0	0	0	1	1	0	0	0	1	1	4	0	0	2
Carvacrol	3	0	0	0	3	0	3	1	0	0	0	0	0	0	0	0	1	0	0
5,6-epxoycarvone	2	0	2	0	0	0	0	0	2	0	1	1	0	1	0	1	0	0	1
Dihydrocarvone	2	0	3	0	0	0	0	0	2	0	1	1	0	0	0	0	0	0	1
Carvomenthol	3	0	3	0	0	0	0	1	3	0	0	0	0	0	0	0	1	0	0
Components in the itaconic acid case																			
itaconic acid	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2	2	0
2-methylsuccinic acid	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	2	0
4-hydroxy-3-methylbutanoci acid	1	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	1	0
2-methylbutane-1,4-diol	1	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0
3-methyldihydrofuran-2(3H)-one	1	0	2	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1
3-methylenedihydrofuran-2,5-dione	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	2
3-methyldihydrofuran-2,5-dione	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	2
3-methyltetrahydrofuran	1	0	3	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0

Table 3.11: Estimated price ($/mol) for chemicals involved in this chapter

Compounds	Price $/mol(from GAMs)
D-Limonene	0.421
Limonene oxide	0.511
Limonene dioxide	0.631
alpha-terpineol	0.54
p-cymene	0.537
carvone	0.438
Limonene dicarbonate	0.784
Carvacrol	0.703
5,6-epxoycarvone	0.651
Dihydrocarvone	0.483
Carvomenthol	0.614
itaconic acid	0.208
2-methylsuccinic acid	0.221
4-hydroxy-3-methylbutanoci acid	0.208
2-methylbutane-1,4-diol	0.195
3-methyldihydrofuran-2(3H)-one	0.267
3-methylenedihydrofuran-2,5-dione	0.259
3-methyldihydrofuran-2,5-dione	0.284
3-methyltetrahydrofuran	0.25

3.4. Concluding Remarks

In this chapter, an MILP optimisation model along with a conceptual superstructure for the preliminary process synthesis of optimal chemical reaction networks integrated with distillation sequencing is presented. The model allows us to quickly scan all the reaction pathways as well as different distillation sequences. Recycles of unreacted chemicals and direct removal of pure products or wastes are also implemented for the model to be more robust and to account for incomplete conversion. Another unique aspect of this model is that the innovative formulation of the separation mechanism, which not only allows the detection of components entering each distillation stage but also provides constraints for the system to logically select both light and heavy keys. The use of bi-linear conditions for the bypass mass balance within each separation unit also increases the flexibility and robustness of the algorithm. Two case studies have been solved to demonstrate the effectiveness and applicability of the developed approach. The case studies covered two popular bio-derived chemicals with the objective of maximising the plant economic potential. The solutions of the case studies illustrate the ability of the proposed method to generate both the optimal reaction network and process flowsheet.

In the next chapter, the implementation of energy balances to model process synthesis with simultaneous heat integration is discussed.

4. Chapter IV: Process Synthesis with Simultaneous Heat Integration

As mentioned in the literature review, heat integration plays a vital role in process synthesis and design as it has been proven to boost the energy utilisation efficiency of chemical processes significantly. Therefore, this chapter presents a state-of-the-art mixed integer linear programming (MILP) optimisation method, which incorporates simultaneous heat integration with the process synthesis optimisation approach introduced in Chapter III. In this innovative approach, the temperature variable is discretised to be used as a new set domain instead of being treated as a continuous variable to maintain the model's linearity. This chapter starts with defining the problem to be tackled in section 4.1 followed by the mathematical formulation in section 4.2. The limonene case study, which is used in the previous chapter, is used in section 4.3 to demonstrate the effect of heat integration on the optimal solutions. Sensitivity analyses are conducted in section 4.4 to investigate the model's reaction to several key parameters. The results are summarised in the concluding remarks in section 4.5.

4.1. Problem Statement

The same problem as the previous chapter is solved but with the addition of new sets of constraints that are related to heat integration and require the following additional data:

- additional reaction thermodynamic data (reaction temperature and standard enthalpy of reaction),
- thermodynamic data of each reactant and product in the reaction network (e.g. enthalpy of vaporisation, standard enthalpy of formation and specific heat capacity),
- and utility costs of steam and cooling water.

The new optimal solution now not only includes the flow charts and other results mentioned in chapter III but also provides the following heat integration solutions:

- amount of external utilities required (cooling water and steam),
- process flowsheet including reaction, separation sequences and heat exchangers,
- and the optimal heat integration cascade

to maximise the economic potential of the process. Figure 4.1 summarises the general relationship between Framework I and II.

The same set of assumptions used in chapter III is also applicable in this chapter. However, besides market prices of chemicals, thermodynamic properties including standard enthalpy of formation, enthalpy of evaporation and specific heat capacity, if unavailable online, are all estimated via the group contribution method (Zábranský and Růžička 2004, Verevkin et al. 2013). In addition, thermodynamic data of a mixture is simplified to be equal to the sum of the properties of the individual chemicals (i.e. perfect mixing assumption). Therefore, the temperatures of separation inlet streams are always adjusted

to the boiling point of the lightest component (the component in the mixture with the lowest boiling point) to keep the streams in the liquid phase. Heat integration is performed based on the principle developed by Douglas (1988) but takes place simultaneously rather than sequentially. The model is designed to obtain a simultaneous optimal heat integration cascade of the process while maintaining the integrity of the original framework. The results can help to gain further insight of the optimal solution and to pave the way for more rigorous process analysis.

The additional equations are presented in the following sections.

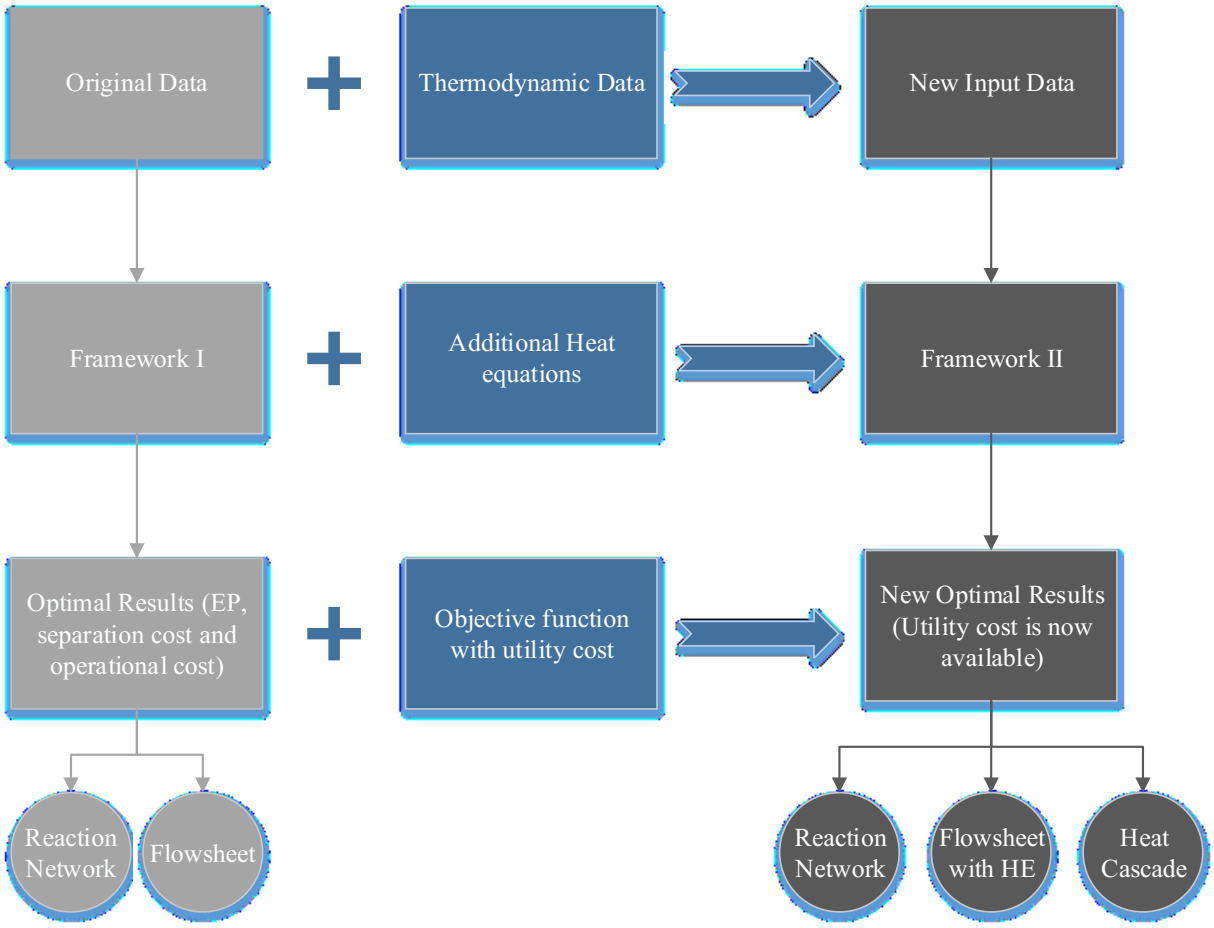


Figure 4.1: Implementation of the new constraints and objective function

4.2. Mathematical Formulation

All new equations are linear using the following notation:

Indices	Descriptions
i	Components
j	Reactions
k	Master Stages
l	Separation Stages
t	Discretised temperature points

p	Discretised split fraction points

Superscript-Mstream (Master stage streams)

A	Addition streams to each reaction stage
$Qnext$	Reactor outlet by pass to next reaction stage
$Qsepin$	Reactor outlet to the separation system
QR	Reactor outlet bypass to direct removal

Superscript-Dstream (Lower level separation stage streams)

$Tout$	Distillate of each distillation column to next reaction stage
TR	Distillate to direct removal of each distillation column
$Bout$	Bottom product of each distillation column to next reaction stage
BR	Bottom product to direct removal of each distillation column

Binary Variables

$\hat{x}_{j,t}$	$= 1$ if the reaction temperature of j is at t; 0 otherwise
$\hat{y}_{k,t}$	$= 1$ if the temperature of reaction stage k is at t; 0 otherwise
$y_{k,t}^{Qin}$	$= 1$ for all temperature t within the range of 0 ℃ to the temperature of the reactor inlet stream at stage k; 0 otherwise
$y_{k,t}^{Qout}$	$= 1$ for all temperature t within the range of 0 ℃ to the temperature of the reactor outlet stream at stage k; 0 otherwise
EXO_k	$= 1$ is reactor stage k is exothermic; 0 otherwise
$\widehat{ys}_{k,t}^{Mstream}/\widehat{ys}_{k,l,t}^{Dstream}$	$= 1$ if the starting temperature of the stream is at t; 0 otherwise
$\widehat{ye}_{k,t}^{Mstream}/\widehat{ye}_{k,l,t}^{Dstream}$	$= 1$ if the end temperature of the stream is at t; 0 otherwise
$HOT_k^{Mstream}/$ $HOT_{k,l}^{Dstream}$	$= 1$ if the inlet temperature of the heat exchanger of the stream is higher than the outlet temperature; 0 otherwise
$TI_{k,t}^{Mstream}/TI_{k,l,t}^{Dstream}$	$= 1$ if temperature point t lies within the starting and end temperature of the stream; 0 otherwise
$BpM_{i,t}$	$= 1$ if the boiling point of component i is at t; 0 otherwise
$TopCT_{i,k,l}$	$= 1$ if component i is the one with the lowest boiling point in separation stage l of k; 0 otherwise
$N_{i,k,l}^{fin}$	$= 1$ component i exists in the inlet stream to separation column l of stage k; 0 otherwise
$Yb_{i,k,l}$	$= 1$ if component i is chosen as the heavy key at separation stage l and in master stage k; 0 otherwise
$N_k^{Mstream}/N_{k,l}^{Dstream}$	$= 1$ if the flowrate of the stream is not zero; 0 otherwise

Continuous Variables

$A_{i,k}$	Flowrate of component i added to the master mixing stage k
$\Delta EC_{k,l}$	Condenser cooling duty of column l at stage k
$\Delta ER_{k,l}$	Reboiler heating duty of column l at stage k
ΔE_k	Isothermal energy change of reaction stage k
E_k^{in}	Flow energy of the inlet stream to reactor at stage k
E_k^{out}	Flow energy of the outlet stream of reactor at stage k
$E_{k,l}^{fT}$	Flow energy of the distillate of column l at stage k
$E_{k,l}^{fB}$	Flow energy of the bottom product of column l at stage k
$E_{k,l}^{fin}$	Flowrate energy of the separation inlet stream to column l at stage k
T_k^{Re}	Reaction temperature of each reaction stage k
$T_{k,l}^{CD}$	Operating temperature of condenser in column l of k
$T_{k,l}^{RB}$	Operating temperature of reboiler in column l of k
$Ts_k^{Mstream}/Ts_{k,l}^{Dstream}$	Temperature of streams prior to heat exchangers
$Te_k^{Mstream}/Te_{k,l}^{Dstream}$	Temperature of streams leaving heat exchangers
$SHA_{k,t}^{Mstream}/$ $SHA_{k,l,t}^{Dstream}$	The actual heat surplus of different streams
$SH_{k,t}^{Mstream}/$ $SH_{k,l,t}^{Dstream}$	The heat surplus used for heat integration analysis of each stream
$FCp_{k,t}^{Mstream}/FCp_{k,l,t}^{Dstream}$	Flow heat capacity of the streams
$EP3$	Economic potential of the process including energy cost
$EP2$	Economic potential of the process without energy cost
Z	Objective function
$energy\ cost$	Total energy cost of the process
CWC^T	Total cooling water cost of the process
CWC^{unit}	Unit cost of cooling water
SC^T	Total steam cost of the process
SC^{unit}	Unit cost of steam
SH_t^{Total}	Total heat surplus of the process at temperature t
CH_t	Cascaded heat of the process at temperature t
$CH_t^{afterheating}$	Cascaded heat after minimum heating is provided at temperature t

$CH_t^{aftercooling}$	Cascaded heat after minimum heating and cooling are provided at temperature t
Q_{min}^{cold}	Minimum cooling duty required of the process
Q_{min}^{hot}	Minimum heating duty required of the process

The material balances, separation rules and the concepts of master reaction stages k along with affiliated distillation stages l have already been discussed. In this extended method, heat exchangers are installed for the cooling or heating of various streams prior to mixing or right after splitting points. In addition to the energy change in streams, the enthalpy change of reactors (ΔE_k), cooling duty of condensers ($\Delta EC_{k,l}$) and heating duty of reboilers ($\Delta ER_{k,l}$) are taken into consideration for the heat integration as they also play important roles in the heat cascade. The general concepts, as illustrated in Figure 4.2, are as follows:

(i) the addition stream of each stage k ($A_{i,k}$) is firstly heated to match the reaction temperature of that stage (T_k^{Re}) before entering the mixer (mix_k),

(ii) enthalpy change of each reaction is calculated by performing an energy balance around the reactor which is subject to either constant cooling or heating to maintain isothermal operation,

(iii) the temperature of the separation inlet stream ($Q_{i,k}^{sepin}$) is adjusted to be equal to the boiling point of the lightest component in the stream by a heat exchanger before entering the separation system,

(iv) on the other hand, the temperature of the reactor outlet bypass stream Q_k^{next} is adjusted to match the reactor temperature of the next stage (T_{k+1}^{Re}) via a heat exchanger prior to mixing (mix_{k+1})

(v) the direct removal stream of the reaction outlet ($Q_{i,k}^{R}$) is assumed to be cooled to ambient temperature ($T^{\emptyset}$) by a cooler

(vi) the distillate of each column is condensed via a total condenser with a cooling duty of $\Delta EC_{k,l}$ while the bottom stream exits via a partial reboiler with a heating duty of $\Delta ER_{k,l}$

(vii) finally, the product streams ($f_{i,k,l}^{TR}/f_{i,k,l}^{BR}$) of each column are cooled to ambient temperature and the temperatures of distillation products that enter the next reaction stage ($f_{i,k,l}^{Tout}/f_{i,k,l}^{Bout}$) are adjusted to T_{k+1}^{Re} before mixing.

4.2.1. General Approach

Based on the principle of Douglas (1988), heat integration must follow two rules: (1) heat can be transferred from hot streams in one temperature interval (i.e. $T_1 \leq t \leq T_2$) to cold streams in that same interval; (2) to ensure driving force for heat transfer, each hot stream temperature should be reduced by

the minimum approach temperature before intervals are constructed. As mentioned above, the temperature variable is discretised to form a new index $t \in T$. As a result, two new binary variables known as the temperature point (TP) variables are introduced for each *stream* in this paper to identify the respective temperatures of inlet ($\widehat{ys}_t{}^{stream}$) and outlet ($\widehat{ye}_t{}^{stream}$) streams of each heat exchanger (i.e. if the temperature of the inlet stream of the heat exchanger is at temperature t, then $\widehat{ys}_t{}^{stream} = 1$, and 0 otherwise). The temperature interval of each heat exchanger ($TI_t^{stream} = 1$ if the temperature range includes temperature point t) is determined by the inlet and outlet temperatures via a constraint. Usually, temperature intervals are modified before being used for heat integration analysis (i.e. subtraction by the minimum approach temperature for hot streams). However, unlike conventional heat integration problems, the property of each stream involved in heat integration (hot or cold) in this problem is a variable and is revealed as a part of the solution. Therefore, evaluating the heat surplus directly over the modified intervals is not achievable. To mitigate the issue, the actual heat surpluses of each stream (SHA_t^{stream}) at different temperature points (t) are firstly calculated:

$$SHA_t^{stream} = FCp_t^{stream} TI_t^{stream}, \forall t \quad (4.1)^*$$

where FCp_t^{stream} is the flow heat capacity of each stream calculated as the sum of the product of the flowrate and specific heat capacity of all components in that stream at temperature t (The same definition is applicable for flow heat capacity of each stream throughout this chapter).The adjusted cascaded heat surpluses (SH_t^{stream}) based on the stream property (HOT^{stream} as a variable) are then given as follows:

$$SH_t^{stream} = -SHA_t^{stream}, \forall\, t, if\ HOT^{stream} = 0 \quad (4.2)$$

$$SH_t^{stream} = SHA_{t+T_{gradient}}^{stream}, \forall\, t, if\ HOT^{stream} = 1 \quad (4.3)$$

where $T_{gradient}$ is the minimum approach temperature allowed for energy to transfer between streams.

Note that all equations marked with * in the following sections are bi-linear equations and can be formally linearised in the same way as in Chapter III to fit in the MILP model.

4.2.2. Objective Function

In this work, the primary objective is to combine heat integration with process synthesis to produce the optimal process flowsheet and heat cascade. Generally, an optimal or near optimal heat exchanger network exhibits the following characteristics: (a) minimal utility cost; (b) a minimal number of heat exchanger units; (c) minimal investment cost for heat exchangers. However, it is possible to have conflicts among these characteristics. Therefore, it is assumed in this chapter that (a) has precedence over (b), and (b) over (c) (Biegler et al. 1997). The general expression for the objective function is given as follows:

$$EP3 = EP2 - energy\ cost \quad (4.4)$$

$$Z = \max EP3 \tag{4.5}$$

where $EP2$ is the economic potential of the process based on the analysis of the raw material and separation cost and has been discussed in the previous chapter. The energy cost, on the other hand, includes both the total cost of cooling water (CWC^T) and steam (SC^T), i.e.

$$energy\ cost = CWC^T + SC^T \tag{4.6}$$

To calculate the total cost of utilities, the total heat surplus at each temperature t (SH_t^{Total}) is firstly calculated via the following equation:

$$SH_t^{Total} = \sum_k (SH_{k,t}^{Re} + SH_{k,t}^{A} + SH_{k,t}^{Qnext} + SH_{k,t}^{QR} + SH_{k,t}^{Qsepin}) + \sum_{k,l} (SH_{k,l,t}^{RB} + SH_{k,l,t}^{CD} + SH_{k,l,t}^{Tout} + SH_{k,l,t}^{Bout} + SH_{k,l,t}^{TR} + SH_{k,l,t}^{BR}), \forall t \tag{4.7}$$

The cascaded heat (CH_t) at each temperature interval t is thus:

$$CH_t = \sum_{t'>t} SH_{t'}^{Total}, \forall t \tag{4.8}$$

The minimum heating duty required for the system (Q_{min}^{hot}) is provided to the cascade to neutralise the temperature interval t with the most significant energy deficit and is formulated as follows:

$$CH_t^{afterheating} = CH_t + Q_{min}^{hot}, \forall t \tag{4.9}$$

$$CH_t^{afterheating} \geq 0 \tag{4.10}$$

The minimum cooling duty of the system (Q_{min}^{cold}) is measured as the cooling duty required to meet the largest heat surplus of the cascade given as follows:

$$CH_t^{aftercooling} = CH_t^{afterheating} - Q_{min}^{cold}, \forall t \tag{4.11}$$

$$CH_t^{aftercooling} \leq 0 \tag{4.12}$$

where $CH_t^{afterheating}$ and $CH_t^{aftercooling}$ each represents the cascaded heat surplus of each temperature interval after minimum heating and cooling are provided to the process respectively. The total cost of cooling water and stream are thus given as follows:

$$CWC^T = CWC^{unit} Q_{min}^{cold} \tag{4.13}$$

$$SC^T = SC^{unit} Q_{min}^{hot} \tag{4.14}$$

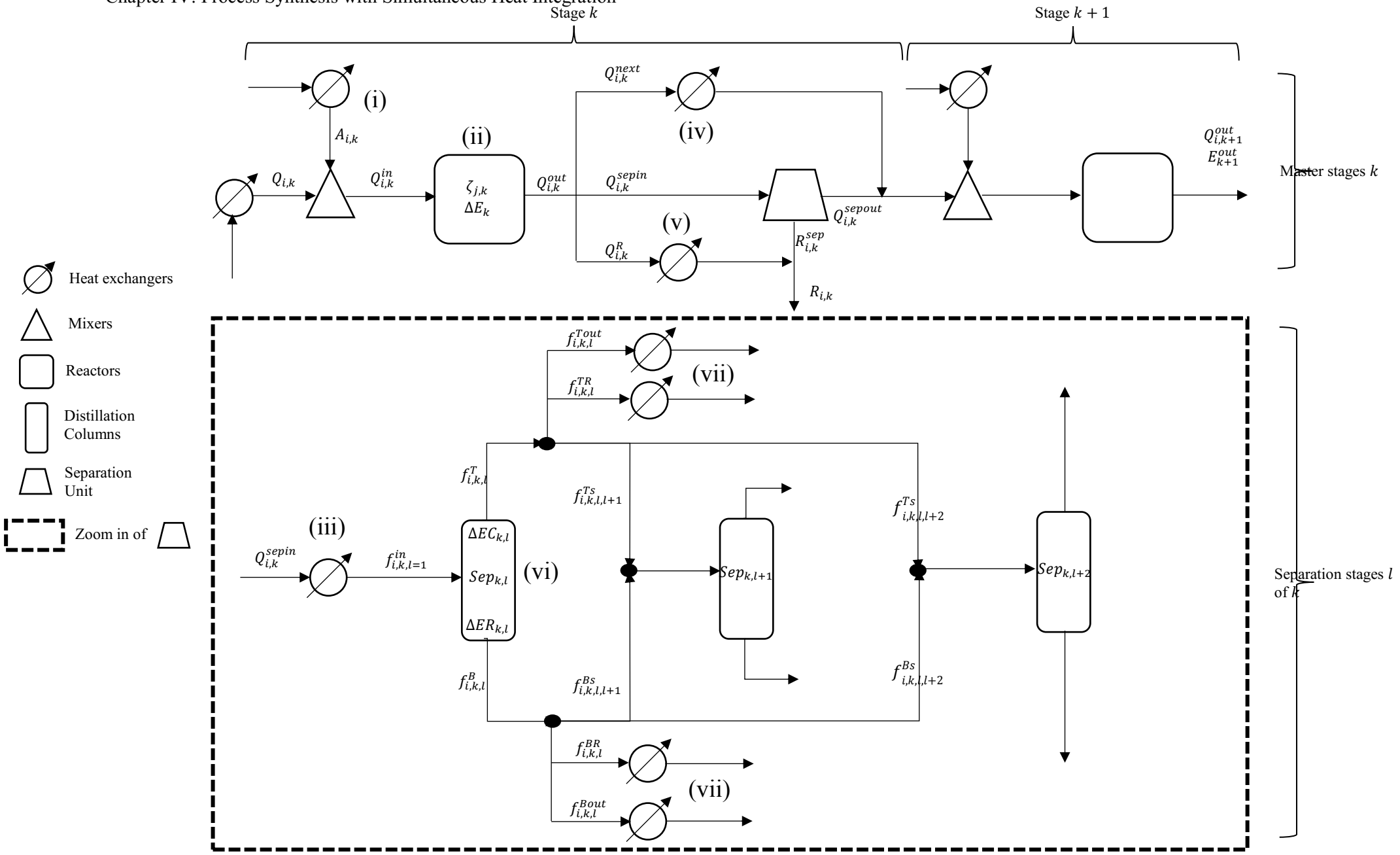


Figure 4.2: Conceptual diagram of the process superstructure

4.2.3. Reactor Heat Surplus

Each reaction stage operates at a specific temperature (T_k^{Re}) which is determined by the temperature of the reaction that occurs at that stage (TR_j). Therefore, a new binary variable known as the temperature point of each reaction ($\hat{x}_{j,t}$) is introduced. It is used to identify (in a binary term) the temperature of each reaction, i.e. if reaction j occurs at temperature t, the binary variable $\hat{x}_{j,t} = 1$, and 0 otherwise:

$$\hat{x}_{j,TR_j} = 1, \forall\, j \tag{4.15}$$

The temperature point of that reaction stage ($\hat{y}_{k,TR_j}$) is equal to $\hat{x}_{j,TR_j}$ only if reaction j occurs in stage k (i.e. $y_{j,k} = 1$), and only one temperature can be assigned to each reactor, i.e.

$$\hat{y}_{k,t} \geq \hat{x}_{j,t} + y_{j,k} - 1, \forall j, k, t \tag{4.16}$$

$$\sum_t \hat{y}_{k,t} \leq 1, \forall k \tag{4.17}$$

The actual temperature of each reaction stage k (T_k^{Re}) is thus written as:

$$T_k^{Re} = \sum_t t \cdot \hat{y}_{k,t}, \forall k \tag{4.18}$$

Therefore, the temperature of the inlet flow to the reactor ($Q_{i,k}^{in}$) is preheated to T_k^{Re}. The binary variable $y_{k,t}^{Qin}$ is defined so that $y_{k,t}^{Qin} = 1$ for all temperature t within the range of 0 ℃~T_k^{Re}:

$$y_{k,t}^{Qin} \geq y_{k,t+1}^{Qin}, \forall k, t \tag{4.19}$$

$$y_{k,t}^{Qin} \geq \hat{y}_{k,t}, \forall k, t \tag{4.20}$$

By assuming the reference temperature in this work to be 0 ℃, the relative flow energy (E_k^{in}) of the reaction inlet stream, which also includes the energy change involved in phase change (e.g. vaporisation) is thus given as follows:

$$E_k^{in} = \sum_{i,t} \left(Cp_{i,t} + HoV_{i,t}\right) y_{k,t}^{Qin} Q_{i,k}^{in}, \forall k \tag{4.21}*$$

where $Cp_{i,t}$ accounts for both liquid and gaseous phase specific heat capacity of component i at temperature t and $HoV_{i,t}$ is the enthalpy of vaporisation of component i at temperature t (Note, $HoV_{i,t}$ = enthalpy of component i if the temperature is equal to the boiling point of component i, 0 otherwise).

Similarly, the relative flow energy of the reaction outlet stream (E_k^{out}) is formulated as follows:

$$E_k^{out} = \sum_{i,t} \left(Cp_{i,t} + HoV_{i,t}\right) y_{k,t}^{Qout} Q_{i,k}^{out}, \forall k \tag{4.22}*$$

with $y_{k,t}^{Qout} = y_{k,t}^{Qin}$ as reactions are isothermal.

The overall energy balance of each reaction stage is simply:

$$\sum_i C_{pli}\left(Q_{i,k}^{in} T^{\emptyset}\right) - E_k^{in} + \sum_j \Delta HR_j^{\emptyset} \xi_{j,k} + E_k^{out} - \sum_i C_{pli}\left(Q_{i,k}^{out} T^{\emptyset}\right) = \Delta E_k, \forall k \quad (4.23)$$

where $\Delta HR_j^{\emptyset}$ is the standard heat of reaction j, $T^{\emptyset}$ is the temperature under standard conditions and is set to be equal to the ambient temperature, $\xi_{j,k}$ is the extent of reaction j in stage k, and ΔE_k is the enthalpy change of reaction stage k. The actual heat surplus of each reaction stage at temperature t ($SHA_{k,t}^{Re}$) is defined as follows:

$$SHA_{k,t}^{Re} = -\Delta E_k \cdot \hat{y}_{k,t}, \forall k, t \quad (4.24)^*$$

In order to fit the heat surplus of each reaction into the overall heat integration analysis, each reaction is either treated as a hot stream (exothermic) or cold stream (endothermic) and releases or absorbs energy at one particular temperature. The reaction stage heat surplus ($SH_{k,t}^{Re}$), which takes the reactor's exothermicity into consideration, is thus given as:

$$SH_{k,t}^{Re} = SHA_{k,t+T_{gradient}}^{Re} EXO_k - SHA_{k,t}^{Re}(1 - EXO_k), \forall k, t \quad (4.25)^*$$

where EXO_k is a binary variable defined as the exothermicity of each reaction stage and is constrained as follows:

$$EXO_k \geq y_{j,k}, \forall k, j \in ExoR \quad (4.26)$$

$$1 - EXO_k \geq y_{j,k}, \forall k, j \in EndoR \quad (4.27)$$

4.2.4. Reboilers and Condensers

For the analysis of reboiler and condenser utilities, two fundamental assumptions are made. The first one is that the operating temperature of the reboiler is set to be equal to the boiling point of the heavy key component of that column to ensure that none of the heavy non-key components are returned to the column via boil-up. Therefore, the temperature point of each reboiler ($\hat{y}_{k,l,t}^{RB}$) is defined as follows:

$$\hat{y}_{k,l,t}^{RB} = \sum_i Yb_{i,k,l}\, BpM_{i,t}, \forall k, l, t \quad (4.28)$$

where $Yb_{i,k,l}$ is the binary variable used for the selection of the heavy key in each separation stage (i.e. $Yb_{i,k,l} = 1$ if component i is the heavy key in column l of k) and $BpM_{i,t}$ is defined so that $BpM_{i,t} = 1$ if the boiling point of component i is t, 0 otherwise. The second assumption is that the operating temperature of each condenser is set to be equal to the boiling point of the lightest component (i.e. most volatile component) in that column to allow total condensation of all components in the distillate:

$$\hat{y}_{k,l,t}^{CD} = \sum_i TopCT_{i,k,l}\, BpM_{i,t}, \forall k, l, t \quad (4.29)$$

where $\hat{y}_{k,l,t}^{CD}$ is the temperature point of each condenser and $TopCT_{i,k,l}$ is the binary variable used to identify the component with the lowest boiling point each active column (i.e. $TopCT_{i,k,l} = 1$ if component i is the lightest component in column l of k), which is defined as the following:

$$1 - N_{i,k,l}^{in} + \sum_{i'<i} N_{i',k,l}^{in} \geq 1 - TopCT_{i,k,l}, \forall i, k, l \tag{4.30}$$

$$TopCT_{i,k,l} \leq N_{i,k,l}^{in}, \forall i, k, l \tag{4.31}$$

$$\sum_{i} TopCT_{i,k,l} \leq SA_{k,l}, \forall k, l \tag{4.32}$$

These three equations essentially state that for an active distillation column (Eq. (4.32)), if component i exists in the column (Eq. (4.31)) and all other components i' that have lower boiling points than that of i are not present (Eq. (4.30)) , then component i is the lightest component in that stage, i.e. $TopCT_{i,k,l} = 1$. In addition, condensers and reboilers remain non-operational in columns that are not in operation (i.e.$\sum_t \hat{y}_{k,l,t}^{RB} \leq SA_{k,l}, \forall k, l$ and$\sum_t \hat{y}_{k,l,t}^{CD} \leq SA_{k,l}, \forall k, l$).

The utility requirements of the condenser ($\Delta EC_{k,l}$) and the reboiler ($\Delta ER_{k,l}$) are related by the energy balance around each separation column, i.e.

$$\Delta ER_{k,l} = -\Delta EC_{k,l} + E_{k,l}^{fT} + E_{k,l}^{fB} - E_{k,l}^{fin}, \forall k, l \tag{4.33}$$

where $E_{k,l}^{fT}$ and $E_{k,l}^{fB}$ are the relative flow energy of the distillate and bottom product stream respectively defined as the sum of the product of the corresponding flowrate ($f_{i,k,l}^{T}/f_{i,k,l}^{B}$) and liquid specific heat capacity (C_{pli}) of component i multiplied by the operating temperature of the condenser ($T_{k,l}^{CD}$) and rebolier ($T_{k,l}^{RB}$):

$$E_{k,l}^{fT} = \sum_{i} f_{i,k,l}^{T} C_{pli} T_{k,l}^{CD}, \forall k, l \tag{4.34}*$$

$$E_{k,l}^{fB} = \sum_{i} f_{i,k,l}^{B} C_{pli} T_{k,l}^{RB}, \forall k, l \tag{4.35}*$$

where the operational temperatures of the condenser and rebolier can be written in terms of the temperature point variables (i.e. $T_{k,l}^{CD} = \sum_t t \cdot \hat{y}_{k,l,t}^{CD}, \forall k, l$ and $T_{k,l}^{RB} = \sum_t t \cdot \hat{y}_{k,l,t}^{RB}, \forall k, l$). These two equations are also bi-linear. On the other hand, the relative flow energy of the inlet stream to each distillation column ($E_{k,l}^{fin}$) is calculated by the following equation:

$$E_{k,l}^{fin} = \sum_{i} f_{i,k,l}^{in} C_{pli} T_{k,l}^{fin}, \forall k, l \tag{4.36}*$$

where $f_{i,k,l}^{in}$ is the inlet flowrate of component i to separation stage l of k and $T_{k,l}^{fin}$ is the temperature of that stream.

In the case where $l = 1$, the starting temperature of the separation inlet stream ($f_{i,k,l=1}^{in}$) is dependent on the temperature of the reactor outlet which may need a heat exchanger to adjust the temperature so that all components in the stream are in the liquid phase. Also as mentioned above, the temperature after the adjustment is equal to the boiling point of the lightest component in the stream, which, in fact, is also equal to the operating temperature of the condenser of that column (i.e. $T_{k,l=1}^{fin} = T_{k,l=1}^{CD}$). On the other hand, in the case where $l > 1$, each separation inlet stream ($f_{i,k,l>1}^{in}$) comes from either from the top (temperature is equal to the lightest component's BP) or bottom product (temperature is equal to the heavy key's BP) of previous separation stages. As a result, the flow temperature ($T_{k,l>1}^{fin}$) is always equal to the boiling point of the lightest component in that stream and requires no temperature adjustment (i.e. the relation: $T_{k,l}^{fin} = T_{k,l}^{CD}$ always holds for any $l > 1$). The cooling duty ($\Delta EC_{k,l}$) for a total condenser is calculated based on Douglas (1988) and is simply given as follows:

$$\Delta EC_{k,l} = -\sum_{i}(Rf + 1) f_{i,k,l}^{T} hv_i \,, \forall k, l \tag{4.37}$$

where Rf is the reflux ratio, and hv_i is the heat of evaporation of component i.

Similarly to the reaction stages, substances in condensers act as heat sources whereas those in reboilers act as heat sinks at their respective temperatures, so that they can be treated as hot and cold streams respectively. As a result, the expression for the heat surplus of each condenser ($SH_{k,l,t}^{CD}$) and reboiler ($SH_{k,l,t}^{RB}$) in terms of the actual heat surplus ($SHA_{k,l,t}^{CD}$/ $SHA_{k,l,t}^{RB}$) at each temperature t are written as:

$$SH_{k,l,t}^{RB} = -SHA_{k,l,t}^{RB} = -\Delta ER_{k,l}\hat{y}_{k,l,t}^{RB}, \forall k, l, t \tag{4.38}*$$

$$SH_{k,l,t}^{CD} = SHA_{k,l,t+T_{gradient}}^{CD} = -\Delta EC_{k,l}\hat{y}_{k,l,t+T_{gradient}}^{CD}, \forall k, l, t \tag{4.39}*$$

4.2.5. Master Stage Streams (Mstream)

There are total four types of streams in each master stage k (*Mstream*) that may require temperature adjustment, namely the addition stream ($A_{i,k}$), reactor outlet bypass to the next master stage (RON-$Q_{i,k}^{next}$), to the separation stage (ROS-$Q_{i,k}^{sepin}$) and to the direct removal (ROR-$Q_{i,k}^{R}$). As mentioned in section 4.2.1, the inlet (Ts_k^{stream}) and outlet temperatures (Te_k^{stream}) of the heat exchanger of each master stream are defined either by the user or the model, i.e.

$$Ts_k^A = T^{\emptyset} N_k^A, \forall k \tag{4.40}$$

$$Te_k^A = T_k^{Re} N_k^A, \forall k \tag{4.41}*$$

$$Ts_k^{Qnext} = T_k^{Re}\, N_k^{Qnext}, \forall k \tag{4.42}*$$

$$Te_k^{Qnext} = T_{k+1}^{Re}\, N_k^{Qnext}, \forall k \tag{4.43}*$$

$$Ts_k^{Qsepin} = T_k^{Re}\, N_k^{Qsepin}, \forall k \tag{4.44}*$$

$$Te_k^{Qsepin} = T_{k,l=1}^{fin}\, N_k^{Qsepin}, \forall k \quad (4.45)^*$$

$$Ts_k^{QR} = T_k^{Re}\, N_k^{QR}, \forall k \quad (4.46)^*$$

$$Te_k^{QR} = T^{\emptyset}\, N_k^{QR}, \forall k \quad (4.47)^*$$

where $N_k^{Mstream}$ is the binary variable representing the existence of each master stage stream (i.e. $N_k^{Mstream} = 1$ if there is flow in that stream).

The respective binary expressions for the starting and ending temperature points of each master stage stream, $\widehat{ys}_{k,t}^{Mstream}$ and $\widehat{ye}_{k,t}^{Mstream}$, are thus given as follows:

$$Ts_k^{Mstream} = \sum_t t \cdot \widehat{ys}_{k,t}^{Mstream}, \forall k \quad (4.48)$$

$$Te_k^{Mstream} = \sum_t t \cdot \widehat{ye}_{k,t}^{Mstream}, \forall k \quad (4.49)$$

with both of the binary variables constrained so that only one temperature point can be selected for active streams (i.e. $\sum_t \widehat{ys}_{k,t}^{Mstream} \leq N_k^{Mstream}, \forall k$ and $\sum_t \widehat{ye}_{k,t}^{Mstream} \leq N_k^{Mstream}, \forall k$). Among those four types of streams, addition streams are assumed to be cold as all reactions in this work are assumed to have a higher operating temperature than the ambient temperature. ROR streams are set to be hot as direct removal streams are always subject to cooling. On the other hand, RON and ROS streams are either hot streams if the exit temperature of the heat exchanger is lower than that of the inlet, or cold streams otherwise, i.e.

$$Ts_k^{Qnext} - Te_k^{Qnext} \leq M \cdot HOT_k^{Qnext}, \forall k \quad (4.50)$$

$$Te_k^{Qnext} - Ts_k^{Qnext} \leq M\left(1 - HOT_k^{Qnext}\right), \forall k \quad (4.51)$$

$$HOT_k^{Qnext} \leq N_k^{Qnext}, \forall k \quad (4.52)$$

$$Ts_k^{Qsepin} - Te_k^{Qsepin} \leq M \cdot HOT_k^{Qsepin}, \forall k \quad (4.53)$$

$$Te_k^{Qsepin} - Ts_k^{Qsepin} \leq M\left(1 - HOT_k^{Qsepin}\right), \forall k \quad (4.54)$$

$$HOT_k^{Qsepin} \leq N_k^{Qsepin}, \forall k \quad (4.55)$$

where $HOT_k^{Mstream} = 1$ if that stream is a hot stream, 0 otherwise.

Based on the above information, the temperature interval where each stream is located ($TI_{k,t}^{Mstream}$), can be determined (i.e. if the stream' starting temperature is T, then $\widehat{ys}_{k,T}^{Mstream} = 1$ and exits the heat exchanger at temperature $T + x$ (i.e. $\widehat{ye}_{k,T+x}^{Mstream} = 1$), $TI_{k,t}^{Mstream} = 1$ for all $T \leq t \leq T + x$). For user-defined cold streams and cold streams determined by the optimisations (i.e. $HOT_k^{Mstream} = 0$):

$$TI_{k,t}^{Mstream} = TI_{k,t-1}^{Mstream} + \widehat{ys}_{k,t}^{Mstream} - \widehat{ye}_{k,t-1}^{Mstream}, \forall k, t \quad (4.56)$$

On the other hand, $TI_{k,t}^{Mstream}$ is defined as follows for hot streams:

$$TI_{k,t}^{Mstream} = TI_{k,t-1}^{Mstream} + \widehat{ye}_{k,t}^{Mstream} - \widehat{ys}_{k,t-1}^{Mstream}, \forall k, t \quad (4.57)$$

The flow heat capacity of each master stream ($FCp_{k,t}^{Mstream}$) is simply:

$$FCp_{k,t}^{Mstream} = \sum_{i} Mstream_{i,k} Cp_{i,t} \; + \sum_{i} Mstream_{i,k} HoV_{i,t}, \forall k, t \quad (4.58)$$

The actual heat surplus of each master stage stream ($SHA_{k,t}^{Mstream}$) is given as:

$$SHA_{k,t}^{Mstream} = FCp_{k,t}^{Mstream} TI_{k,t}^{Mstream}, \forall k, t \quad (4.59)^*$$

4.2.6. Separation Stage Streams (Dstream)

As mentioned in Chapter III, the product streams of each distillation column can have three end destinations: (i) the inlet streams to subsequent separation stages ($f_{i,k,l,l'}^{Ts}/f_{i,k,l,l'}^{Bs}$), which as mentioned above, do not require any heating or cooling; (ii) the direct removal streams ($f_{i,k,l}^{TR}/f_{i,k,l}^{BR}$), which are cooled before storage (direct removal streams); (iii) the streams that enter the next reaction stage ($f_{i,k,l}^{Tout}/f_{i,k,l}^{Bout}$), which may require temperature adjustment to match the temperature of the reactor (top/bottom out streams). The following sections present the heat surplus formulation of the distillation products in detail.

The inlet ($Ts_{k,l}^{Dstream}$) and outlet temperatures ($Te_{k,l}^{Dstream}$) of the heat exchanger of each distillation split stream are given as follows:

$$Ts_{k,l}^{TR} = T_{k,l}^{CD} N_{k,l}^{TR}, \forall k, l \quad (4.60)^*$$

$$Te_{k,l}^{TR} = T^{\emptyset} N_{k,l}^{TR}, \forall k, l \quad (4.61)$$

$$Ts_{k,l}^{BR} = T_{k,l}^{RB} \; N_{k,l}^{BR}, \forall k, l \quad (4.62)^*$$

$$Te_{k,l}^{BR} = T^{\emptyset} N_{k,l}^{BR}, \forall k, l \quad (4.63)$$

$$Ts_{k,l}^{Tout} = T_{k,l}^{CD} N_{k,l}^{Tout}, \forall k, l \quad (4.64)^*$$

$$Te_{k,l}^{Tout} = T_{k+1}^{Re} N_{k,l}^{Tout}, \forall k, l \quad (4.65)^*$$

$$Ts_{k,l}^{Bout} = T_{k,l}^{RB} N_{k,l}^{Bout}, \forall k, l \quad (4.66)^*$$

$$Te_{k,l}^{Bout} = T_{k+1}^{Re} N_{k,l}^{Bout}, \forall k, l \quad (4.67)^*$$

where $N_{k,l}^{Dstream}$ are binary variables represents the existence of each distillation stage stream.

The respective binary expressions for the starting and ending temperature points of each distillation stream, $\widehat{ys}_{k,l,t}^{Dstream}$ and $\widehat{ye}_{k,l,t}^{Dstream}$, are thus given as follows:

$$Ts_{k,l}^{Dstream} = \sum_{t} t \cdot \widehat{ys}_{k,l,t}^{Dstream}, \forall k, l \quad (4.68)$$

$$Te_{k,l}^{Dstream} = \sum_{t} t \cdot \widehat{ye}_{k,l,t}^{Dstream}, \forall k, l \quad (4.69)$$

$$\sum_t \widehat{ys}_{k,l,t}^{Dstream} \leq N_{k,l}^{Dstream}, \forall k,l \tag{4.70}$$

$$\sum_t \widehat{ye}_{k,l,t}^{Dstream} \leq N_{k,l}^{Dstream}, \forall k,l \tag{4.71}$$

In addition, the stream properties of $Tout$ and $Bout$ streams are determined by the following constraints:

$$Ts_{k,l}^{Tout} - Te_{k,l}^{Tout} \leq M \cdot HOT_{k,l}^{Tout}, \forall k,l \tag{4.72}$$

$$Te_{k,l}^{Tout} - Ts_{k,l}^{Tout} \leq M\left(1 - HOT_{k,l}^{Tout}\right), \forall k,l \tag{4.73}$$

$$Ts_{k,l}^{Bout} - Te_{k,l}^{Bout} \leq M \cdot HOT_{k,l}^{Bout}, \forall k,l \tag{4.74}$$

$$Te_{k,l}^{Bout} - Ts_{k,l}^{Bout} \leq M\left(1 - HOT_{k,l}^{Bout}\right), \forall k,l \tag{4.75}$$

The product streams are defined by the user to be hot streams.

Similar to Equations (4.56) and (4.57), the temperature intervals of the heat exchangers of each distillation stream are given as:

$$TI_{k,l,t}^{Dstream} = TI_{k,l,t-1}^{Dstream} + \widehat{ys}_{k,l,t}^{Dstream} - \widehat{ye}_{k,l,t-1}^{Dstream}, \forall k,l,t, cold\ streams \tag{4.76}$$

$$TI_{k,l,t}^{Dstream} = TI_{k,l,t-1}^{Dstream} + \widehat{ye}_{k,l,t}^{Dstream} - \widehat{ys}_{k,l,t-1}^{Dstream}, \forall k,l,t, hot\ streams \tag{4.77}$$

The flow heat capacity of each distillation stream ($FCp_{k,l,t}^{Dstream}$):

$$FCp_{k,l,t}^{Dstream} = \sum_i Dstream_{i,k} Cp_{i,t} \ + \sum_i Dstream_{i,k} HoV_{i,t}, \forall k,l,t \tag{4.78}$$

and the actual heat surplus ($SHA_{k,l,t}^{Dstream}$):

$$SHA_{k,l,t}^{Dstream} = FCp_{k,l,t}^{Dstream} TI_{k,l,t}^{Dstream}, \forall k,l,t \tag{4.79}*$$

Overall, the objective function (4.5) and together with the constraints from Eq. (4.4)-(4.79) form a mixed integer linear programming (MILP) model which can be solved globally using the optimisation solver CPLEX in GAMS (version 24.2.3) to obtain the optimal process configuration and the grand composite curve of the integrated process. To illustrate the proposed approach, a case study is solved in the following sections.

4.3. Case study

In this section, the same case used in 3.3.1 of Chapter III is used for illustration purpose.

4.3.1. Case Outline

D-Limonene is a renewable chemical with numerous applications in the cosmetics, medicine, perfume and food industries (Ciriminna et al. 2014). It is currently being produced as a side product from the citrus juice industry (Lange 2015). Recently, extensive research has been carried out to study the possibility for limonene to be used as a platform chemical for the synthesis of higher value chemicals such as p-cymene (Martín-Luengo et al. 2008), perillic acid (Ferrara et al. 2013) and even polymers

(Byrne et al. 2004) thanks to its competitive price, biodegradability and low toxicity. With increasing effort being committed to limonene related reaction pathways, it would be beneficial for researchers if they can identify the optimal pathways from a set of potential candidate reactions before further experiments are conducted. Therefore, in this work, the proposed limonene reaction network given in chapter III (Figure 4.3) is a suitable candidate for the case study. The effects of additional constraints

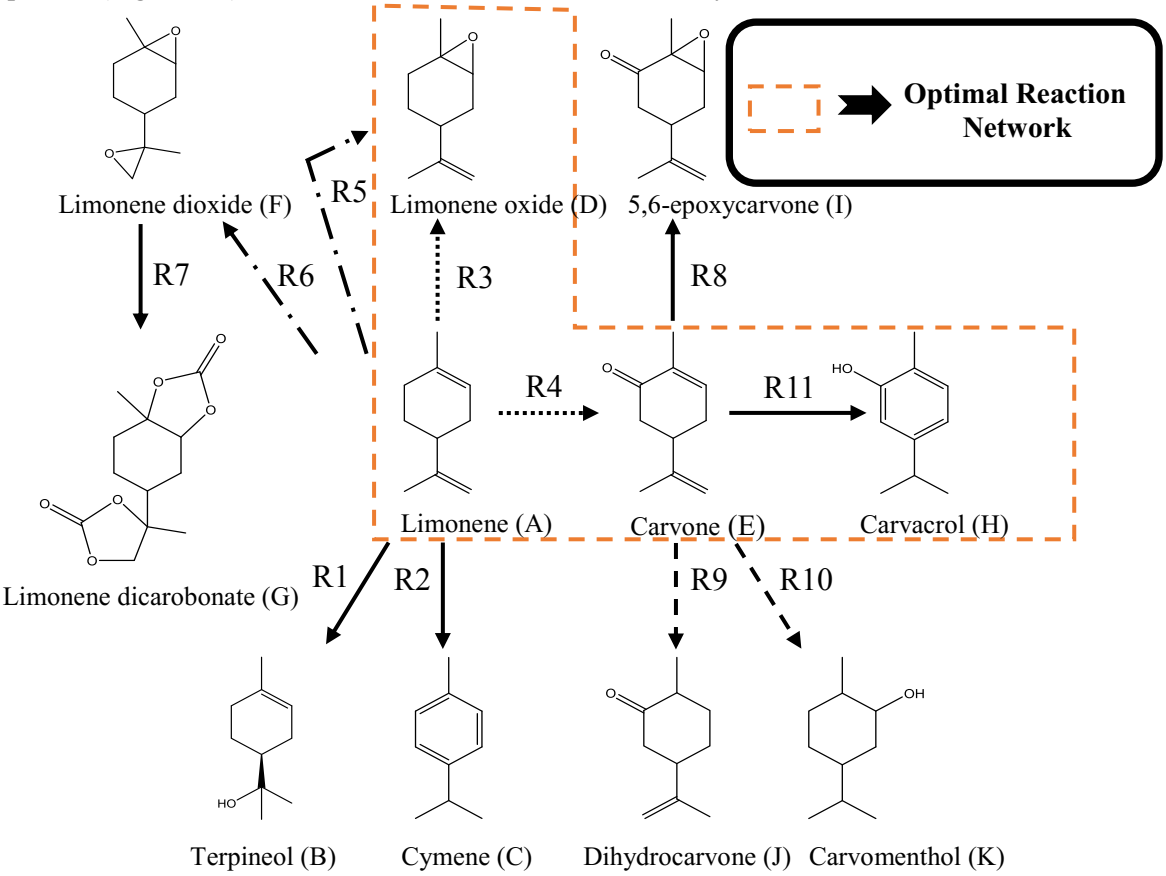


Figure 4.3: Proposed reaction network-reaction. Reactions R3 and R4, R5 and R6, R9 and R10 are all competitive reactions. Optimal reaction network is outlined with the dashed circle.

considering heat integration are thus demonstrated.

The same operational parameters from Chapter III are used in this problem with the addition of some energy parameters. Firstly, the costs of steam and cooling water are taken to be \$11.2 per 10^6 kJ and \$3.91 per 10^6 kJ respectively (Mascia et al. 2007). The reflux ratio of each distillation column, which is given in Eq. (4.37), is initially assumed to be 1.3 and is subject to further sensitivity analysis to determine its impact on the optimal solution.

In addition, thermodynamic properties of the components and reactions shown in Figure 4.3 are estimated via various group contribution methods. The additive value of each group that contributes to specific gas heat capacity of each component is estimated via one of the follow equations depending on the group structure (Bureš et al. 1981):

$$\Delta Cp_i = a_i + b_i T + c_i T^2 \tag{4.80}$$

$$\Delta Cp_i = a_i + b_i T + \frac{c_i}{T^2} \tag{4.81}$$

$$\Delta Cp_i = a_i + b_i \left(\frac{c_i}{T}\right)^2 \frac{e^{-\frac{c_i}{T}}}{\left(1 - e^{-\frac{c}{T}}\right)^2} \tag{4.82}$$

where ΔCp_i is the additive contribution of each chemical group i towards the heat capacity, a_i, b_i and c_i are constants related to each group and T is the temperature in K.

The liquid phase heat capacity of each component given in Figure 4.3, on the other hand, is calculated using Group Additivity method developed by Zábranský and Růžička (2004), which is given as follows:

$$\frac{Cp_i}{R} = \sum_{j=1}^{k} n_j \Delta Cp_j \tag{4.83}$$

$$\Delta Cp_i = a_i + b_i \frac{T}{100} + d_i \left(\frac{T}{100}\right)^2 \tag{4.84}$$

In equation (4.83), R is the gas constant, n_j is the number of additivity units of type j, ΔCp_j is a dimensionless value of the additivity unit of type j, and k is the total number of additivity units in a molecule. In equation (4.84), T is temperature in K and a_i , b_i , d_i are adjustable parameters.

Other thermodynamic properties of each component such as the standard enthalpy of formation, which is used to calculate the standard enthalpy of each reaction, and enthalpy of evaporation are calculated using the Group-Contribution Approach proposed by Verevkin et al. (2013), which has satisfactory level of accuracy ($\leq 3\%$ for most of the tested component, with a few at about 10%). Although there are other estimation methods that are applicable (e.g. Joback and Reid (1987), Marrero and Gani (2001)) the investigation of the estimation differences between those methods is not within the scope of this thesis. The calculated properties are summarised in Table 4.1 and Table 4.2. Furthermore, the temperature is initially discretised into intervals of 10 °C to save computational effort (e.g. the reaction temperature of R11 is at 65 °C, which lies in the 7th interval), and intervals of 5℃ and 2 ℃ are also analysed for comparison in the sensitivity analysis.

Table 4.1: Thermodynamic properties of reactions involved in the case study

Reaction No. (j)	Conversion (mol%)	Selectivity (mol%)	Reaction Temperature TR_j (°C)	Estimated standard heat of reaction, $\Delta H_j^{\emptyset}$(kJ/mol)
R1	93	-	25	-34.29
R2	100	-	170	3.54
R3	100	59.5	75	-230.77
R4	100	40.5	75	-186.5
R5	98	74.5	40	-230.77
R6	98	25.5	40	-455.77
R7	100	-	140	-252.41
R8	32	-	50	-170.3
R9	90	62.6	100	-49.23
R10	90	37.4	100	-232.04
R11	90	-	75	-44.91

Table 4.2: Thermodynamic properties of components involved in the case study

Component (i)	Liquid phase specific heat capacity, Cp_{li} $(Jmol^{-1}K^{-1})$	Gas phase specific heat capacity, Cp_{gi} $(Jmol^{-1}K^{-1})$	Latent heat of Evaporation, HoV_i (kJ/mol)
D-Limonene (A)	300.2	88.5	43.2
α-Terpineol (B)	412.5	96.4	67.3
P-cymene (C)	307.4	74.5	49.0
Limonene oxide (D)	229.6	92.0	38.5
Carvone (E)	320.2	85.3	55.5
Limonene dioxide (F)	228.5	97.1	37.3
Limonene dicarbonate (G)	370.3	125.9	68.2
Carvacrol (H)	369.2	82.0	75.0
5,6-epoxycarvone (I)	286.3	90.6	49.9
Dihydrocarvone (J)	285.1	87.9	49.7
Carvomenthol (K)	425.3	96.6	68.0

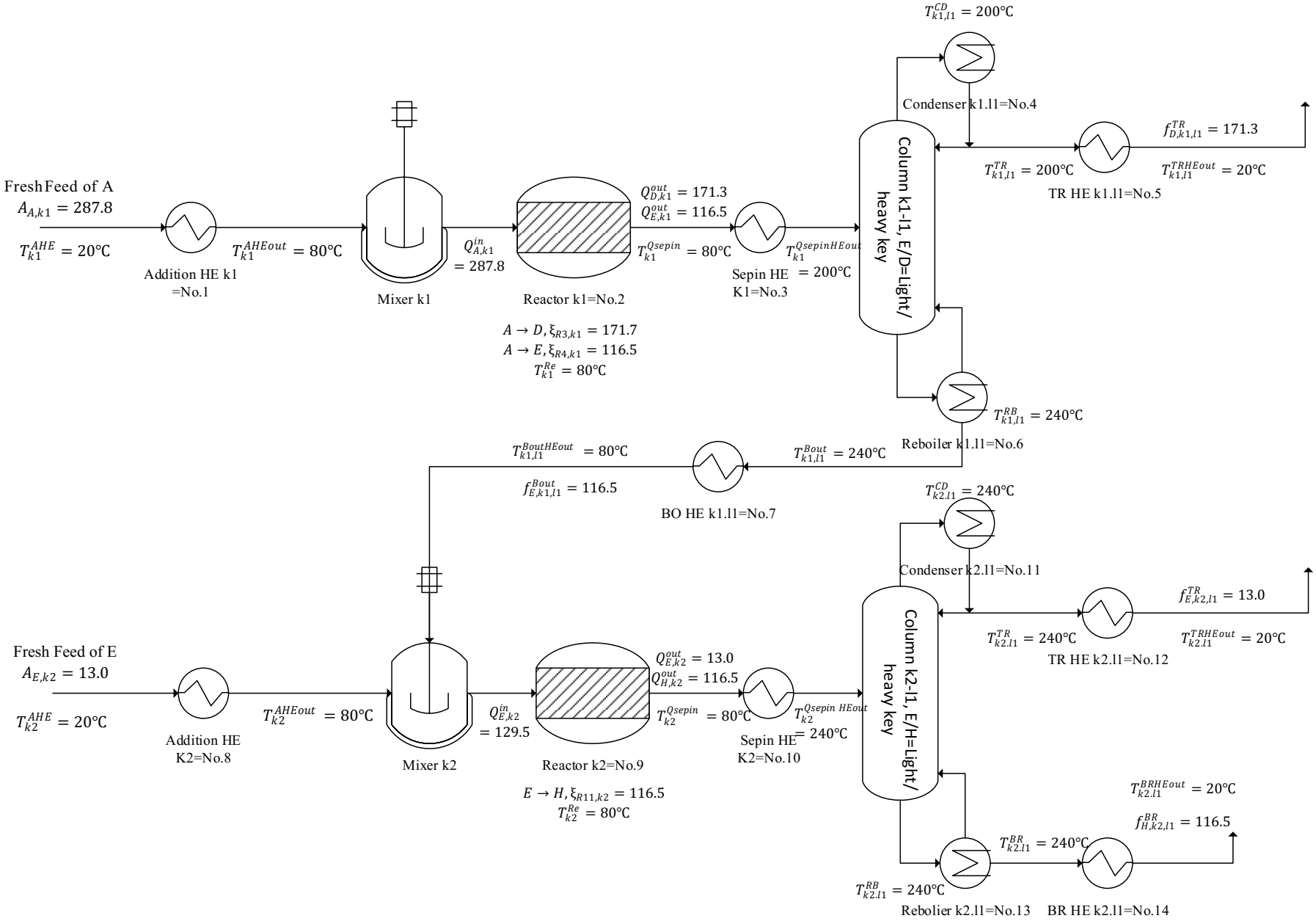


Figure 4.4: Optimal process flowsheet of the case study, all flowrate are in the unit of mol/min (temperature intervals of 10 °C)

4.3.2. Initial Solutions

The optimal reaction network identified by the model remains the same as that of Chapter III and is shown as the network within the dashed envelope of Figure 3.4 with carvacrol (H) being selected as the optimal product. In general, reaction No.4 is chosen to be the first step over the others that originate from limonene (A) to produce the intermediate product carvone (E). Meanwhile, reaction No.3 also takes place as a competitive reaction to produce the by-product limonene oxide (D). Pure carvone is separated from limonene oxide to undergo further isomerisation to form the final product carvacrol. Both carvacrol and limonene oxide are sold at their market prices. The corresponding optimal process flowsheet with heat integration implemented is shown in Figure 4.4. There are 14 utility 'streams' in total shown in Table 4.3 including the enthalpy change units (reactor, reboiler and condensers). As can be seen that the fresh feed of limonene (A) is first heated from 20 °C to 80 °C to match the operational temperature of reactor k1, where R3 and R4 take place. The reactor outlet, which is composed of a mixture of limonene oxide (D) and carvone (E), is initially at 80 °C and then heated to 200 °C before entering the sub-distillation column. The reboiler and condenser of the first column (k1-l1) operate at 240 and 200 °C respectively. The distillate containing pure limonene oxide leaves the condenser at 200 °C and is cooled to 20 °C before storage. On the other hand, the bottom product (carvone), exiting the reboiler at 240 °C, is cooled to 80 °C to match the operating temperature of the 2nd reactor (80 °C) and is then mixed with the fresh feed of carvone (also preheated to 80°C) before entering the 2nd reactor. The product of the 2nd reactor, which consists of both carvone and carvacrol (H), leaves the reactor at a temperature of 80 °C and is preheated to 240 °C prior to separation. One thing to note here is that the

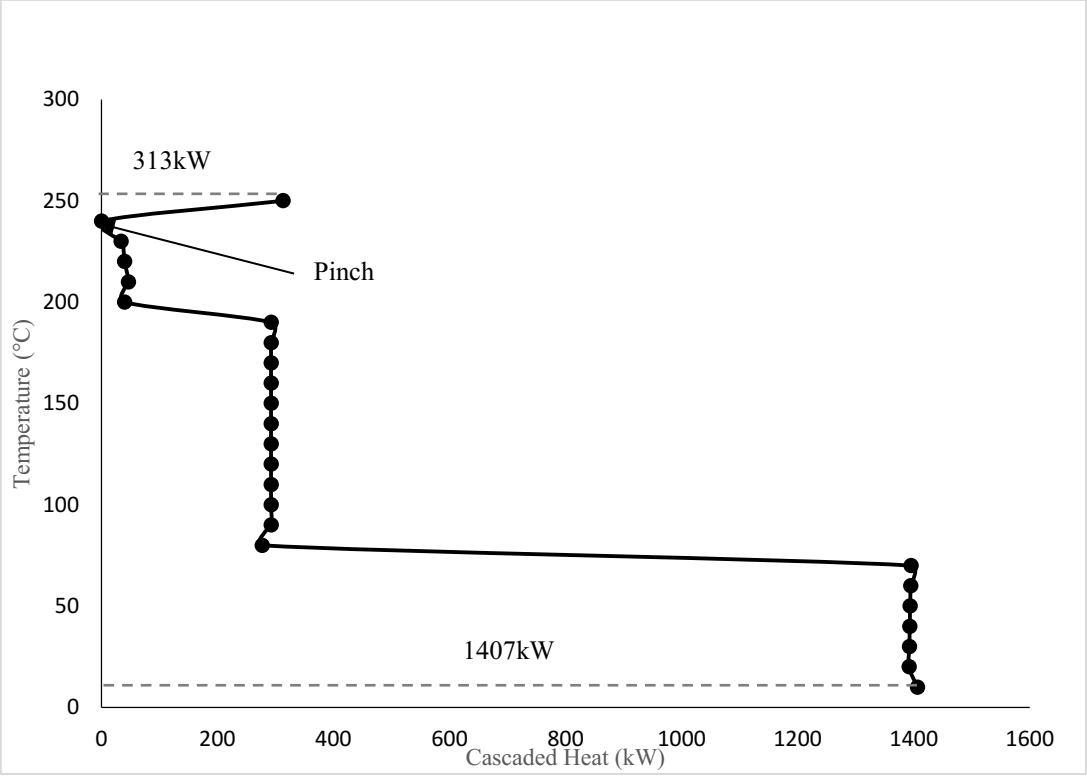


Figure 4.5: The grand composite curve ($\Delta \boldsymbol{T} = \mathbf{10}\ \mathbf{°C}$)

boiling points of component E and H are very close (231 and 238 °C respectively) and are both located in the 24th interval in the case of temperature interval of 10 °C. This indicates that the reboiler and condenser of the 2nd column both operate at 240 °C, which is technically not possible for any distillation columns. Although increasing the number of temperature intervals can resolve this issue by assigning the boiling points of E and H to different temperature intervals, the set of solutions still exposes the model's lack of capability in dealing with azeotropic distillation. Even though this topic is not the focus of this thesis, future work should be conducted. Nevertheless, this particular scenario is acceptable to serve as an estimated solution to the complex model. Finally, the top and bottom products of column k2-l1 containing pure carvone and carvacrol respectively are cooled to 20 °C before storage.

Table 4.3: Inlet/outlet temperatures of each heat exchanger or enthalpy change unit

Stream No.	Utility stream characterisation	Inlet temperature (°C)	Outlet temperature (°C)	Flowrate (mol/min)
1	Stream	20	80	287.8
2	Reactor k1	80	80	-
3	Stream	80	200	287.8
4	Condenser k1.l1	200	200	-
5	Stream	200	20	171.3
6	Reboiler k1.l1	240	240	-
7	Stream	240	80	116.5
8	Stream	20	80	13.0
9	Reactor k2	80	80	-
10	Stream	80	240	129.5
11	Condenser k2.l1	240	240	-
12	Stream	240	20	13
13	Reboiler k2.l1	240	240	-
14	Stream	240	20	116.5

The optimal economic potential is $20.5 million per annum with an annual utility cost of $0.275 million. The grand composite curve of the whole process is given in Figure 4.5. The optimal solution indicates that the minimum heating and cooling required for the system is 313 kW and 1407 kW respectively with a pinch point at 240 °C. The total utility saving compared with the case where there is no heat integration (i.e. when all the energy required is provided by auxiliary utilities) is around 40 %. The initial solutions are further analysed by performing sensitivity analysis in section 4.4. Detailed model statistics and solution summary are given in Table 4.4.

Table 4.4: Model statistics and solution summary of the case study

Single Variables	26,684	**MIP Iterations**	261,324
Discrete Variables	7,965	**MIP Nodes Explored**	1,745
Relative Tolerance	0.1	**CPU time (s)**	33.9
Best Possible	20.5×10^6	**Final Solve**	20.5×10^6
Absolute Gap	0	**Relative Gap**	0

4.3.3. Column Pressure Modification

However, it is noticed that the energy integration potential is relatively low due to a high pinch point caused by the fact that there is a reboiler operating at 240 °C which consumes only auxiliary heating utilities. On the other hand, the massive leap in cascaded heat from 80 °C to 70 °C indicates that two highly exothermic reactors both operate at 80 °C and release far more energy than any other streams. To match the reboiler's heating duty with the cooling duty required by the reactors, the pressures of the distillation columns are reduced to lower the operating temperatures of reboilers and condensers by 170 °C. The resulting grand composite curve is shown in Figure 4.6.

As can be seen, the new pinch point is at 80 °C, and a significant amount of heat produced by reactors is utilised by the low-pressure reboilers at 70 °C. As a result, the minimum cooling duty required of the process is reduced from 1407 to 1143 kW (an 18.7% drop) whereas the heating duty drops from 313 to 21.3 kW (a 93% improvement). In summary, the economic potential of the low-pressure process is $20.8 million per annum with an annual utility cost of $0.144 million. The total energy saving is improved from 40% to approximately 69%.

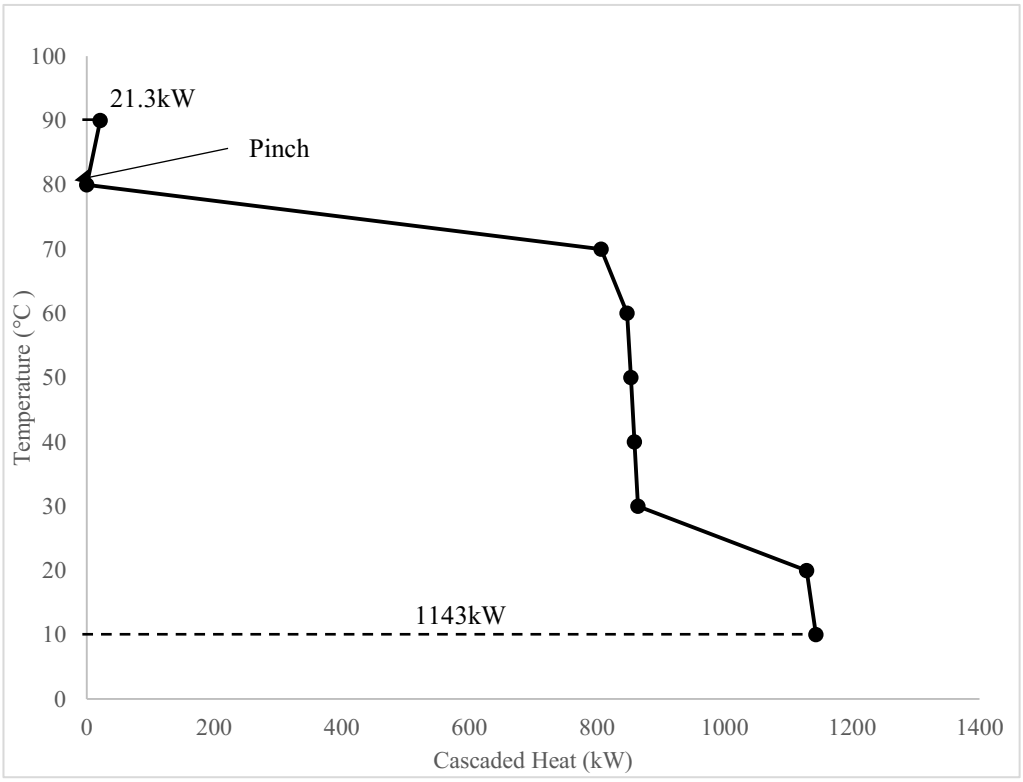


Figure 4.6: The grand composite curve ($\Delta \boldsymbol{T = 10}$ **°C**) with column depressurisation

4.4. Sensitivity Analysis

Sensitivity analysis of the model on several parameters such as the interval size of the discretisation, reflux ratio and reaction temperatures are performed in this section to evaluate the model's reliability towards uncertainties.

4.4.1. Interval Size

In the above case study, the original problem with temperature interval size of 1 °C takes hours to solve, whereas by increasing the size of each temperature interval from 1 °C to 10 °C, the solution time is reduced from hours to minutes (shown in Table 4.5). However, it may lower the resolution of the results and can be an issue for distillation columns with close reboiler and condenser operating temperatures. One can also argue that a more accurate solution can be obtained if the interval is small enough. Therefore, sensitivity analysis is performed on the interval size to evaluate the validity of the initial interval size. The interval size is firstly set to 5 °C before being further reduced to 2 °C, and the grand composite curves for each of the three cases are shown in Figure 4.7. Solutions of key variables, such as the pinch point, the minimum cooling and heating requirement for the system ($Q_{min}^{cooling}$ and $Q_{min}^{heating}$) are summarised in Table 4.6. It is shown that the pinch points shift from 240 to 230 °C for both of the later cases. The two temperature points of the $\Delta T = 10℃$ case where there is a significant change in cascaded heat (200 °C and 70 °C) shift slightly downwards. These changes can be interpreted as follows: with a reduced interval size, more points are evaluated by the solver thus providing a more precise value for the cascaded heat of each point. On the other hand, due to the following facts in this case study: (i) the whole heat integration network is below the pinch; (ii) the largest heat surplus and deficit of the process integration network are caused by the highly exothermic reactions and high temperature reboilers; (iii) the reactors and reboilers operate at different temperatures that are not affected by the resolution of the temperature interval, changing the size of the temperature interval barely affects both

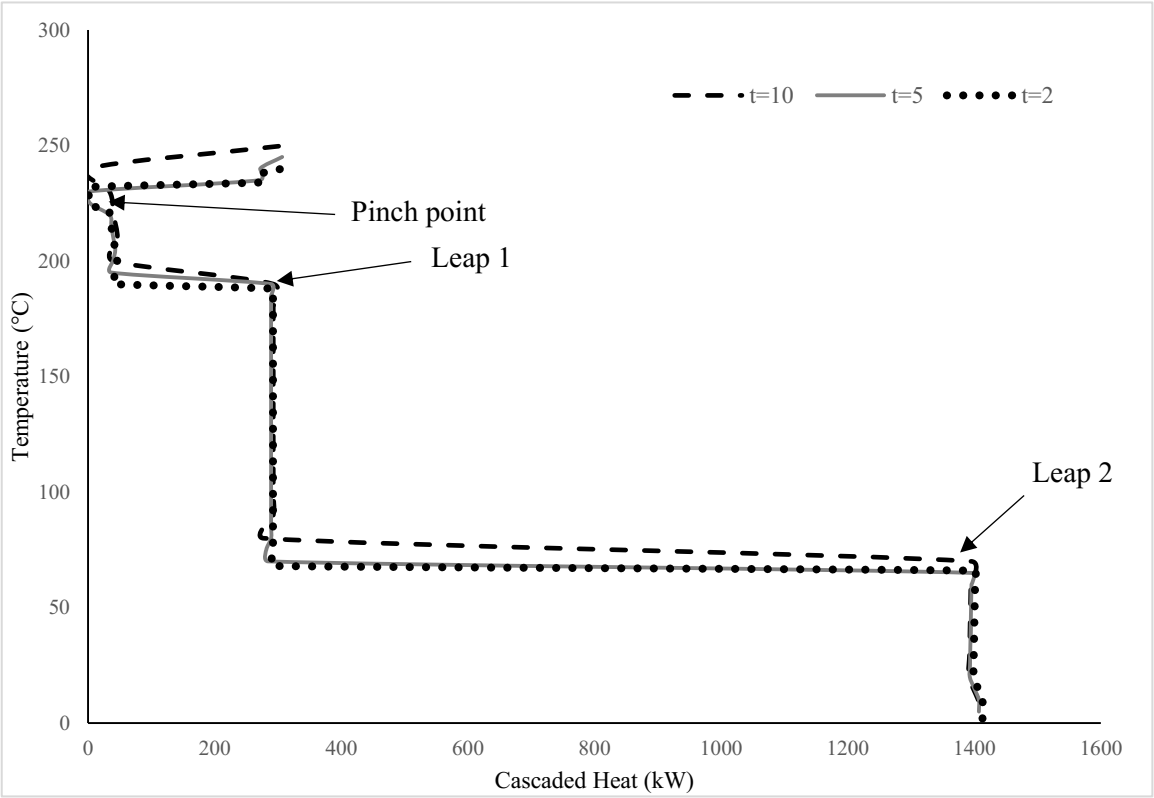


Figure 4.7: Grand composite curve for the sensitivity analysis on the interval size with original $\boldsymbol{TR_i}$ and $\boldsymbol{Rf = 1.3}$

the minimum cooling and heating duties. Also, the sensitivity analysis indicates that changing the interval size does not alter the optimal process flow diagram.

In summary, the solution obtained under the 10℃ interval assumption is validated, but more precise solutions can always be acquired by narrowing the interval size if CPU time is not a major concern for designers.

Table 4.5: Summary of problems sizes and solving time for cases with different interval sizes

Cases/Variables	Number of Variables	Number of Equations	CPU time (s)
$\Delta T = 10$	26,684	62,700	33.9
$\Delta T = 5$	50,044	115,814	105.9
$\Delta T = 2$	114,284	261,876	710.4

Table 4.6: Summary of key variables for cases with different interval sizes

Cases/Variables	$Q_{min}^{cooling}(kW)$	$Q_{min}^{heating}(kW)$	Pinch Point (℃)
$\Delta T = 10$	1406	313	240
$\Delta T = 5$	1407	306	230
$\Delta T = 2$	1413	307	230

4.4.2. Reflux Ratio

The reflux ratio of each column is usually calculated based on the minimum reflux ratio which can be obtained via complex correlations such as Underwood's equation (Underwood 1949). However, the nonlinearity of the equation will overcomplicate the problem exponentially. Therefore, instead of having reflux ratio as a decision variable, it is assumed in this paper that the reflux ratios for all columns

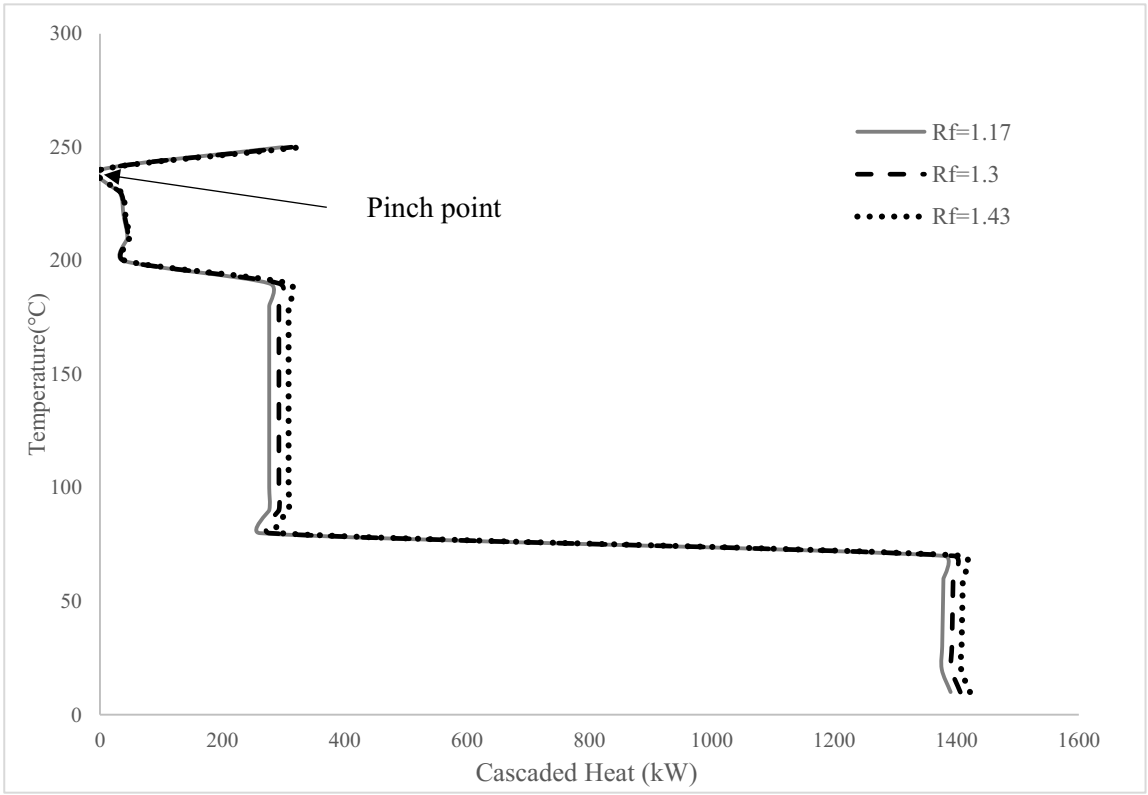


Figure 4.8: Grand composite curve for the sensitivity analysis on the reflux ratio (Rf) with original $\boldsymbol{TR_j}$ and $\boldsymbol{\Delta T = 10}$℃

are the same and the value is set to be 1.3 (Rf_0) initially. To examine the model's sensitivity to this assumption, the same method is performed on the same case but with 10% change in the reflux ratio (i.e. $Rf = 0.9Rf_0$ or $Rf = 1.1Rf_0$). The cascaded heat for the three scenarios is given in Figure 4.8. As the reflux ratio increases from 1.17 to 1.43 (10% each time), the overall grand composite curve shifts slightly towards right meaning a slightly higher cascaded heat at each temperature. Moreover, the minimum cooling and heating duties required are also increased (Table 4.7) thus leading to a rise in the overall utility cost of the process. The reason for this is that as reflux ratio increases, both the reboiler and condenser of each column will process more fluid increasing the utility requirements for each unit (Eq. (4.33) and (4.37)). Overall, changing the reflux ratio will affect the overall process economics, but the influence can be considered negligible and it does not alter the optimal flowsheet of the process. Therefore, it is safe to use reflux ratio as a parameter for the preliminary analysis.

Table 4.7: Summary of key variables for cases with different reflux ratios

Cases/Variables	$Q_{min}^{cooling}(kW)$	$Q_{min}^{heating}(kW)$	Pinch Point (℃)
$Rf = 1.17$	1390	297	240
$Rf = 1.3$	1406	313	240
$Rf = 1.43$	1422	329	240

4.4.3. Reaction Temperature

Another parameter that requires sensitivity analysis is the reaction temperature of each reaction (TR_j). Since the data for all reactions are gathered from the literature, most of which are at laboratory scale, the conditions of each reaction, if scaled up, may vary from the original literature value. As a result, it

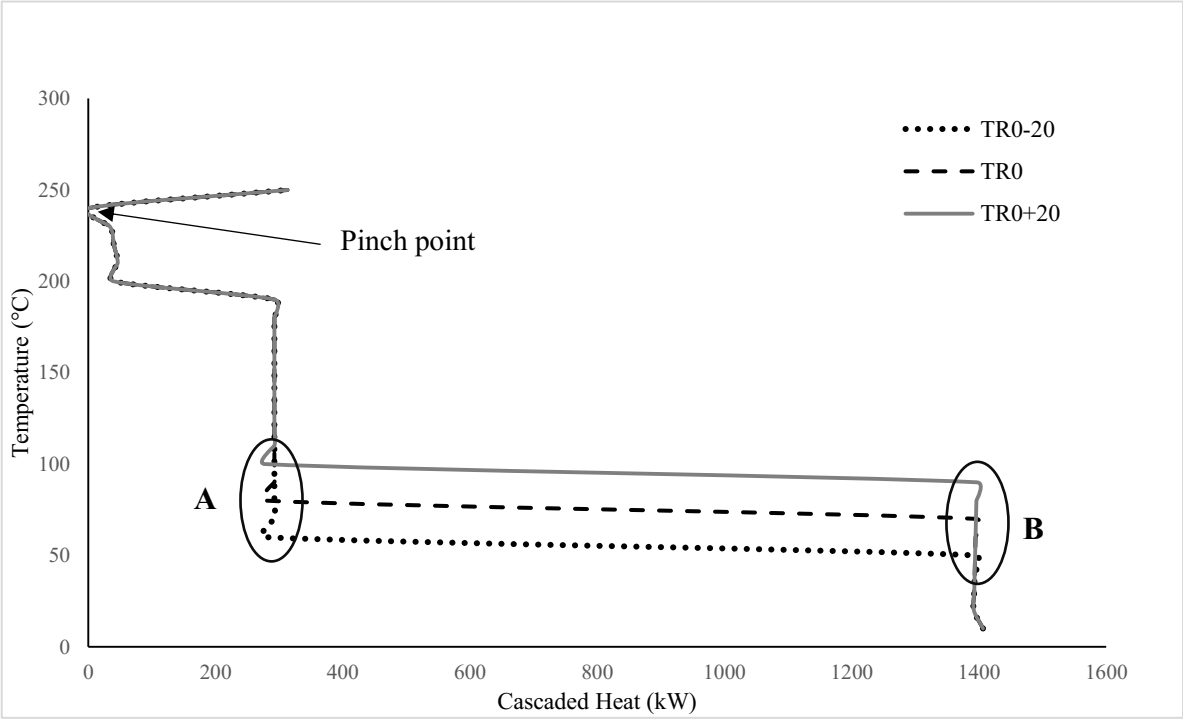


Figure 4.9: Grand composite curve for the sensitivity analysis reaction temperature ($\boldsymbol{TR_j}$) with $\boldsymbol{\Delta T = 10}$℃ and $\boldsymbol{Rf = 1.3}$

is necessary to identify the impact that reaction temperature has on the heat integration analysis and the optimal process flowsheet. As illustrated in Figure 4.9, the reaction temperature of each reaction (TR_{j0}) is firstly reduced by 20 °C (dotted line) and before increased to $TR_{j0} + 20$ °C (grey line). As can be seen in Table 4.8, the changes inflict no change on the process utility and the pinch temperature. This is because that even with different reaction temperatures, the active reactions of the optimal solution still operate well below the pinch temperature. However, as mentioned above, the temperatures at which the most significant change in heat surplus in the grand composite curve starts and ends (Point A and B) are dependent on the temperatures of the active reactions in the optimal solution. Therefore, it is shown in the diagram that changing the reaction temperatures from $TR_{j0} - 20$ to $TR_{j0} + 20$ °C also shifts both the turning point A and B by 40 °. It implies the scenario where an alternative cooling agent may have to be used to maintain the reactors at isothermal conditions if the reaction temperature is relatively low (i.e. cooling water does not satisfy the minimum approach temperature). Overall, for cases with a lower pinch point or reactions operating around the pinch point, it is possible to improve the optimal solution (i.e. reducing utility requirement) by varying the reaction temperatures. Nevertheless, in this case, neither the process economy nor the flowsheet is affected by the change in reaction temperatures.

Table 4.8: Summary of key variables for cases with different reaction temperatures

Reaction Temperature(℃)	$Q_{min}^{cooling}(kW)$	$Q_{min}^{heating}(kW)$	Pinch Point (℃)
$TR_j = TR_{j0} - 20$	1406	313	240
$TR_j = TR_{j0}$	1406	313	240
$TR_j = TR_{j0} + 20$	1406	313	240

4.5. Concluding Remarks

In this chapter, the methodology introduced in chapter III has been modified to allow the simultaneous process synthesis with automatic heat integration. This work aims to investigate the effect of heat integration on the energy consumption while screening for the most promising bio-based chemical and its production route from alternative candidate products and reaction pathways. The modified process superstructure is presented with the addition of heat exchangers. New sets of constraints are added to the original mixed integer linear programming (MILP) optimisation formulation to track and evaluate the energy change of each stream. The continuous temperature variable is discretised to maintain the model linearity. This avoids any local optimum induced by nonlinear equations and keeps the model's solution CPU time to a relatively low magnitude. The objective is to maximise the economic potential of the process subject to constraints with respects to the screening of reaction pathways, separation sequencing, heat integration and product selection. The proposed optimisation model was tested using a case study which involves the reaction pathways starting from the popular bio-derived D-limonene. Good quality global optimal solutions including the optimal flowsheet, process economic potential,

pinch temperature, minimum cooling and heating duties, are obtained from a single model. The initial results show that energy saving due to simultaneous heat integration was close to 40 % for this particular case which is further improved to approximately 69 % when distillation columns are operating at lower pressures. Several sensitivity analyses are performed to validate the assumptions used in the framework. The sensitivity analysis shows that the optimal process flow sheet is not significantly affected by varying the size of temperature interval, reflux ratio or reaction temperatures. On the other hand, variations in reflux ratio and temperature interval size do have an impact, although small, on the overall process utility requirement. The pinch temperature, however, only changes when reducing the size of temperature interval.

This model at this stage has already shown great potential for the process synthesis optimisation of newly developed reaction networks. It can provide valuable results even if provided with only limited information and help researchers to eliminate unnecessary risks at the earliest stage of process design. To build upon this work in the next chapter, a more sophisticated case study involving the production of polymers will be used to demonstrate the model's applicability further. Moreover, separation technologies such as precipitation and filtration will also be included. Most importantly, the solution CPU time is significantly reduced by applying feasible region partition method, which will be thoroughly discussed in the next chapter.

5. Chapter V: Enhancement of Computational Efficiency via the Partitioning of the Feasible Region

In the previous chapter, the proposed framework has shown its capability to solve process synthesis problems with simultaneous heat integration using an MILP model. The temperature discretisation allows more robust heat integration and provides a starting point for the generation of heat exchanger networks. However, the computational time (CPU time) grows exponentially when the number of temperature intervals increases. Although the computational time for $\Delta T = 2℃$ is still manageable (~30 minutes), it could easily explode to days or even becomes unsolvable if the interval size is further refined. Therefore, in this chapter, a new solution approach is used, which is based on the concept of the partition of the feasible region, to decompose the original problem into subproblems. The subproblems are submitted to a computational cluster and solved in parallel with inter-job communications to accelerate the solution speed.

The background of this chapter is first explained in section 5.1, followed by the problem statement in section 5.2. The methodology concept is presented in section 5.3 while a case study (section 5.4) is used to demonstrate the algorithm's applicability. Finally, this chapter concludes in section 5.5.

5.1. Background

The theory behind the discretisation of continuous variables to formulate MILP models is firstly presented in section 5.1.1 followed by the issues encountered when one tries to determine the proper interval size for different problems. Secondly, the primary solution strategy used to solve MILP problems (the best-bound branch and bound strategy) and its limitations towards complex MILP problems with a large number of integer variables, are explained in section 5.1.2. Finally, the concept of the feasible region of the discretised variables and the graphical representation of the optimal solutions are introduced in section 5.1.3.

5.1.1. Discretisation

The general expression of an MINLP formulation can be written as follows:

$$
\begin{aligned}
&Maximise\ z = c_1 x + c_2 E \\
&subject\ to: \\
&\quad f(x, Y) = 0; \\
&\quad g(x, Y) \leq 0; \\
&\quad w(t, Y) = 0; \\
&\quad h(x, t) - E = 0; \\
&x \in [0, X], t \in [0, T], E \in R; \\
&\quad Y \in \{0,1\}
\end{aligned}
\qquad \text{MINLP-P1}
$$

where c_1 and c_2 are parameters; x, t and E are continuous variables; Y is a binary decision variable; z is the objective function. Functions $f(x,Y)$, $g(x,Y)$ and $w(t,Y)$ are linear constraints while function $h(x,t)$ is nonlinear. The nonlinearity can be eliminated by modifying the continuous variable t into a discretised variable, which can only take integer values. As shown in Figure 5.1, each integer point n corresponds to an integer value and is assigned with a binary variable b_n. The temperature can only take the t value of one point if the binary variable of that point is equal to 1, i.e. $b_n = 1$ (e.g. if $b_{n=2} = 1$, then $t = 10$).

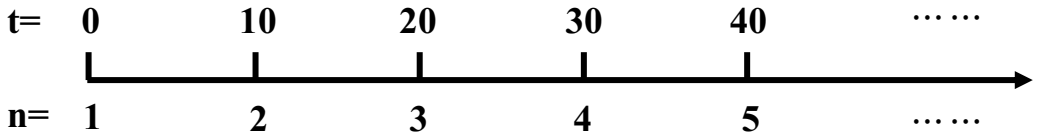


Figure 5.1: Variable discretisation using binary variables (a)

The equality constraint $h(x,t) - E = 0$ contains a nonlinear term in $h(x,t)$, which can now be decomposed into three parts: 1) a linear function - $d(b_n,t) = 0$ that relates the binary variable to the continuous variable; 2) a reformulated equality constraint - $h'(x,b_n) - E = 0$, which replaces the original nonlinear constraint by linking x with the new binary variable b_n; 3) a binary constraint - $p(b_n) = 0$. The MINLP formulation (P1) is thus reformulated to a MILP problem (P2) as follows:

$$
\begin{aligned}
&Maximise\ z = c_1 x + c_2 E \\
&subject\ to: \\
&\quad f(x,Y) = 0; \\
&\quad g(x,Y) \le 0; \\
&\quad w(t,Y) = 0; \\
&\quad h'(x,t) - E = 0; \\
&\quad d(b_n,t) = 0; \\
&\quad p(b_n) = 0; \\
&x \in [0,X], t \in [0,T,\Delta T = 10], E \in R; \\
&\quad Y, b_n \in \{0,1\}
\end{aligned}
\qquad \text{MILP-P2}
$$

The discretisation approach demonstrated in Figure 5.1 only allows integer values that are multiples of 10 to be used as the value of the variable. It means that if variable t in the optimal solution turns out to be equal to 10 (or its multiples), the discretisation interval is small enough for the given problem. However, if the optimality of the original nonlinear problem is achieved when $t = 5$, the interval size of the linear model has to be reduced by at least a half (as shown in Figure 5.2) to be able to obtain the

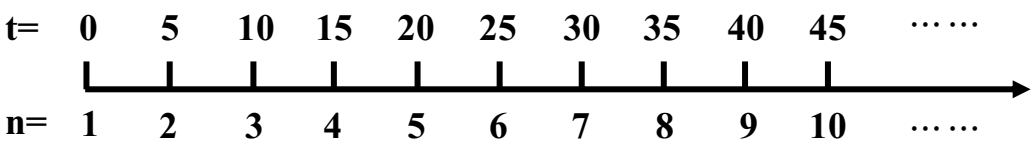


Figure 5.2: Variable discretisation using binary variables (b)

precise solution. However, such a decision-making process highly depends on the user's knowledge of the nonlinear problem and in reality, it is often the simplified linear model that serves to provide an estimated initial solution to the nonlinear model. Therefore, a smaller interval is generally preferred to be implemented in the linear model to improve the quality of the initial solution.

5.1.2. Branch and Bound

The solver used in this thesis is CPLEX, and it employs the traditional branch and bound algorithm for MILP problems. Before CPLEX starts the branch and bound process on the binary variables, it first solves the relaxed LP. During the branch and bound process, the iteration follows the default best-bound search strategy, which always explores the nodes with the best LP solution unless it is fathomed. There are three scenarios when a node is fathomed, namely:

- if the solution of a sub-problem at the node is infeasible;
- if the value of the objective function at the node is worse than the cut-off;
- if the linear programming relaxation at the node is an integer solution.

For complex MILP formulations, the problem arises when the number of integer variables is too high for the solver to handle, and the branch and bound process may be significantly slowed down due to the large number of branches. Detailed illustration is given in the following sections.

The solver's CPU time depends on the number of iterations the branch and bound search requires to detect and eliminate all the fathomed nodes before locating the optimal integer solution. However, for some complicated MILP problems, the actual solution node may not be processed after more complex nodes with better initial bounds are checked first. As a result, the solver may take a rather large number

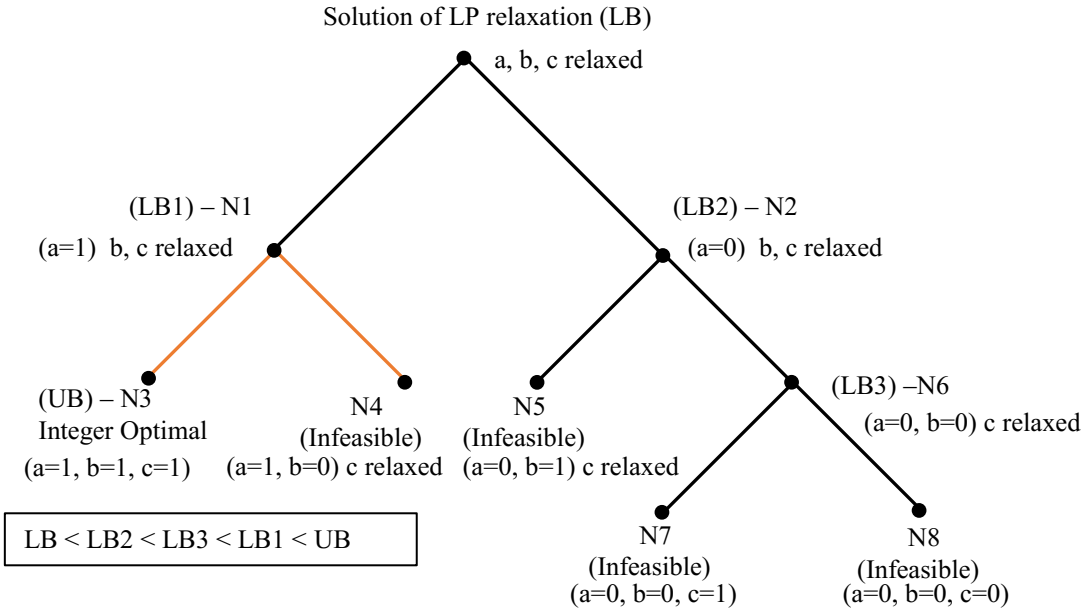


Figure 5.3: An illustration of the branch and bound process for MILP problems with three binary variables

of iterations and CPU time to find the solution node. For example, in Figure 5.3, the branch and bound tree (for a minimisation problem with three binary variables – a, b and c) searches through all eight nodes until it finds the optimal integer solution at node N3. Node N2 is the first to branch based on the best-bound first heuristic. However, after a full tree search, all the subnodes of node N2 turn out to be infeasible and thus the algorithm goes back to the other branch of the initial LP solution (N1) and finds the optimal integer solution at node N3. In this simple example, it is evident that even with only three integer variables, the complication can still arise in the node elimination process leading to unnecessary but compulsory iterations, which can take up a significant proportion of the CPU time (branch N2).

5.1.3. Feasible Region

The branch and bound process searches for the solution from all the nodes within the problem's feasible region. In mathematical optimisation, a feasible region is the set of all possible values of the selected variables of an optimisation problem that satisfy all of the problem's constraints. It is generally difficult to showcase the feasible region of an entire optimisation problem as the problem can be very complex

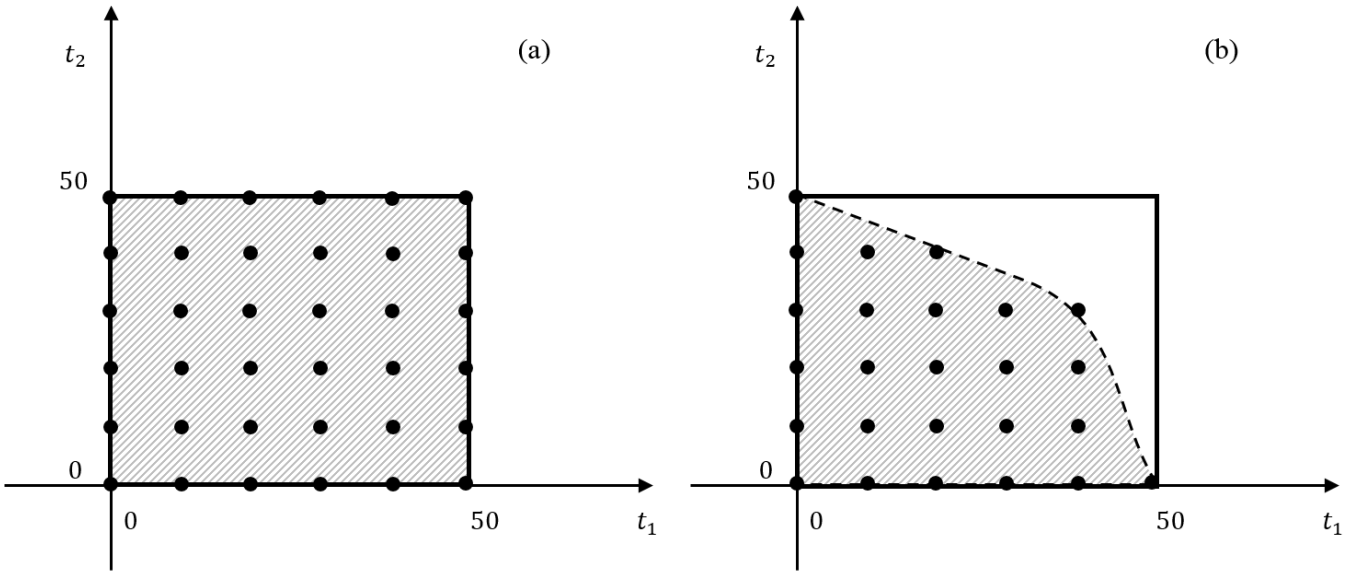


Figure 5.4 a): $t_1 - t_2$ space; b) Feasible region of the whole problem in $t_1 - t_2$ space

Fand involve multi-dimensional feasible space. However, the feasible space of the objective function that consists of only a couple of variables is easier to visualise (a 2-d space). For instance, the continuous feasible space of P1 is comprised of only variables t_1 and t_2 is shown as the shadow area in Figure 5.4a. In P2, both t_1 and t_2 are discretised, their feasible space or in this case the feasible points are given as the dots in Figure 5.4a. Both variables are within the range of $[0,50]$ with an incremental value of $\Delta T = 10$ (i.e. both variables can only take values of 10 or its multiples). If there are a few other constraints in the $t_1 - t_2$ space, the feasible region can then be represented as the dot points Figure 5.4b. For each feasible point, there are two associated binary variables (one each for t_1 and t_2) and are written as $b_{1,n}$

and $b_{2,n}$. As discussed in section 5.1.2, each binary variable adds multiple of $\times 2$ to the maximum number of branch and bounds nodes.

The objective function is concave (for a maximisation problem) and linear in the $t_1 - t_2$ space, and the optimal solution in the $t_1 - t_2$ feasible region ($(t_1, t_2) = (30,40)$) is shown via the graphical method in Figure 5.5a. Now if the feasible space is divided into four identical compartments (Figure 5.5b), it is clear that the optimal solution lies in the second compartment (ii). This suggests that if the bounds of the variable pair $[t_1, t_2]$ are tighten to $[(30,40,50), (30,40,50)]$, the problem can still achieve its optimal solution without going through all the other three redundant compartments. By doing so, the sub-optimal parts of the feasible region (i, iii, and iv) are removed. As a result, the number of potential nodes searched is significantly lowered and hence, in theory, the solver's CPU time is reduced.

However, it is not efficient to check the optimality of each sub-space of the feasible region by solving them sequentially. Therefore, the challenge remains for researchers to solve large-scale MILP problems with discretised variables within a reasonably small time window.

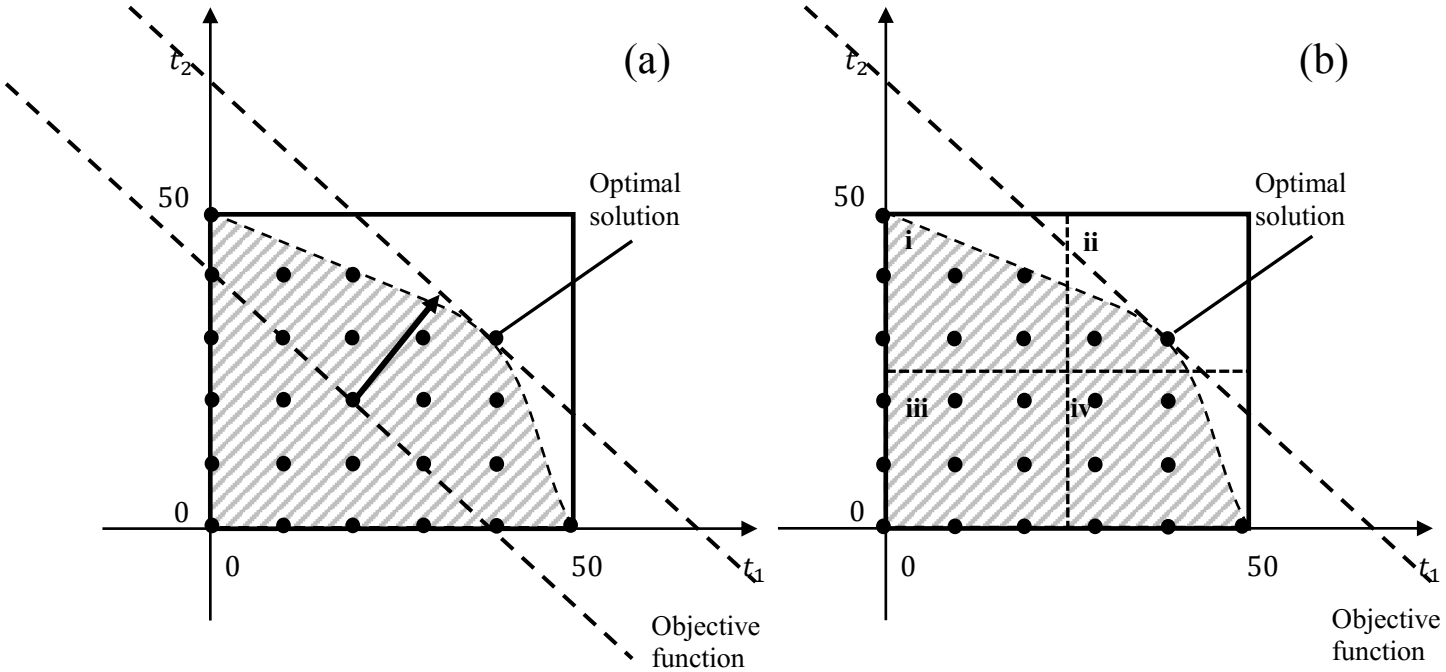


Figure 5.5 a): Objective function and optimal solution using graphical method; b) feasible region partition

5.2. Problem Statement

In optimisation problems where a continuous variable is discretised into a set of binary variables, the quality of the solution is sacrificed to allow more flexible formulation and to reduce the computational time. Such an example is given as the case study in chapter IV. Nevertheless, the number of discretised intervals can only reach a certain point before the MILP problem becomes too complicated to solve. Therefore, the problem to be solved in this chapter can be described as follows:

Given are:

- a mixed integer linear programming (MILP) optimisation problem;

- one or more than one of the continuous variables of the problem have been discretised to avoid nonlinearity;
- for each point of the discretised variable, a binary variable is assigned;
- the accuracy of the MILP model depends on the number of intervals defined, i.e. the smaller the interval size, the more accurate the results;
- the solution computational time increases exponentially when the number of discretisation intervals increases and eventually get outcompeted by the solution time of the nonlinear problem;

to increase the quality of the solution by increasing the number of intervals while maintaining the computational time at a reasonably low level (i.e. minutes to hours).

5.3. Methodology

The feasible region partition method is first explained in section 5.3.1. The issue arises when subproblems are more difficult to solve than the original problem. The issue is tackled by combining the feasible region partition method with inter-job communication, which is discussed in section 5.3.2.

5.3.1. Partition of the Feasible Region

First, the variable t in the original MILP problem P2 is reformulated to two explicit discretised variables (t_1, t_2) and P2 becomes P3 which is written as follows:

$$
\begin{aligned}
&Maximise\ z = c_1 x + c_2 E \\
&subject\ to: \\
&\quad f(x, Y) = 0; \\
&\quad g(x, Y) \le 0; \\
&\quad w(t_1, t_2, Y) = 0; \\
&\quad h'(x, t_1, t_2) - E = 0; \\
&\quad d\left(b_{1,n}, b_{2,n}, t_1, t_2\right) = 0 \\
&x \in [0, X], t_1, t_2 \in [0{,}50, \Delta T = 10], E \in R; \\
&\quad Y, b_{1,n}, b_{2,n} \in \{0{,}1\}
\end{aligned}
\qquad \text{P3}
$$

Second, the feasible space of each of the selected discretised variables is divided into two equal-sized subdomains, i.e. [0, 10, 20] and [30, 40, 50]. As a result, the original problem becomes four sub-problems (SP1, SP2, SP3 and SP4) with mutually exclusive feasible spaces in t_1 and t_2 as shown in Table 5.1.

Table 5.1: Modified variable range of the sub-problems

SP1	SP2	SP3	SP4
$t_1 \in [0{,}10{,}20]$	$t_1 \in [30{,}40{,}50]$	$t_1 \in [0{,}10{,}20]$	$t_1 \in [30{,}40{,}50]$
$t_2 \in [30{,}40{,}20]$	$t_2 \in [30{,}40{,}50]$	$t_2 \in [0{,}10{,}20]$	$t_2 \in [0{,}10{,}20]$

Each of the four independent sub-problems has a reduced number of integer variables as each only represents a quarter of the original problem. They are submitted to the grid and solved in parallel, and

the graphical representation of the solution of each subproblem is given in Figure 5.6. As can be seen, the point in the $t_1 - t_2$ space that corresponds to the optimal objective value of each sub-problem is given as the intersection point between the objective function and the t_1 axis (labelled as I in Figure 5.6). The optimal solution of SP2 (I_2) is greater than the solutions of all other three sub-problems (i.e. $I_2 > I_1 = I_4 > I_3$), and therefore is the optimal solution of the entire problem. In theory, each of the reduced-size subproblem should take less CPU time to solve than the original problem, and the partition

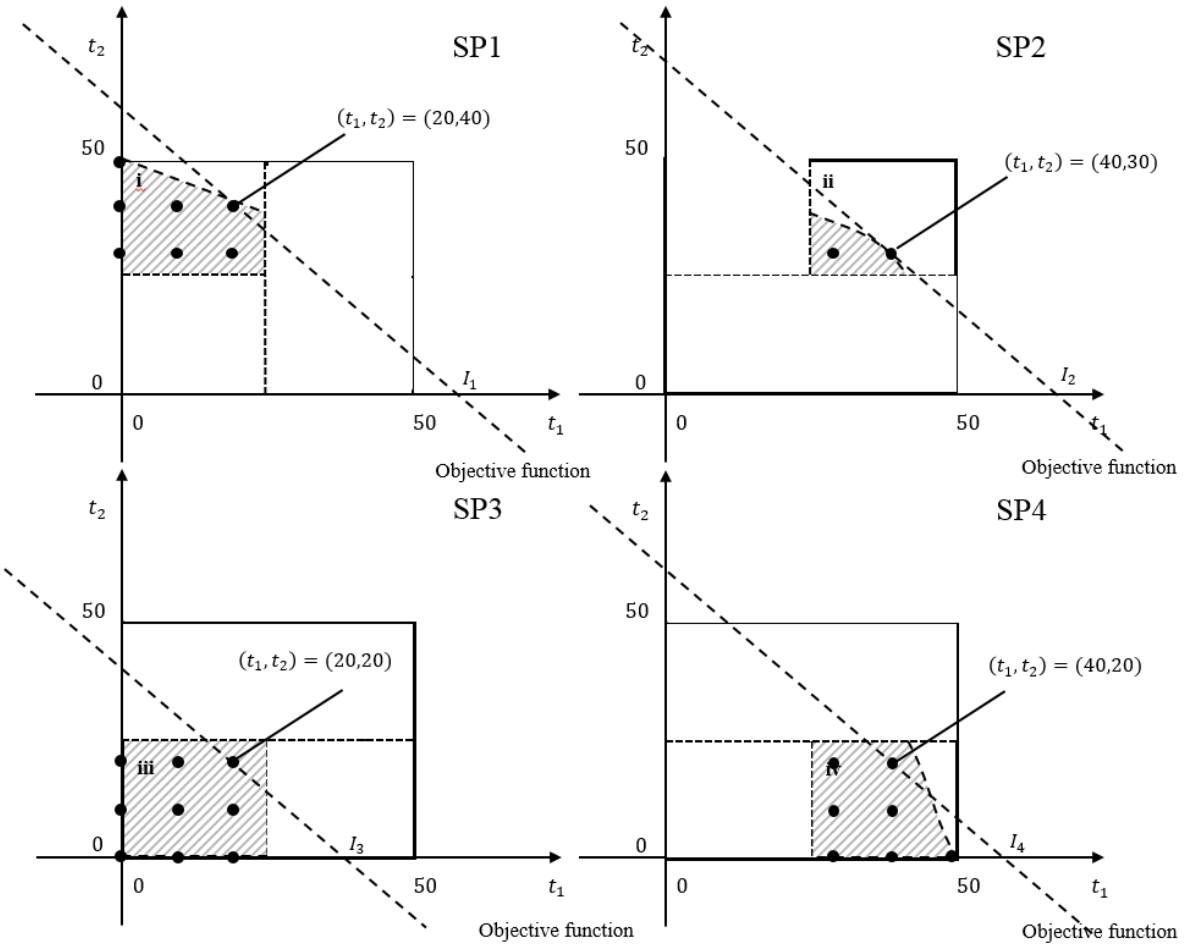


Figure 5.6 i), ii), iii), iv): Graphical illustration of each sub-problem and their respective optimal solutions

of the feasible region of the original problem should result in much less computational effort. However, the solution speed of each sub-problem is not necessarily correlated with the relative size of the problem, and can vary depending on problem structures (i.e. some subproblems may be very easy to solve, while others require more computational effort). In some cases, the subproblem with the best solution can be the bottle-neck problem, while the other subproblems are solved with relative ease, and their solutions are readily available at an early stage of the solution process (a detailed example is given in section 5.3.2). As a result, the CPU time of the entire solution process depends on the bottle-neck problem (i.e. sub-problem with the longest solution time). In other words, the feasible region partition method does not guarantee a faster solution speed than solving the original problem.

Therefore, in the next section, an inter-job communication mechanism is introduced to allow the best incumbent value among all the sub-problems to be periodically fed to the other problems as their 'cutoff' values, and any unpromising subproblems can thus be terminated immediately (i.e. any problems with the current best solution that is worse than the 'cutoff' value).

5.3.2. Communication between Jobs

In some branch and bound cases, one branch may seem more promising than the other because of their respective best bound values, but it does not necessarily yield the best integer solution. For example, Figure 5.7 illustrates a full enumeration solution process of a maximisation problem where the early promising node N3 (with better bound) yields an optimal integer solution (N9) that is worse than the integer solution of node N12 (from N4). The actual branch and bound process based on the best-bound search (usually the default setting in CPLEX) is shown as the dashed line and is described as follows:

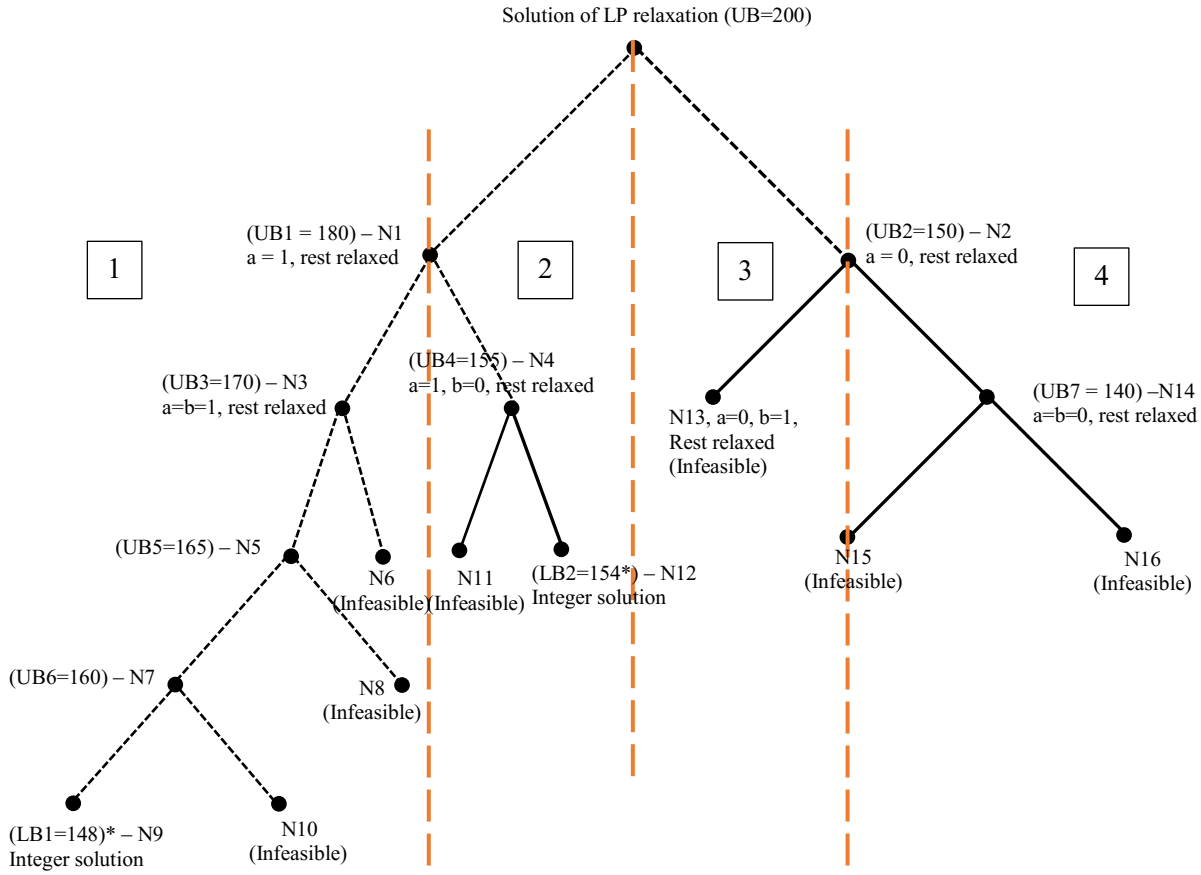


Figure 5.7: Full solution enumeration of the example maximisation problem

- LP relaxation is first branched to nodes N1 and N2,
- Since UB1>UB2, node N1 firstly branches to N3 and N4 while N2 remains open,
- Node N3 is a better bound than node N4 and thus continues branching,
- N6 turns out to be infeasible and is fathomed, and node N5 with a bound value of 165 becomes the new upper bound of the problem,

- N5 further branches to nodes N7 and N8; while node N8 is an infeasible node, N7 replaces N5 to be the new upper bound,
- Node N9, which is a subnode of N7, is both a feasible integer solution and a new incumbent value,
- In a perfect branch and bound process, the remaining open nodes, N2 and N4, both of which have better bounds than the incumbent value, will be further branched until fathomed,
- In this case, since the relative stopping tolerance, i.e. the relative gap between the best integer objective and the objective of the best node remaining, is set to be 0.1 and therefore, both N2 and N4 will not be processed since the relative gap between the current incumbent and the bound values of those two nodes are within tolerance (i.e. $\frac{UB4-LB1}{UB6}, \frac{UB2-LB1}{UB6} < 0.1$)
- As a result, the branch and bound process is terminated with an optimal solution of 148.
- However, there exists a better integer solution that also satisfies the default relative gap tolerance (N12) and will not be found unless the stopping tolerance is reduced.

To avoid the scenario mentioned above, the partition of feasible region combined with the inter-job communication mechanism is implemented. The original problem is divided into four subproblems (divided by the dashed line shown in Figure 5.7) based on the values of binary variables a and b, which is equivalent to manually branching on a and b. The four problems are solved in parallel on a cluster. Every time one of the problems (j) is solved, and if its solution (Z_j^{COMP}) is better than the current one (Z^{INCU}), the new one will replace the old one as the new 'cutoff' value and is stored in a file ($Z^{INCU} = Z_j^{COMP}$). Meanwhile, all the other unfinished jobs constantly read the 'cutoff' value (cutoff $= Z^{INCU}$), and if the current best bound is worse than the 'cutoff' value, the problem is declared to be fathomed and thus terminated. On the other hand, if the current best bound is better, then the 'cutoff' value serves as an early lower bound, which could potentially increase the solution speed by eliminating nodes that are worse than the 'cutoff' value. When the second branch is solved, its final integer solution (OBJ=154) becomes the 'cutoff' value, and the other three problems, if still unsolved, will immediately terminate once the 'cutoff' value is updated. Therefore, in this scenario, the CPU time saved compared with the direct branch and bound process is equal to the time required to branch from Node N3 to N9.

The procedure of solving all subproblems of a maximisation problem with inter-job communication is given as follows with a diagram illustration shown in Figure 5.8 (Detailed implementation in GAMS in available in Appendix D):

SET $Z^{INCU} = -inf$;

SUBMIT subproblems $j \in J$ to the computational cluster;

WHILE the number of unsolved problems ≥ 0:

SOLVE subproblems $j \in J$ in parallel using CPLEX

RETRIEVE incumbent values of each subproblems j (Z_j^{COMP})

IF $Z_j^{COMP} \geq Z^{INCU}$

UPDATE $Z^{INCU} = Z_j^{COMP}$

END

FOR unsolved problems

READ Z^{INCU} as the new 'cutoff'

END

END

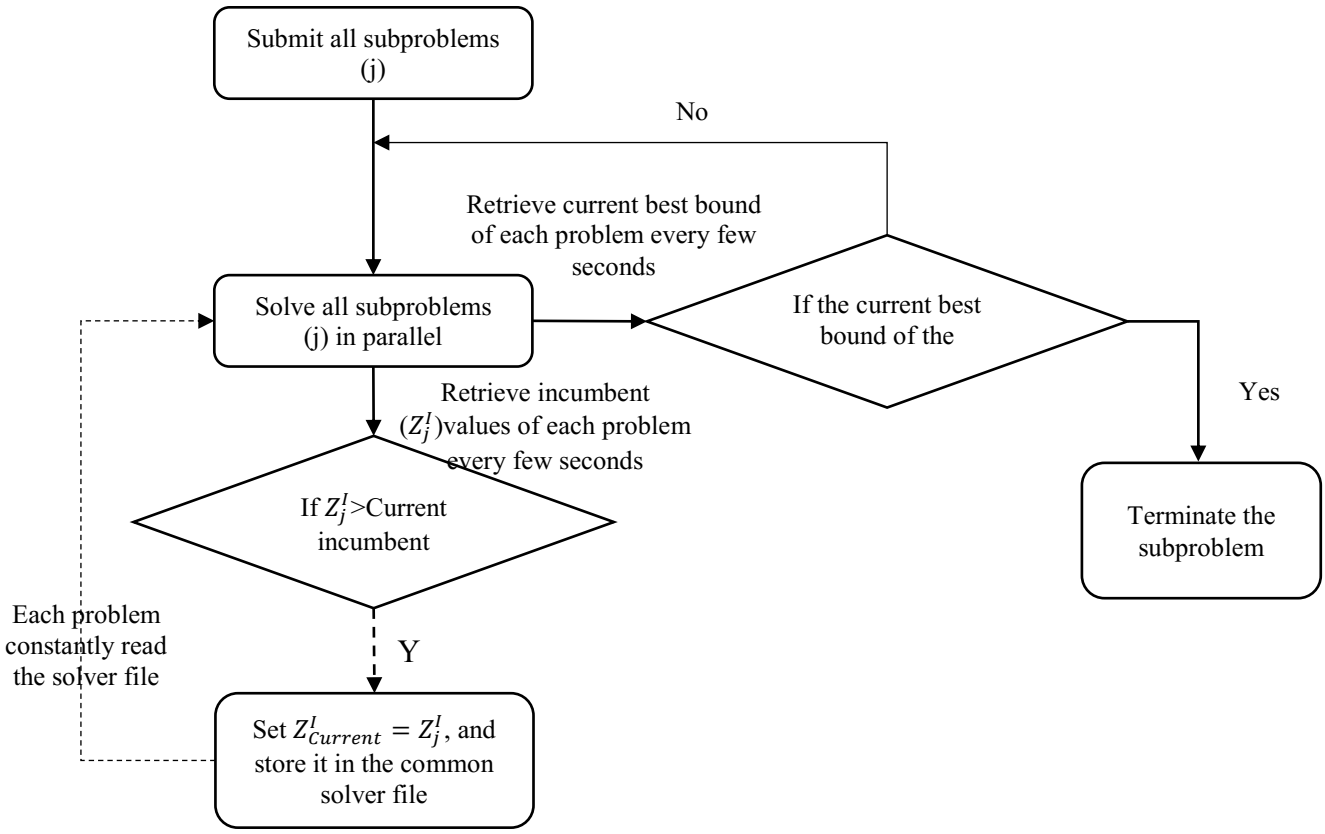


Figure 5.8. Solution procedure of the proposed method

5.4. Case study

For demonstration purposes, a new case study that concerns the reaction pathways towards polymers is presented. The complete reaction network is given in Figure 5.9. It is shown that d-limonene can self-polymerises to form a polyalkane (poly-limonene) at around 85 ℃ (Singh and Kamal 2012). Free radical coplymerisation between d-limonene and other alkenes (e.g. styrene and methyl methacrylate) is also possible through reactions R2 and R3 (Sharma and Srivastava 2003, Sharma and Srivastava 2004). Limonene oxide, which is one of the valuable derivatives of d-limonene (R4), can react with carbon dioxide (R6) to form polycarbonate (Byrne et al. 2004) and with anhydrides (phthalic anhydride and maleic anhydride) to produce polyesters via reactions R7 and R8 (Jeske et al. 2007, Nejad et al. 2013). On the other hand, another derivative from d-limonene (R9), p-cymene can be converted to terephthalic acid (R10), which can be further alkalised to form dimethyl terephthalate (R11). The resulting salt is an ingredient in reacting with diols (R12 and R13) to produce polyesters (Berti et al. 2010, Colonna et al. 2011). The characteristics of the reactions and the properties of the chemicals/polymers are all estimated via the previously mentioned group contribution method and the data are summarised in Appendix B.1.

In this case study, the mathematical formulations of Chapter III and Chapter IV are used. However, since some of the streams may contain polymers, distillation can no longer be used as the only separation technology as it cannot separate polymers from monomers. Therefore, a conceptual dissolution-precipitation mechanism adopted from Pappa et al. (2001) is integrated to allow polymer recovery before distillation. The new polymer separation unit is illustrated in Figure 5.10, and the detailed formulation of the separation unit is given in Appendix B.2.

The reactor effluent of each stage k contains unreacted monomers and product polymer, which can be difficult to separate directly due to its viscosity and complicated interaction with the solution. Therefore, a versatile organic solvent ($Q_{i,k}^{solvent}$) is selected to dissolve the polymer/monomer mixture ($Q_{i,k}^{polysep}$) in the dissolution tank. After dissolution, the resulting solution ($Q_{i,k}^{solution}$) enters a selective precipitation tank. A designated anti-solvent, which can significantly reduce the solubility of the target polymer, is added to the solution to precipitate the polymer. The slurry mixture ($Q_{i,k}^{precip}$) contains pure polymer in the solid phase and monomers in the liquid phase and is passed through a filtration unit. The filtered polymer product ($Q_{i,k}^{filterB}$) is washed and stored before shipping to the market. The valuable components in the residual liquid mixture ($Q_{i,k}^{filterT}$), in turn, are recycled back to the distillation to be separated.

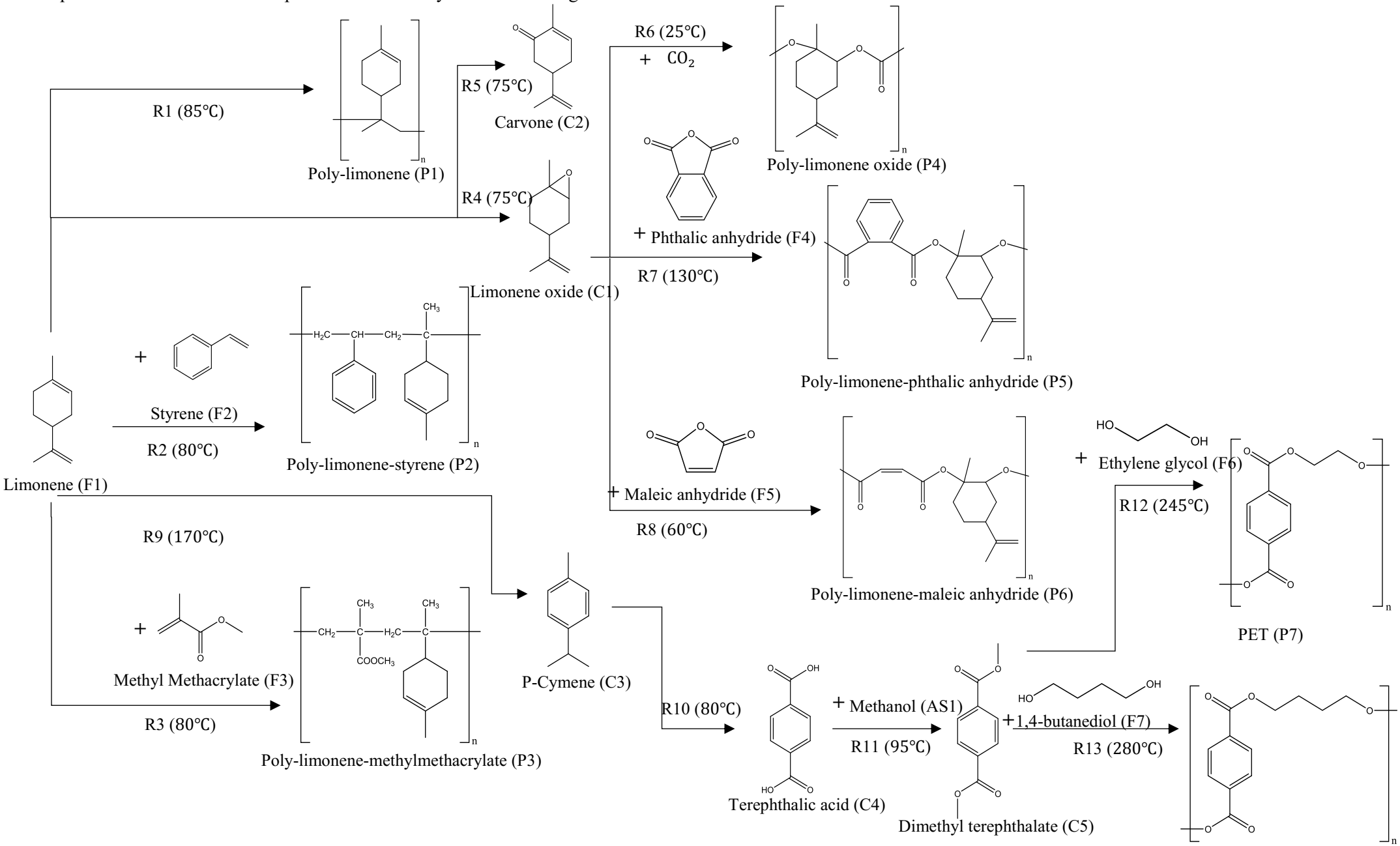


Figure 5.9: Polymer synthesis reaction network for the case study

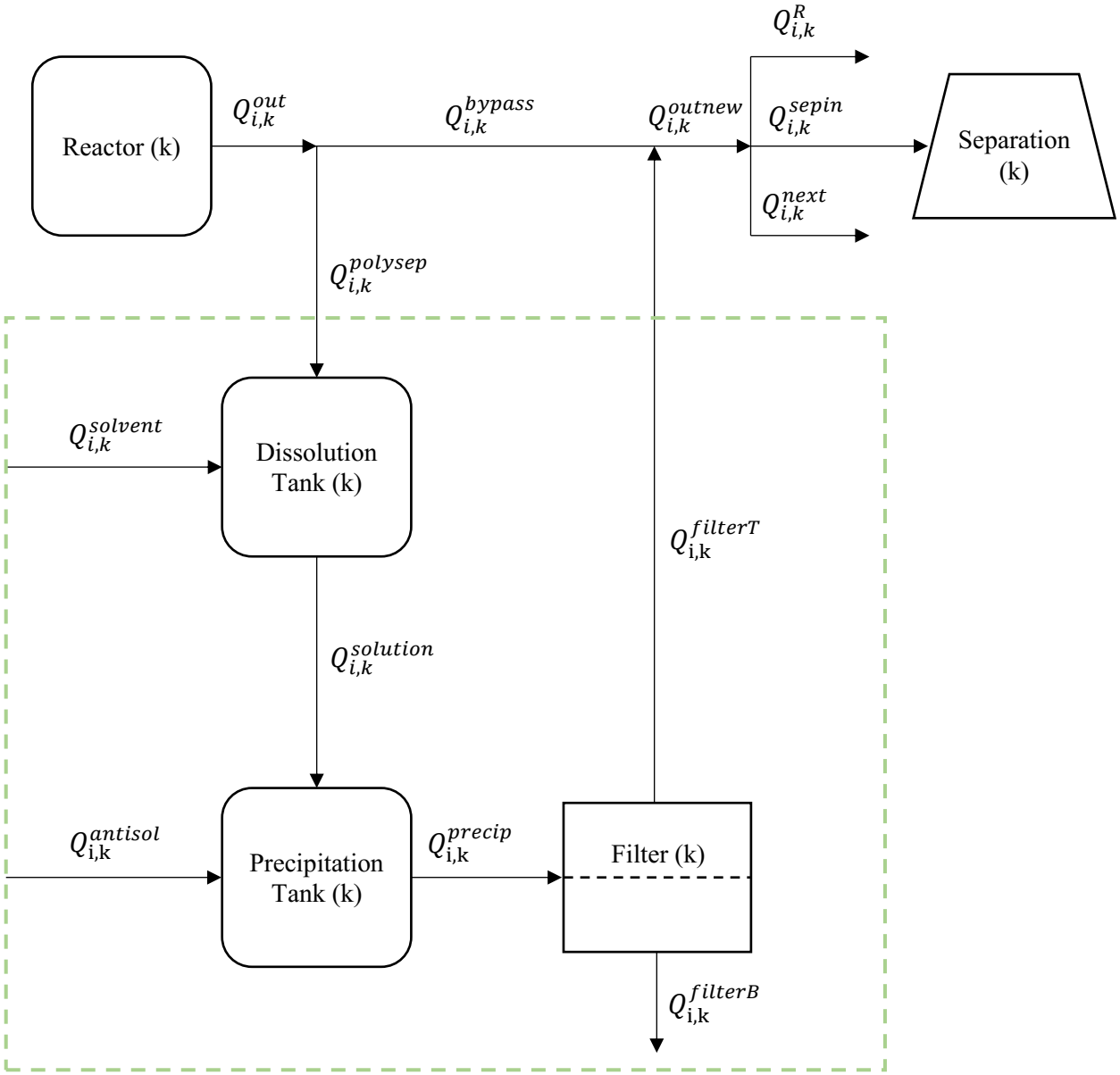


Figure 5.10: Flow diagram of the conceptual dissolution – precipitation separation process

5.4.1. Initial Solutions

The problem is programmed as a discretised MILP problems and solved in GAMS with a discretised temperature interval of $\Delta T = 2$℃. The optimal reaction network is given in Figure 5.11. As can be seen from the diagram, the solution gives that poly-limonene oxide (P4) is the best-choice product among all the candidate polymers. As a result, limonene (F1) is first converted to the monomer limonene oxide (C1) via R4 with a side reaction R5 producing carvone (C2) at the same time. The monomer is then purified and converted to the product polymer (poly-limonene oxide) via R6 with the constant addition of carbon dioxide at 25 ℃.

The corresponding optimal flowsheet of the optimisation problem is given in Figure 4.4. At master stage k1, fresh d-limonene is preheated from 20 °C to 76 °C and fed to the first reactor. Two

simultaneous reactions R4 and R5 in reactor k1 convert the reactant into a mixture of component C1 and C2 at a temperature of 76 °C, which are then heated to 198 °C before entering the distillation system. The distillation system has only one column (k1.l1), which separates carvone from the bottom and limonene oxide from the top. The former is then cooled to 20 °C before being removed to storage while the latter is cooled to 26 °C before mixing at stage k2. The operating temperature of the condenser and the reboiler of column k1.l1 are 198 °C and 232 °C respectively.

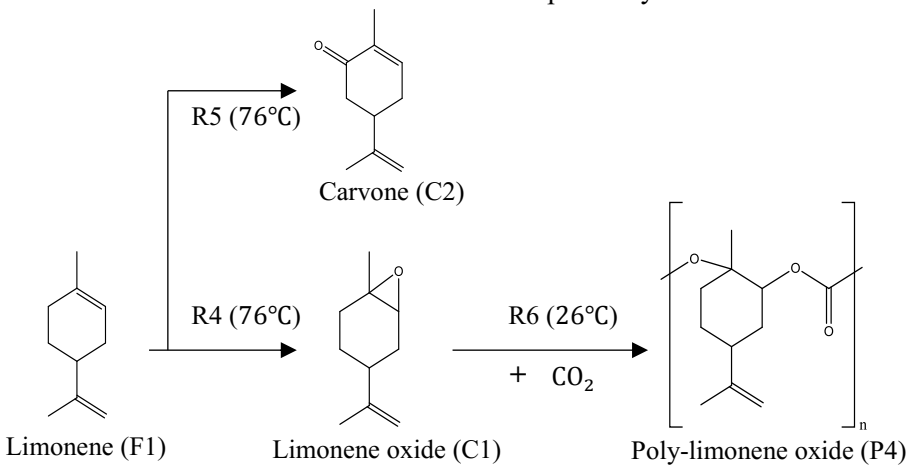


Figure 5.11: The optimal reaction network for the case study

At master stage k2, separated limonene oxide from the previous stage is mixed with the heated recycle stream before entering the reactor, where the polymerisation takes place at 26°C. The undissolved part of the product polymer (P4) forms a slurry with the unreacted limonene oxide. The slurry mixture is then dissolved using organic solvent toluene (S1) in a dissolution tank. To completely isolate the polymer, a specific type of anti-solvent (methanol) is added to the solution mixture in the precipitation tank and resulting in a two-phase mixture. The solid phase, which contains only the polymer, is separated via a filtration unit and stored while the liquid phase is further separated in the distillation system of stage k2. In the first distillation column (k2.l1), the mixture of the anti-solvent and solvent (methanol and toluene) is separated from the top as distillate and fed to the second column. On the other hand, the bottom product is pure limonene oxide and is cooled to be recycled to the reactor. The condenser and the reboiler of the first column operate at 66 °C and 198 °C. In the second column, the anti-solvent (AS1) and the solvent (S1) are removed from the top and bottom respectively. The operating temperature of the condenser and the reboiler are 66 °C and 110 °C. There are 17 utility 'streams' with enthalpy change in the entire process including all the non-heat exchanger units (e.g. reactor, reboiler and condensers) and they are listed in Table 4.3 with the discretised grand composite curve given in Figure 4.5.

Table 5.2: Inlet/outlet temperatures of each 'stream'

Stream No.	Utility Stream Characterisation	Inlet temperature (°C)	Outlet temperature (°C)	Flowrate (mol/min)
1	Stream	20	76	288.4
2	Reactor k1	76	76	-
3	Stream	76	198	288.4
4	Condenser k1.l1	198	198	-
5	Reboiler k1.l1	232	232	-
6	Stream	232	20	116.8
7	Stream	198	26	171.6
8	Stream	20	26	19.1
9	Reactor k2	26	26	-
10	Stream	26	66	70.6
11	Condenser k2.l1	66	66	-
12	Reboiler k2.l1	198	198	-
13	Stream	198	20	19.1
14	Condenser k2.l2	66	66	-
15	Stream	66	20	34.3
16	Reboiler k2.l2	110	110	-
17	Stream	110	20	17.2

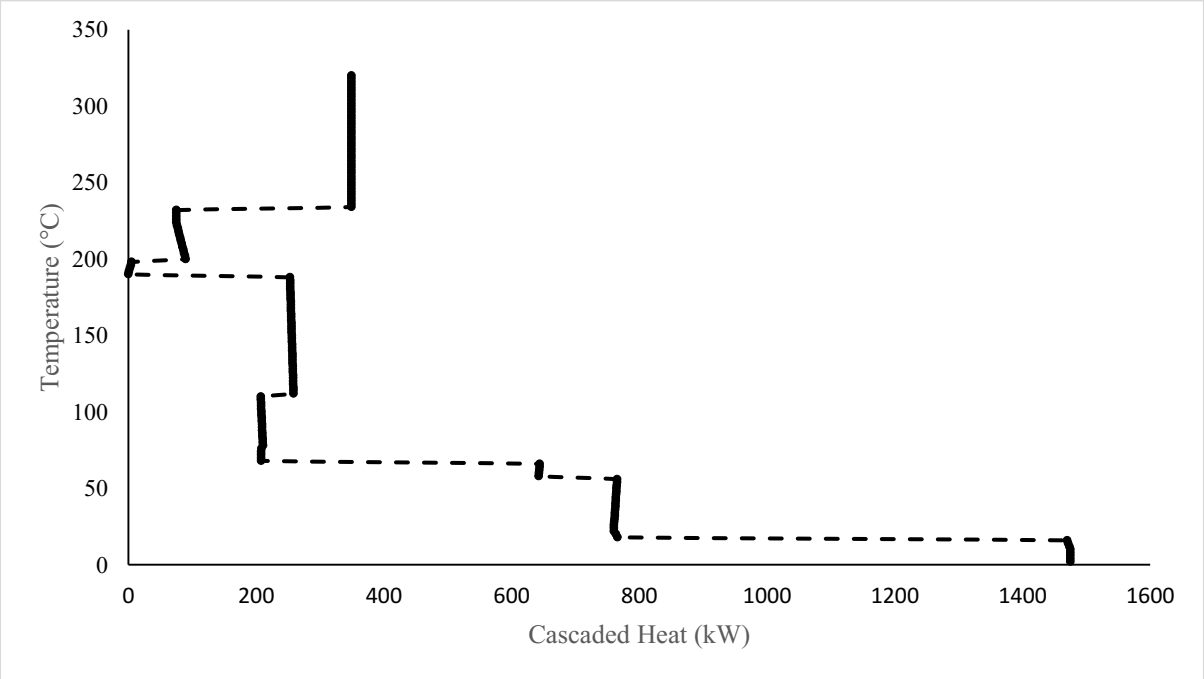


Figure 5.12. The discretised grand composite curve ($\Delta \boldsymbol{T} = \mathbf{2}$ **°C**)

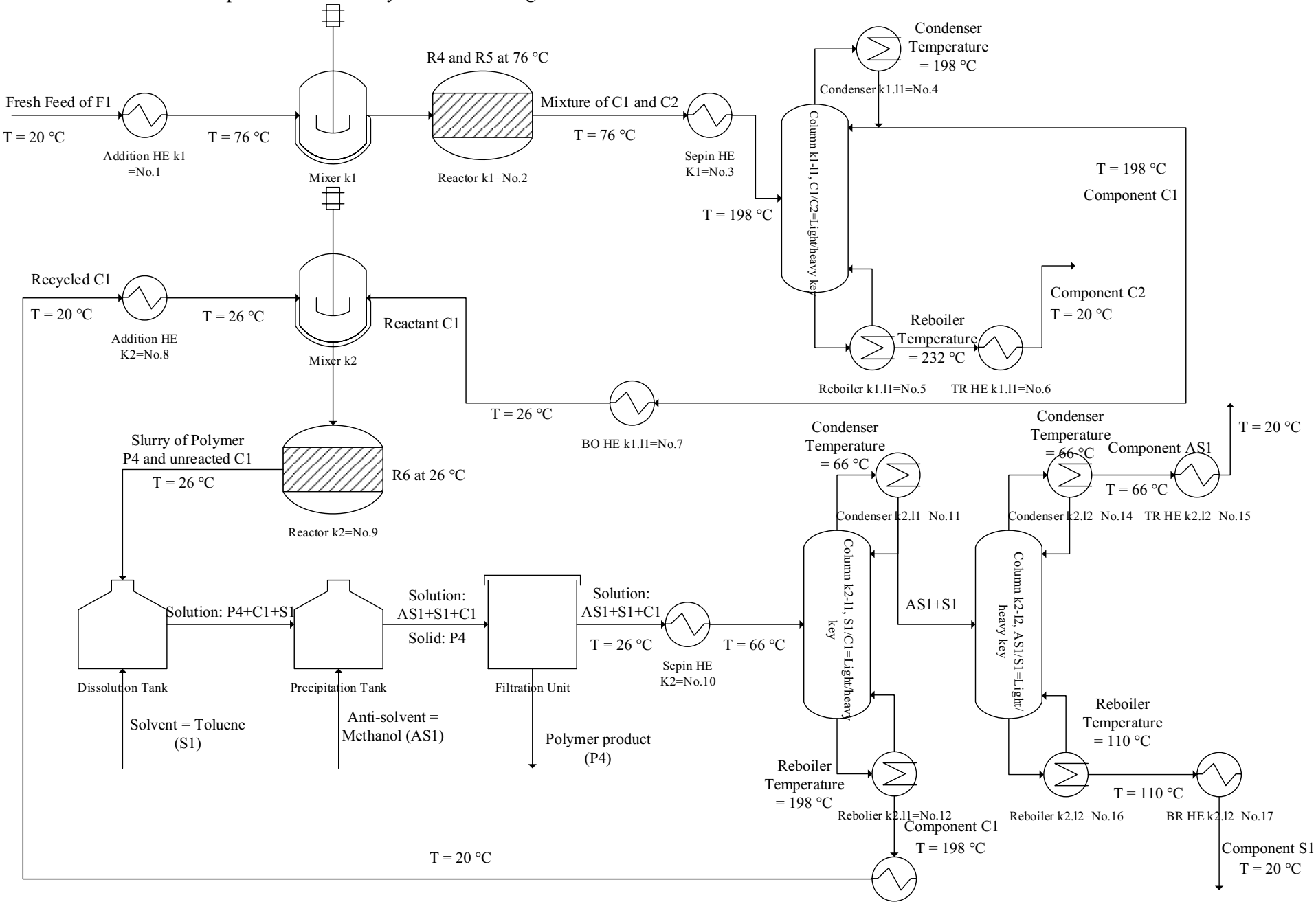


Figure 5.13. Optimal process flowsheet of the case study (temperature intervals of 2 °C)

The gross marginal profit of the optimal process is roughly \$115.2 million per annum (this does not account for the capital cost of the polymer recovery unit). There are three distillation columns, which results in an annualised separation cost of \$0.2 million. The theoretical minimum heating and cooling duties required for heat integration are 349.4 kW and 1475.2 kW respectively. The total utility cost is about \$0.52 million (1/3 heating cost and 2/3 cooling cost). The annual rate of production of the polymer (poly-limonene oxide) is $1.33 \times 10^4\ t$ with an annual consumption rate of d-limonene at $2.0 \times 10^4\ t$.

The relative optimality tolerance of the original problem is set to be 0.1% and the resulting CPU time of the CPLEX solver is 901.3s. The detailed statistics of the model is given in Table 5.3.

Table 5.3: Model statics of the original problems

Single Variables	147,460	**MIP Iterations**	1,266,799
Discrete Variables	64,004	**MIP Nodes Explored**	964
Relative Tolerance	0.001	**CPU time (s)**	901.3
Best Possible	115.3×10^6	**Final Solve**	115.2×10^6

The problem is then solved using different values of relative tolerance, and the results are listed in Table 5.4. As can be seen, the CPU time is reduced from 901s to 779s when the tolerance is increased from 0.001 to 0.01. This is because the final solve for the latter scenario stops at an incumbent of 114.7×10^6 rather than 115.2×10^6. As a result, the number of nodes explored and the number of MIP iterations are both reduced considerably, and hence the CPU time. On the other hand, when the tolerance is further increased from 0.01 to 0.1, the final solve incumbent drops significantly but without obvious improvement in CPU time. This can be explained by the little change in the number of nodes and the number of iterations used.

Table 5.4: Comparison of the solution statistics for different values of relative stopping tolerance

Relative Tolerance	MIP Iterations	MIP Nodes Explored	Final Solve	Best Possible	CPU time (s)
0.001	1,266,799	964	115.2×10^6	115.3×10^6	950
0.01	1,006,661	773	114.7×10^6	115.6×10^6	790
0.1	964,118	760	112.0×10^6	115.6×10^6	788

In summary, the CPU time does not reduce significantly compared with the percentage of change in the relative stopping tolerance. On the other hand, the gap between the final objective value and the best possible solution becomes too large to be ignored when the tolerance is increased. Therefore, for this case study, 0.001 is selected to be the relative stopping tolerance to maintain the solution quality.

In the following sections, the feasible region partition method is implemented to systematically reduce the CPU time without compromising the quality of the solution. The variables selected to be partitioned must follow the selection heuristics, which is based on author's knowledge and understanding of the

model: 1) the selected variable's domains include the one that to be partitioned, which in this case is the temperature index t); 2) by partitioning the selected variable over the discretised domain (t), the integrity of each subproblem must be preserved (i.e. only the feasible region is changed, not the constraints); 3) the variable is preferable to be partitioned if it has direct links to other binary variables (i.e. partitioning on one variable is, in fact, partitioning on multiple inter-related variables). In this case study, the variables chosen are the binary variables used for the temperature point of each reaction stage in chapter IV (i.e. $\hat{y}_{k,t}$). The reason for this choice is given as follows:

- First, this binary variable is tightly constrained by the reaction binary variable ($y_{j,k}$), which indicates the on/off status of that stage,
- Second, fixing this variable does not alter the structure of the subproblems as they all have the same formulation but with mutually exclusive feasible regions (This is illustrated in Table 5.5).
- Third, four other binary variables are directly linked to it (i.e. $\widehat{ye}_{k,t}^{A}$, $\widehat{ys}_{k,t}^{Qsepin}$, $\widehat{ys}_{k,t}^{Qnext}$ and $\widehat{ys}_{k,t}^{QR}$) and by fixing part of the binary variable set (i.e. $\hat{y}_{k,t} = 0, \forall\, 0 \leq t \leq 40$), the feasible space of the other four binary variables are automatically fixed.

Table 5.5: Theoretical solution status of example subproblems

	Original solution	**Subproblem 1**	**Subproblem 2**
Partition	$\hat{y}_{k,t}$ not partitioned	$\hat{y}_{k,t} = 0,$ $\forall k = 1, 0 \leq t \leq 80$	$\hat{y}_{k,t} = 0,$ $\forall k = 1, 80 \leq t \leq 160$
Solution status	Integer optimal	Infeasible	Integer optimal
Solution	$\hat{y}_{k=1,t=38} = 1, \hat{y}_{k=2,t=13} = 1$	No solution	$\hat{y}_{k=1,t=38} = 1, \hat{y}_{k=2,t=13} = 1$

The method is first implemented without any communication between each subproblem. The results are then compared against scenarios with inter-job communication.

5.4.2. Variable Partition without Communication

In GAMS, when submitting subproblems to a cluster, problem generation is always sequential. For small problems, it will not have huge effects on the solution time. However, in this case, each subproblem takes about 3-4 seconds to generate regardless of how the problem is partitioned (The partition only fix specific variables but not reduce the total number of variables). As a result, the more subproblems there are, the longer it will take to completely generate all problems. Therefore, in the following sections, CPU time refers solely to the time spent by CPU to solve the problem, whereas execution time refers to the CPU time plus the problem generation time. To determine the optimal partition strategy, the temperature index t of the binary variable of each master stage k ($\hat{y}_{k,t}$) is first partitioned to two equal-sized ranges ($t \in [0, 80]$, and $t \in [81,160]$). Since four redundant master stages k are initially provided, the original problem is thus divided into $2^4 = 16$ subproblems with mutually exclusive feasible space in t. Suppose $[0, 80]$ is integer range R_1 and $[81,160]$ is integer

range R_2 (Grey boxes in Table 5.6), the 16 subproblems can be defined by the different feasible region of binary variable $\hat{y}_{k,t}$ as follows:

Table 5.6: Feasible region of each subproblem

Range for k and t/	$\hat{y}_{k,t} = 0, \forall k, t \; below$			
Subproblems	$k = 1$	$k = 2$	$k = 3$	$k = 4$
1	$t \in R_2$	$t \in R_2$	$t \in R_2$	$t \in R_2$
2	$t \in R_2$	$t \in R_2$	$t \in R_2$	$t \in R_1$
3	$t \in R_2$	$t \in R_2$	$t \in R_1$	$t \in R_2$
4	$t \in R_2$	$t \in R_2$	$t \in R_1$	$t \in R_1$
5	$t \in R_2$	$t \in R_1$	$t \in R_2$	$t \in R_2$
6	$t \in R_2$	$t \in R_1$	$t \in R_2$	$t \in R_1$
7	$t \in R_2$	$t \in R_1$	$t \in R_1$	$t \in R_2$
8	$t \in R_2$	$t \in R_1$	$t \in R_1$	$t \in R_1$
9	$t \in R_1$	$t \in R_2$	$t \in R_2$	$t \in R_2$
10	$t \in R_1$	$t \in R_2$	$t \in R_2$	$t \in R_1$
11	$t \in R_1$	$t \in R_2$	$t \in R_1$	$t \in R_2$
12	$t \in R_1$	$t \in R_2$	$t \in R_1$	$t \in R_1$
13	$t \in R_1$	$t \in R_1$	$t \in R_2$	$t \in R_2$
14	$t \in R_1$	$t \in R_1$	$t \in R_2$	$t \in R_1$
15	$t \in R_1$	$t \in R_1$	$t \in R_1$	$t \in R_2$
16	$t \in R_1$	$t \in R_1$	$t \in R_1$	$t \in R_1$

All subproblems are programmed in GAMS and submitted to the grid to be solved in parallel. The solution statistics of each subproblem is given in Table 5.7. Surprisingly, some of the subproblems (e.g. subproblem 2, 3, 5 and 9) fail to return the model status (i.e. optimal or infeasible) within the pre-set time frame and are thus terminated. On the other hand, most of subproblems are successfully solved with different CPU times. For instance, subproblems 1, 4, 10 and 13 take longer to solve than the original, and only problem 1 produces the same optimal solution as the original problem. The solutions of subproblems 10 and 13 are sub-optimal compared to the solution of problem 1, but both problems take longer to solve. Subproblem 4 is infeasible, and it also takes 400+ more seconds than problem 1 to solve. The model's overall execution time turns out to be 25 minutes (1500 s) as some subproblems timeout at the 25-minute mark (user-defined time limit), with the optimal subproblem (No.1) takes about 18 minutes. In theory, the individual subproblems with a large number of variables fixed should be solved much faster than the original. However, due to the complications arise from the branch and bound processes of the subproblems, the overall execution time is not improved compared to the

original problem (964s). It is clear that simply creating equivalent subproblems of the original problem and solving the subproblems in parallel does not guarantee a better solution time due to the complex nature of the problem structure. Therefore, communication between jobs is implemented to avoid the drawbacks of the feasible region partition method, and the results are shown and discussed in the next section.

Table 5.7: Model statistics of each subproblem

Subproblems	Final Solution ($\times 10^6$)	CPU time (s)	No. of Iterations ($\times 10^6$)	No. of nodes
1	115.2	1065	1.59	1382
2	TO (Timed Out)	-	-	-
3	TO	-	-	-
4	Infeasible	1483	3.16	2813
5	TO	-	-	-
6	0.25	583	1.09	1435
7	0.25	274	0.527	666
8	0.25	85	0.139	618
9	TO	-	-	-
10	0.616	1340	2.33	2670
11	0.616	897	1.76	3296
12	0.616	78	0.126	182
13	0.616	1196	2.39	2765
14	0.616	74	0.104	127
15	0.616	60	0.069	77
16	0.616	31	0.002	-

5.4.3. Variable Partition with Communication

In this section, the communication protocol discussed in section 5.3.2 is implemented, and the solution statistics of each subproblem are given in Table 5.8. The execution time of the entire parallelisation problem is 474 seconds which is approximately a third of the case with no communication between subproblems. Similar to the solution without inter-job communication, several problems are also quickly solved, and therefore provide early solutions (e.g. problems No. 15 and 16). However, unlike the previous case, the early solutions obtained in this case are used as the 'cutoff' for all the subproblems that are still unfinished. Meanwhile, the incumbent values of those unfinished jobs are constantly checked and the best incumbent among all subproblems, if better than the current 'cutoff', becomes the 'cutoff'. On the other hand, if incumbent of any problem is worse than the current 'cutoff', the problem

is terminated. For example, all the problems from No.2 to No.9 that have solutions worse than the cutoff value of 0.616×10^6 or timed out in the last section, become fathomed thanks to the addition of the 'cutoff' value. In addition, some of the difficult problems (No.10, 11 and 13) become easier to solve than before and still give the same solutions. Problem No.1, which gives the best optimal solution among all subproblems, now only takes 436 seconds to solve compared with 1065 seconds when no communication is available.

Table 5.8: Model statistics of each subproblem (with communication)

Subproblems	Final Solution ($\times 10^6$)	CPU time (s)	No. of Iterations ($\times 10^6$)	No. of nodes
1	115.2	436	0.44	358
2	Infeasible	450	0.298	55
3	Infeasible	381	0.319	153
4	Infeasible	353	0.513	540
5	Infeasible	344	0.423	320
6	Infeasible	341	0.554	731
7	Infeasible	354	0.445	635
8	Infeasible	139	0.174	172
9	Infeasible	441	0.218	33
10	0.616	327	0.511	726
11	0.616	323	0.773	1708
12	0.616	139	0.130	188
13	0.616	315	0.554	775
14	0.616	131	0.139	175
15	0.616	90	0.082	110
16	0.616	39	0.002	-

However, upon repeating the methodology for validation purposes, it is noticed that the results vary significantly even with the same problem and CPU core. Some runs are even worse than the case without any inter-job communication (The CPU times of five additional sample runs with the same parameter are given in Appendix B.3). The possible reasons for this inconsistency are as follows:

- There are no early good incumbent values from all subproblems to completely change the course of the branch and bound process of problem No.1, which is the subproblem that gives the best optimal solution.
- The problems are generated sequentially and the generation time for the same problem depends heavily on the performance of the CPU and can vary slightly during each run based on machine load, CPU temperature, etc.

- This variation can affect the time when the first set of incumbent values are available, which then, in turn, determines the time when the 'cutoff' values are generated and fed to other unfinished problems.
- If the unfinished problems read the proper 'cutoff' value at the right time, their numbers of the branch and bound iterations can be significantly reduced.
- However, if somehow the early 'cutoff' values are not good enough and are read at the wrong time, it can cause the problem to branch further away from the optimal solution and timeout before the solution is found.

To generate good feasible solutions at an early stage and achieve more consistent overall solution time, the partition method is modified to produce more subproblems, but the size of each problem is much smaller.

5.4.4. Variable Partition into Smaller-Size Subproblems

In this section, the t space of the variable $\hat{y}_{k,t}$ is divided into three ranges instead of two (i.e. $t \in [0, 54]$, $t \in [55,108]$ and $t \in [109,160]$) and therefore $3^4 = 81$ subproblems are generated. The subproblems are first solved without inter-job communication and the objective value and the CPU time of each problem are given in Figure 5.14 and Figure 5.15 respectively.

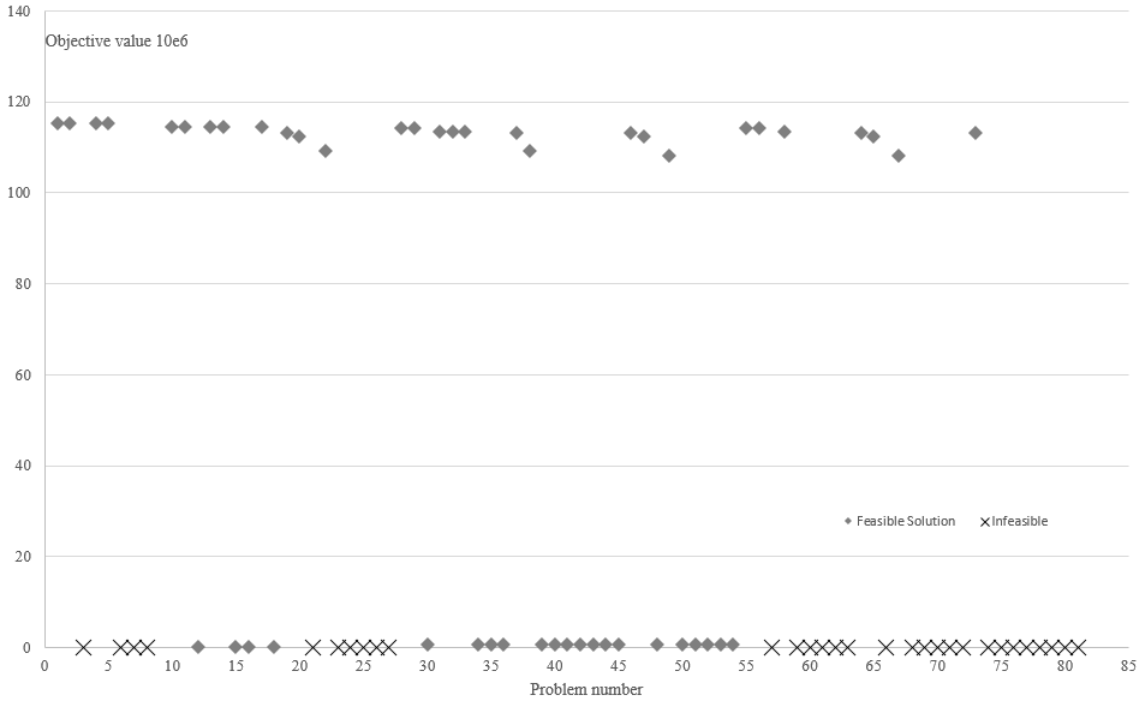


Figure 5.14: Objective value of each subproblem without inter-job communication

As shown in Figure 5.15, the total execution time of the parallelisation is dependent on the most time-consuming job, which is job No.2 in this case and its CPU time is 1145 seconds. Again, although it outperforms the partition approach presented in section 5.4.2 where several subproblems do not solve, it is still not better than the original problem. However, several subproblems do provide good incumbent

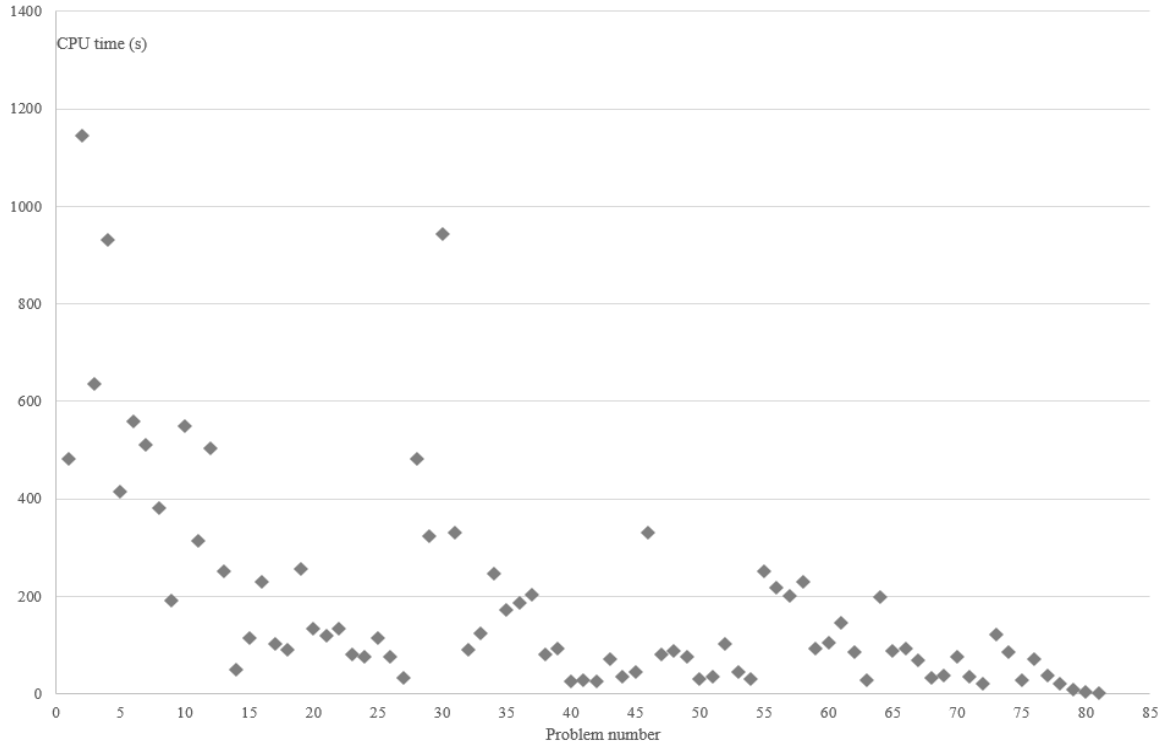


Figure 5.15: CPU time (s) of each subproblem without inter-job communication

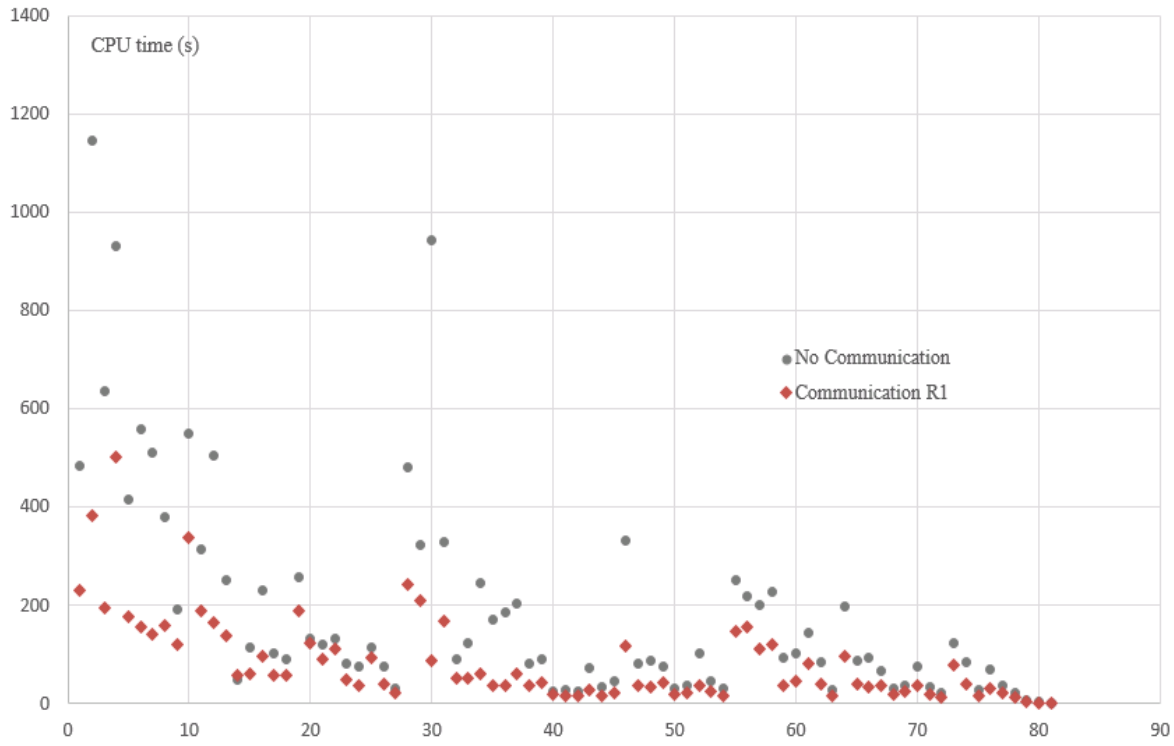


Figure 5.16: Comparison of CPU time (s) between cases with and without inter-job communication

values at a very early stage (e.g. problem No. 14 provides a rather good incumbent of 114.47 at time = 50 s, which is very close to the final solution and problem No.5 is solved within 415 seconds, and its solution turns out to be the best object value). Therefore, inter-job communication is implemented again to test its capability in reducing the solution time of the original problem provided and this time, with

good early incumbent values. The comparison is given in Figure 5.16. It is clearly shown that almost every single problem reduces its CPU time when inter-job communication is introduced (red square). The overall problem solves within 514 seconds with the same optimal solution of 115.2. The massive reduction in solution time is mainly because of the fact that the solution time of the previous difficult problem No.2 is significantly reduced from 1145s to 501s. To validate the method, five more runs with the same problem set-up and computer are conducted, and the detailed model statistics of each run are summarised in Figure B 1 in Appendix B.4. The consistent results prove the capability of the method to significantly reduce the CPU time (in this case, from 901s of the original problem to 514s when the selected variables are partitioned into three feasible spaces in t with inter-job communication) when good and early feasible solutions are combined with parallel computation.

5.5. Concluding Remarks

In this chapter, the combined methodology introduced in chapter III and IV has been modified to accommodate the separation between polymers and unreacted monomers using a conceptual dissolution-precipitation-filtration system. The original problem is then reformulated using a feasible region partition method which essentially breaks the original problem into an equivalent set of subproblems. It is done by dividing the feasible space in the temperature (t) index of the binary variable ($\hat{y}_{k,t}$) into three parts. The reformulation aims to have a better approximated discretised MILP problem by further reducing the size of each temperature interval without causing significant issues to the solvability of the model (i.e. much longer CPU time). The concepts of discretisation of continuous variables, the branch and bound process, and feasible region are first introduced along with the problem statement. The feasible region partition method is then discussed, and the potential benefit of communication between jobs is illustrated. A case study is solved within 901s using GAMS to provide the optimal objective value, reaction network, process diagram and heat. The variable deemed to be the most suitable variable to be partitioned is the temperature point variable of each reaction stage ($\hat{y}_{k,t}$) which affects the temperatures associated in the entire master stage k. It is first partitioned into two equal ranges in the t space, and the equivalent subproblems are solved in parallel without communication. The model times out in 25 minutes (1500 s), which is worse than the original problem. Communication is later added, and the overall execution time of the model is reduced from 901s to 474s. However, the performance of the method when there are only 16 subproblems and without a good early feasible incumbent is not consistent (i.e. the problem's execution time can also increase the time to more than 901s). A second partition mechanism is then implemented, which divides the t space into three parts and forms 81 subproblems. The overall execution time of the parallelisation without inter-job communication is around 1145s, which is still not better than the solution time of the original problem. However, upon activation of inter-job communication, the CPU time drops significantly from 901s to 514s and the method is proven to be stable (i.e. CPU time of each subproblem does not vary

much from run to run) when there are enough subproblems and a good early feasible solution is available.

In summary, it is shown that feasible region partition method alone does not guarantee improvements in solution time. However, when combined with an early feasible solution and inter-job communication, the solution time of complex MILP problems with discretised continues variables can be significantly reduced.

6. Chapter VI: Conclusions and Future Work

This thesis has addressed several problems related to process synthesis - focussing mainly on recently developed reaction pathways of bio-based polymers and their monomers, including optimal design of the process flowsheet with simultaneous heat integration and a decomposition method to improve the computational efficiency of the solution, to fill the gap in the literature work.

In this chapter, the aim is to conclude the work presented in this thesis and provide the potential research directions for the future work.

6.1. Concluding Remarks

In this thesis, mixed integer linear programming based models and solution approaches have been proposed for process synthesis with simultaneous heat integration of several literature examples of reaction networks concerning the synthesis of newly developed bio-based polymers or their monomers.

In Chapter I, a general introduction has been given first related to the global polymers market and its trend and then focussing on the current issues with the exploitation of fossil resources followed by the development of bio-based polymers and its market capacity summary. Finally, background information on process synthesis and the use of optimisation-based mathematical programming approaches used in process synthesis is also introduced.

In Chapter II, a detailed literature review has been presented related to the synthesis of innovative bio-based polymers, process synthesis for individual units and optimisation-based mathematical programming frameworks for the synthesis of an entire process including MILP and MINLP models. A section on well-studied decomposition methods, which is generally used to increase the computational efficiency for the solution of optimisation problems, has also been provided.

In Chapter III, the optimisation-based approach has been presented for the synthesis of the optimal reaction pathways for a given reaction network along with the corresponding process flowsheet including reactor/distillation sequences.

This chapter consists of a methodology section 3.2 where the mathematical formulation is presented, and a case study section 3.3 where two case studies are solved to demonstrate the capability of the model. In section 3.2, an optimisation-based MILP framework has been developed along with a process superstructure for the design of the optimal reaction network integrated with separation sequencing, with an objective to maximise the overall process economic potential. The use of bi-linear constraints for the split fraction at each splitter reduces the model's complexity and at the same time increases its flexibility. The installation of recycles allows the reuse of unreacted chemicals and hence significantly reduces the operational cost of the process. Another innovative mechanism for the design of separation sequences is also implemented to allow the model to identify the components that enter each distillation

stage and provide constraints for the system to logically select both light and heavy keys. In section 3.3, the developed model has been applied to solve the synthesis problem of two case studies, which consist of two different reaction networks obtained from the literature. The two cases cover two popular bio-based chemicals that can be used to produce value-added platform chemicals including monomers, i.e. limonene and itaconic acid. The model results provide insight into the process configurations of the optimal solutions.

Chapter IV presents an extended MILP model for the simultaneous process synthesis with automatic heat integration which has been developed based on the mathematical programming approaches in Chapter III. A modified process superstructure, which includes heat exchangers, reboilers and condensers, is presented. New sets of constraints are incorporated to track and evaluate the energy change of each stream. The linear nature of the model is enforced through the discretisation of the continuous variable-temperature, which avoids the potential non-convexity induced by nonlinear formulations. The model has been applied to the limonene case study given in Chapter III to examine the effect of heat integration on the energy consumption while screening for the most promising bio-based chemical and its production route from alternative candidate products and reaction pathways. The results imply that significant energy saving can be obtained from simultaneous heat integration. The followed up sensitivity analysis validates the assumptions made in the model.

Finally, in Chapter V, the combined model from Chapter III and IV is further adapted to accommodate the separation between polymers and unreacted monomers using a conceptual dissolution-precipitation-filtration system. This framework is then subject to a customised decomposition method to increase the computational efficiency for the solution of the resulting problem with a large number of discretised variables. The decomposition method includes the breakdown of the original problem into an equivalent set of subproblems by the segmentation of choice variables in the discretised temperature space and the implementation of inter-job communication to exchange their best bounds. The framework is applied to a case study which involves a reaction network starting from limonene to produce various types of polymers. The results indicate that significant improvement in the computational efficiency is achieved when inter-job communication is integrated with the decomposition method.

From work presented in this thesis, the mathematical programming techniques, mainly MILP optimisation techniques, can be widely applied to process synthesis problems. The proposed MILP approaches have successfully dealt with the process synthesis problems discussed in this thesis. The work in this thesis has developed a novel approach to tackle literature problems that are in their early stage of development and aims to provide insight knowledge to researchers from an engineering perspective. In addition, the decomposition method proposed manages to lower the increase in solution CPU time in a particular set of MILP problems, where one or more continuous variables are discretised, caused by the increase in the number of integer variables to satisfy the requirement of solution quality.

This thesis a complement to the literature research work on optimisation-based process synthesis and two publications have arisen from work presented in this thesis which can be seen in Appendix D.

6.2. Future Work

The work in this thesis has covered the optimisation-based framework for the process synthesis of newly developed bio-based reaction pathways for the production of polymers, and there are still several research directions that could be deployed for future work as the extension of the current study.

Regarding the combined MILP model developed in Chapter III and IV, as mentioned in the introduction, one limitation of the developed model is the assumption of no azeotrope formation, which is generally not the case in most industrial practices. Therefore, three potential modification approaches are proposed with the first two focusing on the model and the third on the analysis of the solution. The first would be introducing a penalty cost term in the objective function for separating pairs of components that form azeotropes in the same distillation column as the light and heavy key respectively. Secondly, constraints can be imposed to disallow any close boiling-points separation to occur (i.e. $Bp_i - Bp_{i'} > T^{min}$). Finally, another option is to employ solution analysis techniques such as integer cut to generate the next best solution only if azeotropic distillation is present in the optimal solution.

Since the model focuses primarily on analysing new lab-scale reaction routes, one possible future direction could be to consider process uncertainty arise from parameters such as reaction characteristics, material prices and utility costs to reduce the risk when scaling up the optimal process. This would mean that a stochastic programming model would be required. On the other hand, if more data of the reactions being studied becomes available, a more accurate optimisation model can be built by replacing simple reaction and separation units with more complex processing blocks such as simulation-based and surrogate models. This allows the model to deal with more complicated process flowsheets, which could include decision-making processes regarding the selection of biomass, biomass processing technologies (e.g. pyrolysis, steam explosion and hydrolysis) and different types of reactors. Furthermore, the modelling of heat integration is quite basic without any analysis on the stream matching and the optimisation of heat transfer area (i.e. cost of heat exchangers). Therefore, another extension to the model could be the optimisation of the HENS either in sequential order or simultaneously with the process synthesis model.

Regarding the mathematical formulation, several simplifications are used to avoid nonlinearity (e.g. the use of binary split fraction and simplified separation model). To further validate the model, an equivalent MINLP model should be used to examine the effect of those simplifications and how the linear model performs in terms of accuracy and CPU time.

Regarding Chapter V, it is clearly shown that the partitioning approach can significantly reduce the CPU time of the proposed models of Chapter III and IV, but it would also be worth comparing the

performance of the method against other decomposition methods in the literature. In addition, a set of more general rules is required for users to determine the applicability of the partition method to their problems and how to best select the variables to be partitioned to improve the performance. Moreover, the author intends to develop a simple-to-use package and an interface to allow easier access to the partition program, where model developers would be able to plug their models and solve the problems with much improved computational efficiency. In this way, users could examine the effect of changing the decomposition variables by simply altering a few parameters which are related to the decomposition decisions.

Appendix A Supplementary Information for Chapter III

In this appendix, supplementary information for Chapter III is provided.

A.1 Bi-linear Equation & Linearisation

The linearisation of the equations of the separation splits streams are given in this section

$$\zeta_{j,k} \geq \frac{x_j \alpha_j y_{j,k} Q_{i,k}^{in}}{-v_{i,j}}, \forall j,k,i = LR_j, v_{i,j} < 0 \quad \text{A. 1}$$

$$f_{i,k,l}^{TR} = \sum_p dsf_p \, D4_{i,k,l,p}, \forall k,i,l \quad \text{A. 2}$$

$$D4_{i,k,l,p} \leq f^{max} \lambda_{k,l,p}^{TR}, \forall k,i,l \quad \text{A. 3}$$

$$f_{i,k,l}^{T} + f^{max}\left(\lambda_{k,l,p}^{TR} - 1\right) - D4_{i,k,l,p} \leq 0, \forall k,i,l \quad \text{A. 4}$$

$$D4_{i,k,l,p} - f_{i,k,l}^{T} \leq 0, \forall k,i,l \quad \text{A. 5}$$

$$f_{i,k,l}^{Tout} = \sum_p dsf_p \, D5_{i,k,l,p}, \forall k,i,l \quad \text{A. 6}$$

$$D5_{i,k,l,p} \leq f^{max} \lambda_{k,l,p}^{Tout}, \forall k,i,l,p \quad \text{A. 7}$$

$$f_{i,k,l}^{T} + f^{max}\left(\lambda_{k,l,p}^{Tout} - 1\right) - D5_{i,k,l,p} \leq 0, \forall k,i,l,p \quad \text{A. 8}$$

$$D5_{i,k,l,p} - f_{i,k,l}^{T} \leq 0, \forall k,i,l,p \quad \text{A. 9}$$

$$f_{i,k,l}^{BR} = \sum_p dsf_p \, D7_{i,k,l,p}, \forall k,i,l \quad \text{A. 10}$$

$$D7_{i,k,l,p} \leq f^{max} \lambda_{k,l,p}^{BR}, \forall k,i,l,p \quad \text{A. 11}$$

$$f_{i,k,l}^{B} + f^{max}\left(\lambda_{k,l,p}^{BR} - 1\right) - D7_{i,k,l,p} \leq 0, \forall k,i,l,p \quad \text{A. 12}$$

$$D7_{i,k,l,p} - f_{i,k,l}^{B} \leq 0, \forall k,i,l,p \quad \text{A. 13}$$

$$f_{i,k,l}^{Bout} = \sum_p dsf_p \, D8_{i,k,l,p}, \forall k,i,l \quad \text{A. 14}$$

$$D8_{i,k,l,p} \leq f^{max} \lambda_{k,l,p}^{Bout}, \forall k,i,l,p \quad \text{A. 15}$$

$$f_{i,k,l}^{B} + f^{max}\left(\lambda_{k,l,p}^{Bout} - 1\right) - D8_{i,k,l,p} \leq 0, \forall k,i,l,p \quad \text{A. 16}$$

$$D8_{i,k,l,p} - f_{i,k,l}^{B} \leq 0, \forall k,i,l,p \quad \text{A. 17}$$

A.2 Direct Removal Constraints for the Separation Bypass

The direct removal mechanism for the separation outlets is given in this section.

$$N_{i,k,l}^{T} \leq f_{i,k,l}^{T}, \forall k, i, l \quad \text{A. 18}$$

$$N_{i,k,l}^{T} M \geq f_{i,k,l}^{T}, \forall k, i, l \quad \text{A. 19}$$

$$N_{i,k,l}^{B} \leq f_{i,k,l}^{B}, \forall k, i, l \quad \text{A. 20}$$

$$N_{i,k,l}^{B} M \geq f_{i,k,l}^{B}, \forall k, i, l \quad \text{A. 21}$$

$$V_{i,k,l}^{TR} \leq mc_i f_{i,k,l}^{TR}, \forall k, i, l \quad \text{A. 22}$$

$$V_{i,k,l}^{BR} \leq mc_i f_{i,k,l}^{BR}, \forall k, i, l \quad \text{A. 23}$$

$$V_{i,k,l}^{TR} \leq mc_i f^{max} EX_{i,k,l}^{T}, \forall k, i, l \quad \text{A. 24}$$

$$V_{i,k,l}^{BR} \leq mc_i f^{max} EX_{i,k,l}^{B}, \forall k, i, l \quad \text{A. 25}$$

$$EX_{i,k,l}^{T} \leq N_{i,k,l}^{T}, \forall k, i, l \quad \text{A. 26}$$

$$EX_{i,k,l}^{B} \leq N_{i,k,l}^{B}, \forall k, i, l \quad \text{A. 27}$$

$$EX_{i,k,l}^{T} \leq 1 - N_{i',k,l}^{T}, \forall i' \neq i, k, l \quad \text{A. 28}$$

$$EX_{i,k,l}^{B} \leq 1 - N_{i',k,l}^{B}, \forall i' \neq i, k, l \quad \text{A. 29}$$

$$VRtotal_{i,k} = \sum_{l} V_{i,k,l}^{TR} + V_{i,k,l}^{BR}, \forall k, i \quad \text{A. 30}$$

A.3 Separation Coefficient Estimation

The first example used to estimate the separation coefficient is the benzene/toluene distillation system shown in Figure A 1. The inlet flowrate of the distillation column is 100 mol/min with 50% of each benzene and toluene. To analyse the capital cost of installing the column, it is essential to firstly calculate the column height which is related to the number of separation stages required for the system. The relative volatility α_{ij} between benzene and toluene is 2.25 and the minimum number of stages that is required to achieve 99% recovery of benzene at the distillate and 99% recovery of toluene at the bottom is calculated by the Fenske's equation:

$$N_{min} = \frac{\ln\left(\frac{r_{i,D}}{1-r_{i,D}} \cdot \frac{r_{j,B}}{1-r_{j,B}}\right)}{\ln\alpha_{ij}} \quad \text{A. 31}$$

The minimum reflux ratio is secondly calculated by using the simplified version of Underwood's 1st and 2nd equations:

$$\sum_{i=1}^{C} \alpha_i \cdot \frac{z_{i,F}}{\alpha_i - \emptyset} = 1 - q = 0 \quad \text{A. 32}$$

$$\sum_{i=1}^{C} \alpha_i \cdot \frac{x_{i,D}}{\alpha_i - \emptyset} = R_{min} + 1 = 0 \quad \text{A. 33}$$

The actual reflux ratio is commonly known to be equal to$1.1R_{min}$.

The actual number of separation stage N thus can be calculated by Gilliland correlation:

$$\frac{N - N_{min}}{N+1} = 0.75 \cdot \left(1 - \left(\frac{R - R_{min}}{R+1}\right)^{0.5688}\right) \quad \text{A. 34}$$

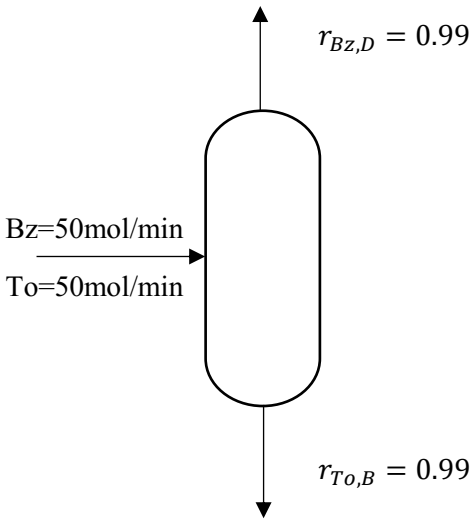


Figure A 1: Bz/To distillation system schematic

On the other hand, the feed location can also be determined via Fenske's equation for the rectifying section, i.e.

$$N_{min}^{Top} = \frac{\ln\left(\left(\frac{x_{i,D}}{x_{j,D}}\right)\cdot\left(\frac{z_{j,F}}{z_{i,F}}\right)\right)}{\ln\alpha_{ij}} \quad \text{A. 35}$$

The analysis below is solely based on the methods developed by Douglas (Douglas 1988).

The height H (in ft) of the distillation column accounting for extra space at the ends of the tower can be calculated as follows:

$$H = 2.3N \quad \text{A. 36}$$

Knowing the estimated column height, the diameter of the column can be calculated given the maximum vapour molar flowrate V in the column:

$$D = 0.0164\sqrt{V}\left(\frac{M_G}{\rho_m}\right)^{\frac{1}{4}} \quad \text{A. 37}$$

where M_G is the molecular weight of the vapour, ρ_m is the molar density of the vapour and the vapour molar flowrate is given by the following equation:

$$V = (R+1)Distillate \quad \text{A. 38}$$

The column installation cost can be thus be calculated as follows:

$$Installed\ Cost\ (\$) = \frac{(M\&S)}{280} \times 101.9D^{1.066}H^{0.802}(2.18 + F_C) \quad \text{A. 39}$$

$$\mathrm{F_C} = \mathrm{F_m F_p} \quad \text{A. 40}$$

and the tray installation cost can be calculated as follows:

$$Installed\ Cost\ (\$) = \frac{(M\&S)}{280} \times 4.8D^{1.55}HF_C \quad \text{A. 41}$$

where F_C is the sum of the material correction factor (F_m), spacing correction factor (F_s) and tray type correction factor (F_t), i.e.

$$F_C = F_S + F_t + F_m \quad \text{A. 42}$$

The installation cost of condenser and reboiler of the column is written as:

$$\text{Installation Cost}\ (\$) = \left(\frac{\mathrm{M\&S}}{280}\right)(101.3)(2.29 + \mathrm{F_C})\left(\mathrm{A_C^{0.65}} + \mathrm{A_R^{0.65}}\right) \quad \text{A. 43}$$

where A_C and A_R are the heat-transfer area of condenser and reboiler respectively and F_C is the correction factor for the condenser and reboiler:

$$\mathrm{A_C} = \left(\left(\frac{\mathrm{Q_C}}{27 * \mathrm{U_C}}\right)\ln\left(\frac{\mathrm{T_b} - 68}{\mathrm{T_b} - 77}\right)\right) \quad \text{A. 44}$$

$$A_R = \left(\frac{Q_R}{U_R \Delta T_M}\right) \quad \text{A. 45}$$

$$F_c = (F_d + F_p)F_m \quad \text{A. 46}$$

For the Bz/To case given in Figure A 1, the value of correction factors and the important parameters that were used for the calculation are summarised in Table A 1 and Table A 2 respectively. The calculated results of the column specification and installation cost are shown in Table A 3. Annualised capital costs of different parts are calculated by dividing the installation cost by 8, which is an estimated value of the plant lifetime in years and the results are shown in Table A 4.

Table A 1: Correction factors summary

F_c			
Column	Tray	Condenser	Reboiler
3.67	2.7	2.39	3.79

Table A 2: Important parameters used in the calculation

	Physical meaning	**Values**	**Unit**	**Rationale**
M&S	Marshall and Swift Code	1352	-	Online source
$\boldsymbol{M_G}$	Molecular weight of benzene	0.172	lb/mol	molecular weight of benzene
$\boldsymbol{\rho_m}$	Molar density of benzene	0.99	Mol/ft^3	NIST
$\boldsymbol{F}$	Feed flowrate	100	mol/min	User defined value
-	Water cost	0.0015	$/gallon	Online source
-	Steam cost	0.0021	$/lb	Online source
$\boldsymbol{T_b}$	Bubble point of benzene	176.2	F	Boling point of benzene
$\boldsymbol{U_c}$	Condenser heat transfer Coefficient	100	Btu/(hr.ft2.F)	Douglas
$\boldsymbol{U_R \Delta T_M}$	Reboiler heat flux	11250	Btu/(hr.ft^2)	Douglas

Table A 3: Results summary

Variables	Unit	Value
N_{min}	-	11.3
Φ	-	1.38
R_{min}	-	1.55
R	-	1.70
N	-	30.0
N^{top}_{min}	-	5.67
N^{top}	-	15.0
H	ft	69.1
Vapour flowrate	mol/hr	8030
D	ft	1.41
Column Install	$	124,000
Tray Install	$	7,200
Qc	BTU/hr	236,820
Qr	BTU/hr	305,133
Ac	ft^2	7.62
Ar	ft^2	27.1
Condenser install	$	8,600
Reboiler install	$	25,400

Table A 4: Annualised cost of the separation column

Column cost ($/yr)	Tray cost ($/yr)	Condenser cost ($/yr)	Reboiler cost ($/yr)	Utility cost($/yr)	TAC($/yr)
15,500	900	1,070	3,200	46,100	66,770

This TAC is then used to estimate the separation cost coefficient σ via the following equation:

$$TAC\ (\$) = \sigma \frac{\sum_i f_i^{in}}{Bp_{To} - Bp_{Bz}} \qquad \text{A. 47}$$

Therefore, $\sigma \approx 339\ \$ \cdot K \frac{hr}{mol}$. The procedure was repeated to obtain the separation cost coefficient for different inlet flowrate with same composition and the results are shown in Table A 5.

Table A 5: Value of σ for the Bz/To case

Inlet flowrate (mol/min)	100	200	300
$\boldsymbol{\sigma}$ ($\$ \cdot K \frac{\boldsymbol{hr}}{\boldsymbol{mol}}$)	339	311	298
Avearge value of σ	316		

The same method was applied on methanol/ethanol and acetone/propyl acetate systems. The relative volatility, boiling point difference, average value of separation cost coefficient of each system is shown in Table A 6.

Table A 6: Separation cost coefficient summary of the three cases

System	Relative volatility	$\Delta BP(K)$	$\sigma(\$ \cdot K\frac{hr}{mol})$
Meol/ethol	1.56	13.7	486
Bz/To	2.25	30.5	316
Act/pacetate	4.74	46.0	146

In the paper it was assumed that all the distillation systems in the case studies have the similar relative volatility as the benzene/toluene system and therefore the separation cost coefficient was taken to be $\underline{316}\ \$ \cdot K\frac{hr}{mol}$.

Appendix B Supplementary Information for Chapter V

In this appendix, supplementary information for chapter V is provided.

B.1 Properties of Reactions and Components involved in the Case Study

In this section, the properties of reactions and components involved in the case study of Chapter V are summarised and given in Table B 1, Table B 2 and Table B 3.

Table B 1: Reaction conditions for all the reactions involved in the case study

Reaction Number	Reaction Temperature (°C)	Standard Enthalpy Change of Reaction (kJ/mol)
R1	85	-88
R2	80	-153
R3	80	-473
R4	75	-23
R5	75	-186
R6	25	-245
R7	130	-58.7
R8	60	-59
R9	170	3.54
R10	80	-718.7
R11	95	-17.48
R12	280	371.54
R13	245	356.92

Table B 2: Thermodynamic properties of the components involved in the case study

Component Symbol	Boiling Point (°C)	Liquid (Solid for Polymers) Heat Capacity (J/mol)	Gas Heat Capacity (J/mol)	Latent Heat (kJ/mol)
AS1	65	80	65	37.6
F3	101	191	193	40.1
S1	110	157	170	37
F2	145	182	151	43.9
F1	176	300	88	43.2
C3	177	307	74	49
F7	197	150	97	65
C1	198	230	92	38.5
F5	202	162	90	90
C2	231	320	85.3	55.5
F6	235	200	173	78
C5	288	392	73.6	80.1
F4	295	220	226	52.1
C4	320	362	55.9	113.9
P1	-	240	-	-
P2	-	386	-	-
P3	-	393	-	-
P4	-	184	-	-
P5	-	360	-	-
P6	-	314	-	-

P7	-	368	-	-
P8	-	328	-	-

Table B 3: Market prices of components involved in the case study

Component Symbol	Market Price ($/mol)	Market Price ($/g)
AS1	0.0032	0.0001
F3	0.23	0.0023
S1	0.0552	0.0006
F2	0.138	0.0013
F1	0.476	0.0035
C3	0.537	0.004
F7	0.031	0.00034
C1	1.52	0.01
F5	0.0784	0.0008
C2	0.51	0.0034
F6	0.0095	0.00015
C5	0.194	0.001
F4	0.1776	0.0012
C4	0.1577	0.00095
P1	0.595	0.0044
P2	0.7675	0.006
P3	0.8825	0.007
P4	1.9	0.0125
P5	2.12	0.014
P6	1.998	0.0135
P7	0.254	0.0014
P8	0.28125	0.0017

B.2 Equations for the Polymer Separation Unit

The material balances of the polymer separation unit and the new objective function, which accounts for the economic value of the polymer produced, are given in this section.

Nomenclature

Symbol involved in all the additional equations are defined as follows:

Sets	Description
So	Component set for solvent
Aso	Component set for antisolvent
Ch	Component set for chemicals involved in reactions (non-polymer)
Pol	Polymer set

Parameters

μ_1	Amount of solvent required to dissolve a unit of polymer
μ_2	Amount of antisolvent required to precipitate a unit of polymer

Binary Variables

b_k^{bypass}	$= 1$ if reaction outlet of stage k bypasses the polymer separation unit; 0 otherwise
$b_k^{polysep}$	$= 1$ if reaction outlet of stage k enters the polymer separation unit; 0 otherwise
$N_{i,k}^{Qout}$	$= 1$ if component i exists in the reaction outlet stream of stage k; 0 otherwise

Continuous Variables

$Q_{i,k}^{bypass}$	Flowrate of component i at stage k that bypassing the polymer separation unit
$Q_{i,k}^{polysep}$	Flowrate of component i at stage k that enters the polymer separation unit
$Q_{So,Pol,k}^{solvent}$	Flowrate of solvent So corresponding to polymer Pol that enters the dissolution tank of stage k
$Q_{i,k}^{solution}$	Flowrate of component i that leaves the dissolution tank of stage k
$Q_{Aso,Pol,k}^{Antisol}$	Flowrate of antisolvent Aso corresponding to polymer Pol that enters the precipitation tank of stage k
$Q_{i,k}^{precip}$	Flowrate of component i leaving the precipitation tank of stage k
$Q_{i,k}^{filterB}$	Flowrate of component i that leaves the filtration unit of stage k as the solid
$Q_{i,k}^{filterT}$	Flowrate of component i that leaves the filtration unit of stage k as the liquid
$Q_{i,k}^{outnew}$	New reaction outlet stream that contains no polymer
$Q_{i,k}^{polymer}$	The production rate of polymer i at stage k
PP	Economic value of polymer of the entire process
SV	Economic value of the solvent used
ASV	Economic value of the antisolvent used

Additional Constraints

The additional mass balances for the polymer separation unit are given as follows:

$$Q_{i,k}^{out} = Q_{i,k}^{bypass} + Q_{i,k}^{polysep}, \forall i,k \quad \text{B. 1}$$

$$N_{i,k}^{Qout} \leq Q_{i,k}^{out}, \forall i,k \quad \text{B. 2}$$

$$N_{i,k}^{Qout} M \geq Q_{i,k}^{out}, \forall i,k \quad \text{B. 3}$$

$$b_k^{polysep} \geq N_{i,k}^{Qout}, \forall i \in Pol, k \quad \text{B. 4}$$

$$b_k^{bypass} + b_k^{polysep} \leq 1, \forall k \quad \text{B. 5}$$

$$b_k^{polysep} \leq \sum_{i \in Pol} N_{i,k}^{Qout}, \forall k \quad \text{B. 6}$$

$$Q_{i,k}^{bypass} \leq b_k^{bypass} Q^{max}, \forall i,k \quad \text{B. 7}$$

$$Q_{i,k}^{polysep} \leq b_k^{polysep} Q^{max}, \forall i,k \quad \text{B. 8}$$

$$Q_{i,i',k}^{solvent} = \mu_1 Q_{i',k}^{polysep}, \forall i \in So, i' \in Pol, k \quad \text{B. 9}$$

$$Q_{i,k}^{solution} = \sum_{i' \in Pol} Q_{i,i',k}^{solvent}, \forall i \in So, k \quad \text{B. 10}$$

$$Q_{i,k}^{solution} = Q_{i,k}^{polysep}, \forall i \in Ch, k \quad \text{B. 11}$$

$$Q_{i,k}^{solution} = Q_{i,k}^{polysep}, \forall i \in Pol, k \quad \text{B. 12}$$

$$Q_{i,i',k}^{antisol} = \mu_2 Q_{i',k}^{solution}, \forall i \in Aso, i' \in Pol, k \quad \text{B. 13}$$

$$Q_{i,k}^{precip} = Q_{i,k}^{solution}, \forall i \neq Aso, k \quad \text{B. 14}$$

$$Q_{i,k}^{precip} = \sum_{i' \in Pol} Q_{i,i',k}^{antisol}, \forall i \in Aso, k \quad \text{B. 15}$$

$$Q_{i,k}^{filterT} + Q_{i,k}^{filterB} = Q_{i,k}^{precip}, \forall i,k \quad \text{B. 16}$$

$$Q_{i,k}^{filterB} = Q_{i,k}^{precip}, \forall i \in Pol, k \quad \text{B. 17}$$

$$Q_{i,k}^{filterT} = Q_{i,k}^{precip}, \forall i \neq Pol, k \quad \text{B. 18}$$

$$Q_{i,k}^{polymer} = Q_{i,k}^{filterB}, \forall i,k \quad \text{B. 19}$$

$$Q_{i,k}^{outnew} = Q_{i,k}^{bypass} + Q_{i,k}^{filterT}, \forall\, i,k \quad \text{B. 20}$$

Objective Function

The modified objective function is given as follows:

$$Z = maxEP4 \quad \text{B. 21}$$

$$EP4 = EP3 + PP - SV - ASV \quad \text{B. 22}$$

$$PP = \sum_{i,k} mc_i\, Q_{i,k}^{polymer} WH \quad \text{B. 23}$$

$$SV = \sum_{i \in So,k} mc_i\, Q_{i,k}^{precip} WH \quad \text{B. 24}$$

$$SV = \sum_{i \in Aso,k} mc_i\, Q_{i,k}^{precip} WH \quad \text{B. 25}$$

B.3 Validation Runs for the 16-Subproblem Case

The CPU times of the sample runs with inter-job communication for the 16-subproblem case are given in this section.

Table B 4: CPU times of sample runs with inter-job communication (16 subproblems)

Subproblems/ CPU time (s)	Run 1	Run 2	Run 3	Run 4	Run 5
1	436	548	714	1464	899
2	450	552	676	1379	702
3	381	485	665	931	1141
4	353	490	666	890	671
5	344	493	782	883	682
6	341	512	472	636	690
7	354	425	464	632	518
8	139	161	116	126	157
9	441	547	640	TO	706
10	327	439	636	937	665
11	323	427	524	857	620
12	139	135	105	152	132
13	315	408	623	856	644
14	131	149	111	138	155
15	90	101	74	100	101
16	39	45	34	34	48

B.4 Validation Runs for the 81-Subproblem Case

The CPU times of the sample runs with inter-job communication for the 81-subproblem case are given in this section.

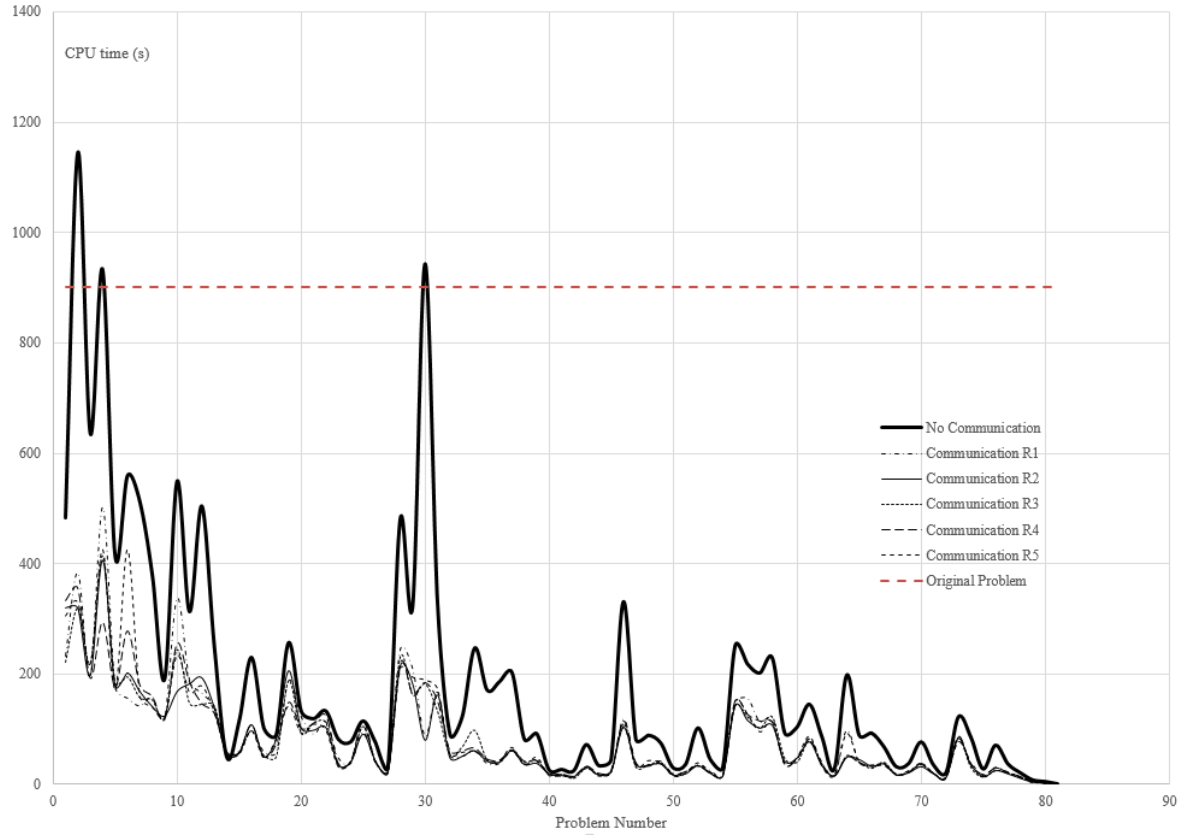


Figure B 1: CPU times of Sample runs with inter-job communication (81 subproblems)

Appendix C GAMS Implementation

```
$set cutoff 'cutlo'
$setglobal ide "ide=%gams.ide% lo=%gams.lo% errorlog=%gams.errorlog% errmsg=1"
$call rm -f bch*.gdx t*.trg incb.gdx incb*.txt allincb.txt GMSbch*.* cplex.*
**------------------Define the sets to be partitioned---------------------
 set
        ss1 /1*3/
        ss2 /1*3/
        ss3 /1*3/
        ss4 /1*3/
        dpart213(ss1,k,t) /1.k1.(1*54),  2.k1.(55*108), 3.k1.(109*160)/
        dpart214(ss2,k,t) /1.k2.(1*54),  2.k2.(55*108), 3.k2.(109*160)/
        dpart215(ss3,k,t) /1.k3.(1*54),  2.k3.(55*108), 3.k3.(109*160)/
        dpart216(ss4,k,t) /1.k4.(1*54),  2.k4.(55*108), 3.k4.(109*160)/
        dpartD(ss1,ss2,ss3,ss4,k,t)
**------------------Define the Parameters---------------------

Parameter
        h(ss1,ss2,ss3,ss4) model handles
        jobssubmitted actual number of submitted jobs
        report(ss1,ss2,ss3,ss4,*)
        ;

Scalar cnt, incumbent
        bestcompleted the best solution of all completed jobs;

        incumbent = -inf;
        bestcompleted = incumbent;

**------------------Set up initial solver's file for each subproblems---------------------
```

```
file fx;


        for (cnt=1 to 81,
        put_utility fx 'ren' / 'cplex.' (100+cnt):0:0;
        put   'userincbicall writeincb.gms lo=0 %ide%'
        / 'userjobid ' cnt:0:0 / 'usergdxprefix %gams.workdir%'
        / 'interactive yes' / 'iafile "%gams.workdir%cplex.op3"' / 'iatriggertime 5'
        / 'iatriggerfile "%gams.workdir%t' cnt:0:0 '.trg"'
        / 'mipemphasis 0'
        / 'indic SurplusRe_A_Cal1(k,t)$ykts(k,t) 1' / 'indic SurplusRe_A_Cal0(k,t)$ykts(k,t) 0'/
          'indic SurplusRe_Cal1(k,t)$EXO(k) 1'/'indic SurplusRe_Cal0(k,t)$EXO(k) 0'/
          'indic FCp_A_Calculation1(k,t)$EofA(k,t) 1'/'indic FCp_A_Calculation0(k,t)$EofA(k,t) 0'/
        'indic Surpluscalc_Dummy1RB(k,l,t)$yRkts(k,l,t) 1'/'indic
Surpluscalc_Dummy0RB(k,l,t)$yRkts(k,l,t) 0'/
        'indic Surpluscalc_Dummy1CD(k,l,t)$yCkts(k,l,t) 1'/'indic
Surpluscalc_Dummy0CD(k,l,t)$yCkts(k,l,t) 0'/
        'indic fTR_FCp_Cal1(k,l,t)$EoffTR(k,l,t) 1'/'indic fTR_FCp_Cal0(k,l,t)$EoffTR(k,l,t) 0'/
        'indic fBR_FCp_Cal1(k,l,t)$EoffBR(k,l,t) 1'/'indic fBR_FCp_Cal0(k,l,t)$EoffBR(k,l,t) 0'/
        'indic SurplusTout_Total_Cal1(k,l,t)$HOTTout(k,l) 1'/'indic
SurplusTout_Total_Cal0(k,l,t)$HOTTout(k,l) 0'/
        'indic SurplusBout_Total_Cal1(k,l,t)$HOTBout(k,l) 1'/'indic
SurplusBout_Total_Cal0(k,l,t)$HOTBout(k,l) 0'/
        'indic SurplusQsepin_Total_Cal1(k,t)$HOTQsepin(k) 1'/'indic
SurplusQsepin_Total_Cal0(k,t)$HOTQsepin(k) 0'/
        'indic FCpQR_Cal1(k,t)$EofQR(k,t) 1'/'indic FCpQR_Cal0(k,t)$EofQR(k,t) 0'/
        'indic SurplusQnext_Total_Cal1(k,t)$HOTQnext(k) 1'/'indic
SurplusQnext_Total_Cal0(k,t)$HOTQnext(k) 0'/
        'preslvnd 0'/'epgap 0.001'/'epopt 1e-009'/'eprhs 1e-009');
       putclose fx;
```

*** ----------Save the overall incumbent in file incb0.txt. This file gets updated----------**

*** ----------by checkincb.gms if we find an imporved incumbent----------**

```
       put_utility fx 'ren' / 'incb0.txt'; putclose fx '0 ' incumbent:18:8;
```

----------Loop over and solve all subproblems in parallel----------

```
profit.solvelink=%solvelink.AsyncGrid%;

loop((ss1,ss2,ss3,ss4),

Option Clear=ykts, Clear=yAkte, Clear=yQsepinkts, Clear=yQnextkts, Clear=yQRkts;

dpartD(ss1,ss2,ss3,ss4,k,t) = dpart213(ss1,k,t) + dpart214(ss2,k,t)+dpart215(ss3,k,t)+
dpart216(ss4,k,t);
ykts.fx(k,t)$(not dpartD(ss1,ss2,ss3,ss4,k,t)) = 0;
yAkte.fx(k,t)$(not dpartD(ss1,ss2,ss3,ss4,k,t)) = 0;
yQsepinkts.fx(k,t)$(not dpartD(ss1,ss2,ss3,ss4,k,t)) = 0;
yQnextkts.fx(k,t)$(not dpartD(ss1,ss2,ss3,ss4,k,t)) = 0;
yQRkts.fx(k,t)$(not dpartD(ss1,ss2,ss3,ss4,k,t)) = 0;

profit.optfile = 100 + (ord(ss1)-1)*27 + (ord(ss2)-1)*9 + (ord(ss3)-1)*3 + ord(ss4);

Solve profit MAXIMIZING Z Using MIP;
h(ss1,ss2,ss3,ss4) = profit.handle;
execute$card(h) 'gams checkincb lo=2 --jobs=81 %ide%';

);
```

----------Retrieve Solutions----------

```
repeat
loop((ss1,ss2,ss3,ss4),

if (handlestatus(h(ss1,ss2,ss3,ss4)) = 2,
profit.handle = h(ss1,ss2,ss3,ss4);
execute_loadhandle profit;
report(ss1,ss2,ss3,ss4,'obj') = profit.objVal;
```

```
report(ss1,ss2,ss3,ss4,'seconds') = profit.resusd;
report(ss1,ss2,ss3,ss4,'modelstat') = profit.modelstat;
report(ss1,ss2,ss3,ss4,'iteration_num') = profit.iterUsd;
report(ss1,ss2,ss3,ss4,'node_num') = profit.nodUsd;

if((profit.modelstat=%modelstat.Optimal% or profit.modelstat=%modelstat.IntegerSolution%) and
Z.l > bestcompleted,
bestcompleted  = profit.objVal;
execute_unload 'incb.gdx', bestcompleted);

display$handledelete(h(ss1,ss2,ss3,ss4)) 'trouble deleting handles' ;

h(ss1,ss2,ss3,ss4) = 0;));

*execute$card(h) 'gams checkincb lo=2 --jobs=81 %ide%';
display$sleep(card(h)*0.2) 'was sleeping for some time';

until card(h) = 0 or timeelapsed > 1800;
report(ss1,ss2,ss3,ss4,'obj')$h(ss1,ss2,ss3,ss4) = na;



Execute_Unload "3by3_Co.gdx", report;

display report, bestcompleted;
```

****----------Write the incumbent to solver's files----------**

```
$escape =
$onecho > writeincb.gms

$setDDlist userjobid ncalls
$set gdxincb %gams.workdir%bchout_i%=userjobid%=.gdx
$set txtincb %gams.workdir%incb%=userjobid%=.txt
Variable Z;
```

```
$gdxin '%gdxincb%' $load Z
File  fx / '%txtincb%' /; PutClose fx '%=userjobid%= ' Z.l:18:8;
$offecho
$escape %
$escape =
$onecho > checkincb.gms
$setDDlist jobs

$call cat incb*.txt > allincb.txt 2> %system.nullfile%
$if NOT exist allincb.txt $exit

file       fx        filehandle for text files
set        s         jobs submitted /0*%=jobs%=/
parameters incumbent previously found
           newincb   new incumbent
           incb(s)   incumbent values reported /
$             onempty oneps
$             include allincb.txt
                                      /;
incumbent = incb('0');
newincb   = smax(s$incb(s), incb(s));

if(newincb > incumbent,
  put_utility fx 'ren' / 'cplex.op3'; put '%cutoff% ' newincb:18:8;
  put_utility fx 'ren' / 'incb0.txt'; put '0 ' newincb:18:8;
*  ------tell everyone that we have a new incumbent by posting the trigger file---------
  loop(s,
    put_utility fx 'ren' / 't' s.tl:0 '.trg'; put 'trigger'));
$offecho
$escape %
```

Appendix D List of Publications

Kong, Q. and N. Shah (2016). "An optimisation-based framework for the conceptual design of reaction-separation processes." Chemical Engineering Research and Design 113: 206-222.

Kong, Q. and N. Shah (2017). "Development of an Optimization-Based Framework for Simultaneous Process Synthesis and Heat Integration." Industrial & Engineering Chemistry Research 56(17): 5000-5013.

Kong, Q. and N. Shah (2017). An Optimisation-based Framework for Simultaneous Process Synthesis and Heat Integration. Computer Aided Chemical Engineering. A. Espuña, M. Graells and L. Puigjaner, Elsevier. 40: 619-624.

References

Achenie, L. E. K. and L. T. Biegler (1986). "Algorithmic synthesis of chemical reactor networks using mathematical programming." Industrial & Engineering Chemistry Fundamentals **25**(4): 621-627.

Achenie, L. K. E. and L. T. Biegler (1990). "A superstructure based approach to chemical reactor network synthesis." Computers & Chemical Engineering **14**(1): 23-40.

Adams, A., J. C. R. Demyttenaere and N. De Kimpe (2003). "Biotransformation of (R)-(+)- and (S)-(−)-limonene to α-terpineol by Penicillium digitatum— investigation of the culture conditions." Food Chemistry **80**(4): 525-534.

Alibaba. (2015). "Global trade center." Retrieved 02/12, 2015, from www.alibaba.com.

Alonso, D. M., S. G. Wettstein and J. A. Dumesic (2013). "Gamma-valerolactone, a sustainable platform molecule derived from lignocellulosic biomass." Green Chemistry **15**(3): 584-595.

Andiappan, V., A. S. Y. Ko, V. W. S. Lau, L. Y. Ng, R. T. L. Ng, N. G. Chemmangattuvalappil and D. K. S. Ng (2014). "Synthesis of sustainable integrated biorefinery via reaction pathway synthesis: Economic, incremental enviromental burden and energy assessment with multiobjective optimization." AIChE Journal: n/a-n/a.

Andrecovich, M. J. and A. W. Westerberg (1985). "An MILP formulation for heat-integrated distillation sequence synthesis." AIChE Journal **31**(9): 1461-1474.

Ashley, V. M. and P. Linke (2004). "A Novel Approach for Reactor Network Synthesis Using Knowledge Discovery and Optimization Techniques." Chemical Engineering Research and Design **82**(8): 952-960.

Bahr, M., A. Bitto and R. Mulhaupt (2012). "Cyclic limonene dicarbonate as a new monomer for non-isocyanate oligo- and polyurethanes (NIPU) based upon terpenes." Green Chemistry **14**(5): 1447-1454.

Baliban, R. C., J. A. Elia and C. A. Floudas (2011). "Optimization framework for the simultaneous process synthesis, heat and power integration of a thermochemical hybrid biomass, coal, and natural gas facility." Computers & Chemical Engineering **35**(9): 1647-1690.

Bao, B., D. K. S. Ng, D. H. S. Tay, A. Jiménez-Gutiérrez and M. M. El-Halwagi (2011). "A shortcut method for the preliminary synthesis of process-technology pathways: An optimization approach and application for the conceptual design of integrated biorefineries." Computers & Chemical Engineering **35**(8): 1374-1383.

Bassett, M. H., J. F. Pekny and G. V. Reklaitis (1996). "Decomposition techniques for the solution of large-scale scheduling problems." AIChE Journal **42**(12): 3373-3387.

Beamon, B. M. (1998). "Supply chain design and analysis:: Models and methods." International Journal of Production Economics **55**(3): 281-294.

Benders, J. F. (1962). "Partitioning procedures for solving mixed-variables programming problems." Numerische Mathematik **4**(1): 238-252.

Beneke, D. A., D. Hildebrandt and D. Glasser (2011). "On Column Profile Maps: An Analysis of Sharp Splits." Industrial & Engineering Chemistry Research **50**(10): 6331-6342.

Beraldi, P., G. Ghiani, A. Grieco and E. Guerriero (2008). "Rolling-horizon and fix-and-relax heuristics for the parallel machine lot-sizing and scheduling problem with sequence-dependent set-up costs." Computers & Operations Research **35**(11): 3644-3656.

Berti, C., E. Binassi, M. Colonna, M. Fiorini, G. Kannan, S. Karanam, M. Mazzacurati, I. Odeh and M. Vanninl (2010). Bio-Based Terephthalate Polyesters, Google Patents.

Biegler, L. T., I. E. Grossmann and A. W. Westerberg (1997). Systematic methods of chemical process design, Prentice Hall PTR.

Boukouvala, F. and M. G. Ierapetritou (2013). "Surrogate-Based Optimization of Expensive Flowsheet Modeling for Continuous Pharmaceutical Manufacturing." Journal of Pharmaceutical Innovation **8**(2): 131-145.

Broeren, M. L. M., K. Molenveld, M. J. A. van den Oever, M. K. Patel, E. Worrell and L. Shen (2016). "Early-stage sustainability assessment to assist with material selection: a case study for biobased printer panels." Journal of Cleaner Production **135**: 30-41.

Brunet, R., D. Cortés, G. Guillén-Gosálbez, L. Jiménez and D. Boer (2012). "Minimization of the LCA impact of thermodynamic cycles using a combined simulation-optimization approach." Applied Thermal Engineering **48**: 367-377.

Brunet, R., G. Guillén-Gosálbez, J. R. Pérez-Correa, J. A. Caballero and L. Jiménez (2012). "Hybrid simulation-optimization based approach for the optimal design of single-product biotechnological processes." Computers & Chemical Engineering **37**: 125-135.

Brunet, R., J. A. Reyes-Labarta, G. Guillén-Gosálbez, L. Jiménez and D. Boer (2012). "Combined simulation–optimization methodology for the design of environmental conscious absorption systems." Computers & Chemical Engineering **46**: 205-216.

Bureš, M., V. Majer and M. Zábranský (1981). "Modification of Benson method for estimation of ideal-gas heat capacities." Chemical Engineering Science **36**(3): 529-537.

Bussieck, M. R., M. C. Ferris and A. Meeraus (2009). "Grid-Enabled Optimization with GAMS." INFORMS Journal on Computing **21**(3): 349-362.

References

Buxton, A., A. G. Livingston and E. N. Pistikopoulos (1997). "Reaction path synthesis for environmental impact minimization." Computers & Chemical Engineering **21**: S959-S964.

Byrne, C. M., S. D. Allen, E. B. Lobkovsky and G. W. Coates (2004). "Alternating Copolymerization of Limonene Oxide and Carbon Dioxide." Journal of the American Chemical Society **126**(37): 11404-11405.

Caballero, J. A. and I. E. Grossmann (2004). "Design of distillation sequences: from conventional to fully thermally coupled distillation systems." Computers & Chemical Engineering **28**(11): 2307-2329.

Caballero, J. A. and I. E. Grossmann (2008). "An algorithm for the use of surrogate models in modular flowsheet optimization." AIChE Journal **54**(10): 2633-2650.

Caballero, J. A., A. Odjo and I. E. Grossmann (2007). "Flowsheet optimization with complex cost and size functions using process simulators." AIChE Journal **53**(9): 2351-2366.

Candela, T. and A. Fouet (2006). "Poly-gamma-glutamate in bacteria." Molecular Microbiology **60**(5): 1091-1098.

Chen, G.-Q. and M. K. Patel (2012). "Plastics Derived from Biological Sources: Present and Future: A Technical and Environmental Review." Chemical Reviews **112**(4): 2082-2099.

Chen, Q. and I. E. Grossmann (2017). "Recent Developments and Challenges in Optimization-Based Process Synthesis." Annual Review of Chemical and Biomolecular Engineering **8**(1): 249-283.

Cherubini, F. (2010). "The biorefinery concept: Using biomass instead of oil for producing energy and chemicals." Energy Conversion and Management **51**(7): 1412-1421.

Ciriminna, R., M. Lomeli-Rodriguez, P. Demma Cara, J. A. Lopez-Sanchez and M. Pagliaro (2014). "Limonene: a versatile chemical of the bioeconomy." Chemical Communications **50**(97): 15288-15296.

Collias, D. I., A. M. Harris, V. Nagpal, I. W. Cottrell and M. W. Schultheis (2014). "Biobased Terephthalic Acid Technologies: A Literature Review." Industrial Biotechnology **10**(2): 91-105.

Colonna, M., C. Berti, M. Fiorini, E. Binassi, M. Mazzacurati, M. Vannini and S. Karanam (2011). "Synthesis and radiocarbon evidence of terephthalate polyesters completely prepared from renewable resources." Green Chemistry **13**(9): 2543-2548.

Contractors, D. S. E. (2016). "Biochemical Production Plants." Retrieved 02/11, 2016, from http://www.dsengineers.com/en/sectors-of-activities/bio-based-products/biochemical-production-plants.html.

Demidova, Y. S., E. V. Suslov, O. A. Simakova, I. L. Simakova, K. P. Volcho, N. F. Salakhutdinov and D. Y. Murzin (2015). "Selective carvone hydrogenation to dihydrocarvone over titania supported gold catalyst." Catalysis Today **241, Part B**: 189-194.

Dimitriadis, A. D., N. Shah and C. C. Pantelides (1997). "RTN-based rolling horizon algorithms for medium term scheduling of multipurpose plants." Computers & Chemical Engineering **21**: S1061-S1066.

Diwekar, U. M., I. E. Grossmann and E. S. Rubin (1992). "An MINLP process synthesizer for a sequential modular simulator." Industrial & Engineering Chemistry Research **31**(1): 313-322.

Doherty, M. F. and J. D. Perkins (1978). "On the dynamics of distillation processes—I." Chemical Engineering Science **33**(3): 281-301.

Doherty, M. F. and J. D. Perkins (1978). "On the dynamics of distillation processes—II. The simple distillation of model solutions." Chemical Engineering Science **33**(5): 569-578.

Doherty, M. F. and J. D. Perkins (1979). "On the dynamics of distillation processes—III." Chemical Engineering Science **34**(12): 1401-1414.

Douglas, J. M. (1988). Conceptual design of chemical processes, McGraw-Hill.

Duran, M. A. and I. E. Grossmann (1986). "Simultaneous optimization and heat integration of chemical processes." AIChE Journal **32**(1): 123-138.

EIA. (2018). "Henry Hub Natural Gas Spot Price." Retrieved 29/08, 2018, from https://www.eia.gov/dnav/ng/hist/rngwhhdA.htm.

EIA. (2018). "Spot Prices for Crude Oil and Petroleum Products." Retrieved 29/08, 2018, from https://www.eia.gov/dnav/pet/pet_pri_spt_s1_a.htm.

Equi, L., G. Gallo, S. Marziale and A. Weintraub (1997). "A combined transportation and scheduling problem." European Journal of Operational Research **97**(1): 94-104.

Essentials, I. E. T. (2007). Biofuel Production.

European Bioplastic (2017). Bioplastics market data 2017.

Fahmi, I. and S. Cremaschi (2012). "Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models." Computers & Chemical Engineering **46**: 105-123.

Fenske, M. R. (1932). "Fractionation of Straight-Run Pennsylvania Gasoline." Industrial & Engineering Chemistry **24**(5): 482-485.

Fernando, S., S. Adhikari, C. Chandrapal and N. Murali (2006). "Biorefineries: Current Status, Challenges, and Future Direction." Energy & Fuels **20**(4): 1727-1737.

Ferrara, M. A., D. S. Almeida, A. C. Siani, L. Lucchetti, P. S. B. Lacerda, A. Freitas, M. R. R. Tappin and E. P. S. Bon (2013). "Bioconversion of R-(+)-limonene to perillic acid by the yeast Yarrowia lipolytica." Brazilian Journal of Microbiology **44**(4): 1075-1080.

Floudas, C. A. and A. R. Ciric (1989). "Strategies for overcoming uncertainties in heat exchanger network synthesis." Computers & Chemical Engineering **13**(10): 1133-1152.

Fornari, T. and G. Stephanopoulos (1994). "SYNTHESIS OF CHEMICAL REACTION PATHS: ECONOMIC AND SPECIFICATION CONSTRAINTS." Chemical Engineering Communications **129**(1): 159-182.

Gilliland, E. R. (1940). "Multicomponent Rectification Estimation of the Number of Theoretical Plates as a Function of the Reflux Ratio." Industrial & Engineering Chemistry **32**(9): 1220-1223.

Govind, R. and G. J. Powers (1981). "Studies in reaction path synthesis." AIChE Journal **27**(3): 429-442.

Grossmann, I. E. (1985). "Mixed-integer programming approach for the synthesis of integrated process flowsheets." Computers & Chemical Engineering **9**(5): 463-482.

Hardcastle, J. L. (2015). Commercial-Scale Biochemical Production Plant Online. Environmental Leader Environmental & Energy Management News.

Hauenstein, O., M. Reiter, S. Agarwal, B. Rieger and A. Greiner (2015). "Bio-based polycarbonate from limonene oxide and CO2 with high molecular weight, excellent thermal resistance, hardness and transparency." Green Chemistry.

Hechinger, M., A. Voll and W. Marquardt (2010). "Towards an integrated design of biofuels and their production pathways." Computers & Chemical Engineering **34**(12): 1909-1918.

Henao, C. A. and C. T. Maravelias (2011). "Surrogate-based superstructure optimization framework." AIChE Journal **57**(5): 1216-1232.

Hengstebeck, R. J. (1961). Distillation: principles and design procedures, Reinhold Pub. Corp.

Holland, S. T., R. Abbas, D. Hildebrandt and D. Glasser (2010). "Complex Column Design by Application of Column Profile Map Techniques: Sharp-Split Petlyuk Column Design." Industrial & Engineering Chemistry Research **49**(1): 327-349.

Holland, S. T., M. Tapp, D. Hildebrandt and D. Glasser (2004). "Column Profile Maps. 2. Singular Points and Phase Diagram Behaviour in Ideal and Nonideal Systems." Industrial & Engineering Chemistry Research **43**(14): 3590-3603.

Holm, J. and U. Lassi (2011). Ionic Liquids in the Pretreatment of Lignocellulosic Biomass. Ionic Liquids: Applications and Perspectives.

Huber, G. W., S. Iborra and A. Corma (2006). "Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering." Chemical Reviews **106**(9): 4044-4098.

IPCC (2014). Climate Change 2014 Synthesis Report.

Isikgor, F. H. and C. R. Becer (2015). "Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers." Polymer Chemistry **6**(25): 4497-4559.

Ivakpour, J. and N. Kasiri (2009). "Synthesis of Distillation Column Sequences for Nonsharp Separations." Industrial & Engineering Chemistry Research **48**(18): 8635-8649.

Iyer, R. R. and I. E. Grossmann (1998). "Synthesis and operational planning of utility systems for multiperiod operation." Computers & Chemical Engineering **22**(7): 979-993.

Jackson, R. (1968). "Optimization of chemical reactors with respect to flow configuration." Journal of Optimization Theory and Applications **2**(4): 240-259.

Jankowski, M. D., C. S. Henry, L. J. Broadbelt and V. Hatzimanikatis (2008). "Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks." Biophysical Journal **95**(3): 1487-1499.

Jeske, R. C., A. M. DiCiccio and G. W. Coates (2007). "Alternating Copolymerization of Epoxides and Cyclic Anhydrides: An Improved Route to Aliphatic Polyesters." Journal of the American Chemical Society **129**(37): 11330-11331.

Ji, S. and M. Bagajewicz (2002). "On the Energy Efficiency of Stripping-Type Crude Distillation." Industrial & Engineering Chemistry Research **41**(23): 5819-5825.

Joback, K. G. and R. C. Reid (1987). "ESTIMATION OF PURE-COMPONENT PROPERTIES FROM GROUP-CONTRIBUTIONS." Chemical Engineering Communications **57**(1-6): 233-243.

Julián-Durán, L. M., A. P. Ortiz-Espinoza, M. M. El-Halwagi and A. Jiménez-Gutiérrez (2014). "Techno-Economic Assessment and Environmental Impact of Shale Gas Alternatives to Methanol." ACS Sustainable Chemistry & Engineering **2**(10): 2338-2344.

Julka, V. and M. F. Doherty (1990). "Geometric behavior and minimum flows for nonideal multicomponent distillation." Chemical Engineering Science **45**(7): 1801-1822.

Julka, V. and M. F. Doherty (1993). "Geometric nonlinear analysis of multicomponent nonideal distillation: a simple computer-aided design procedure." Chemical Engineering Science **48**(8): 1367-1391.

Kelloway, A. and P. Daoutidis (2014). "Process Synthesis of Biorefineries: Optimization of Biomass Conversion to Fuels and Chemicals." Industrial & Engineering Chemistry Research **53**(13): 5261-5273.

Kim, J., S. M. Sen and C. T. Maravelias (2013). "An optimization-based assessment framework for biomass-to-fuel conversion strategies." Energy & Environmental Science **6**(4): 1093-1104.

Kjonaas, R. A. and S. P. Mattingly (2005). "Acid-Catalyzed Isomerization of Carvone to Carvacrol." Journal of Chemical Education **82**(12): 1813.

Kong, Q. and N. Shah (2016). "An optimisation-based framework for the conceptual design of reaction-separation processes." Chemical Engineering Research and Design **113**: 206-222.

Kong, Q. and N. Shah (2017). "Development of an Optimization-Based Framework for Simultaneous Process Synthesis and Heat Integration." Industrial & Engineering Chemistry Research **56**(17): 5000-5013.

Kong, Q. and N. Shah (2017). An Optimisation-based Framework for Simultaneous Process Synthesis and Heat Integration. Computer Aided Chemical Engineering. A. Espuña, M. Graells and L. Puigjaner, Elsevier. **40:** 619-624.

Kostin, A. M., G. Guillén-Gosálbez, F. D. Mele, M. J. Bagajewicz and L. Jiménez (2011). "A novel rolling horizon strategy for the strategic planning of supply chains. Application to the sugar cane industry of Argentina." Computers & Chemical Engineering **35**(11): 2540-2563.

Krishnamurthy, R. and R. Taylor (1985). "A nonequilibrium stage model of multicomponent separation processes. Part I: Model description and method of solution." AIChE Journal **31**(3): 449-456.

Kukhta, N. A., I. V. Vasilenko and S. V. Kostjuk (2011). "Room temperature cationic polymerization of β-pinene using modified AlCl3 catalyst: toward sustainable plastics from renewable biomass resources." Green Chemistry **13**(9): 2362-2364.

Lakshmanan, A. and L. T. Biegler (1996). "Synthesis of Optimal Chemical Reactor Networks." Industrial & Engineering Chemistry Research **35**(4): 1344-1353.

Lange, B. M. (2015). Biosynthesis and Biotechnology of High-Value p-Menthane Monoterpenes, Including Menthol, Carvone, and Limonene. Biotechnology of Isoprenoids. J. Schrader and J. Bohlmann. Cham, Springer International Publishing**:** 319-353.

Laukkanen, T. and C.-J. Fogelholm (2011). "A bilevel optimization method for simultaneous synthesis of medium-scale heat exchanger networks based on grouping of process streams." Computers & Chemical Engineering **35**(11): 2389-2400.

Leboreiro, J. and J. Acevedo (2004). "Processes synthesis and design of distillation sequences using modular simulators: a genetic algorithm framework." Computers & Chemical Engineering **28**(8): 1223-1236.

Lewin, D. R., H. Wang and O. Shalev (1998). "A generalized method for HEN synthesis using stochastic optimization – I. General framework and MER optimal synthesis." Computers & Chemical Engineering **22**(10): 1503-1513.

Li, J., Z. Li, G. Zi, Z. Yao, Z. Luo, Y. Wang, D. Xue, B. Wang and J. Wang (2015). "Synthesis, characterizations and catalytic allylic oxidation of limonene to carvone of cobalt doped mesoporous silica templated by reed leaves." Catalysis Communications **59**: 233-237.

Li, M., S. Hu, Y. Li and J. Shen (2000). "A Hierarchical Optimization Method for Reaction Path Synthesis." Industrial & Engineering Chemistry Research **39**(11): 4315-4319.

Li, X. and A. Kraslawski (2004). "Conceptual process synthesis: past and current trends." Chemical Engineering and Processing: Process Intensification **43**(5): 583-594.

Linnhoff, B. and J. R. Flower (1978). "Synthesis of heat exchanger networks: I. Systematic generation of energy optimal networks." AIChE Journal **24**(4): 633-642.

Lowe, J. R., W. B. Tolman and M. A. Hillmyer (2009). "Oxidized Dihydrocarvone as a Renewable Multifunctional Monomer for the Synthesis of Shape Memory Polyesters." Biomacromolecules **10**(7): 2003-2008.

Lu, J., R. C. Tappel and C. T. Nomura (2009). "Mini-Review: Biosynthesis of Poly(hydroxyalkanoates)." Polymer Reviews **49**(3): 226-248.

Magnanti, T. L., P. Mireault and R. T. Wong (1986). Tailoring Benders decomposition for uncapacitated network design. Netflow at Pisa. G. Gallo and C. Sandi. Berlin, Heidelberg, Springer Berlin Heidelberg**:** 112-154.

Marrero, J. and R. Gani (2001). "Group-contribution based estimation of pure component properties." Fluid Phase Equilibria **183-184**: 183-208.

Martín-Luengo, M. A., M. Yates, M. J. Martínez Domingo, B. Casal, M. Iglesias, M. Esteban and E. Ruiz-Hitzky (2008). "Synthesis of p-cymene from limonene, a renewable feedstock." Applied Catalysis B: Environmental **81**(3–4): 218-224.

Maša Šprajcar, P. H., Andrej Kržan (2012). BIOPOLYMERS AND BIOPLASTICS: PLASTICS ALIGNED WITH NATURE.

Mascia, M., F. Ferrara, A. Vacca, G. Tola and M. Errico (2007). "Design of heat integrated distillation systems for a light ends separation plant." Applied Thermal Engineering **27**(7): 1205-1211.

Matthews, H. D., N. P. Gillett, P. A. Stott and K. Zickfeld (2009). "The proportionality of global warming to cumulative carbon emissions." Nature **459**: 829.

McCabe, W. L. and E. W. Thiele (1925). "Graphical Design of Fractionating Columns." Industrial & Engineering Chemistry **17**(6): 605-611.

McGlade, C. and P. Ekins (2015). "The geographical distribution of fossil fuels unused when limiting global warming to 2 °C." Nature **517**: 187.

Morschbacker, A. (2009). "Bio-Ethanol Based Ethylene." Polymer Reviews **49**(2): 79-84.

Mouawad, J. (2009). OPEC Achieves Cuts in Output, Halting Price Slide. The New York Times.

Naessens, M., A. Cerdobbel, W. Soetaert and E. J. Vandamme (2005). "Leuconostoc dextransucrase and dextran: production, properties and applications." Journal of Chemical Technology & Biotechnology **80**(8): 845-860.

Naka, Y., M. Terashita and T. Takamatsu (1982). "A thermodynamic approach to multicomponent distillation system synthesis." AIChE Journal **28**(5): 812-820.

National Research Council (2005). Sustainability in the Chemical Industry: Grand Challenges and Research Needs - A Workshop Report. Washington, DC, The National Academies Press.

Nejad, E. H., A. Paoniasari, C. G. W. van Melis, C. E. Koning and R. Duchateau (2013). "Catalytic Ring-Opening Copolymerization of Limonene Oxide and Phthalic Anhydride: Toward Partially Renewable Polyesters." Macromolecules **46**(3): 631-637.

Neupane, B., M. Jerrett, R. T. Burnett, T. Marrie, A. Arain and M. Loeb (2010). "Long-Term Exposure to Ambient Air Pollution and Risk of Hospitalization with Community-acquired Pneumonia in Older Adults." American Journal of Respiratory and Critical Care Medicine **181**(1): 47-53.

Ng, D. K. S. (2010). "Automated targeting for the synthesis of an integrated biorefinery." Chemical Engineering Journal **162**(1): 67-74.

Nishida, N., G. Stephanopoulos and A. W. Westerberg (1981). "A review of process synthesis." AIChE Journal **27**(3): 321-351.

Oil & Gas Journal (2009). FACTS: 2009-10 LNG demand flat; supplies to glut markets. Oil & Gas Journal.

Okabe, M., D. Lies, S. Kanamasa and E. Park (2009). "Biotechnological production of itaconic acid and its biosynthesis in Aspergillus terreus." Applied Microbiology and Biotechnology **84**(4): 597-606.

Pantelides, C. C. (1994). Unified frameworks for optimal process planning and scheduling. Proc. Second Conf. on Foundations of Computer-Aided Operations, Rippin.

Papoulias, S. A. and I. E. Grossmann (1983). "A structural optimization approach in process synthesis—II." Computers & Chemical Engineering **7**(6): 707-721.

Pappa, G., C. Boukouvalas, C. Giannaris, N. Ntaras, V. Zografos, K. Magoulas, A. Lygeros and D. Tassios (2001). "The selective dissolution/precipitation technique for polymer recycling: a pilot unit application." Resources, Conservation and Recycling **34**(1): 33-44.

Peng, F. and G. Cui (2015). "Efficient simultaneous synthesis for heat exchanger network with simulated annealing algorithm." Applied Thermal Engineering **78**: 136-149.

Phan, A. N. and A. P. Harvey (2012). "Characterisation of mesoscale oscillatory helical baffled reactor—Experimental approach." Chemical Engineering Journal **180**: 229-236.

PlasticsEurope (2014). Plastics – the Facts 2014, Association of Plastics Manufacturers.

Pohorecki, R., J. Bridgwater and M. M. R. G. C. Gallegos (2010). Chemical Engineering and Chemical Process Technology - Volume II: Unit Operations – Fluids and Solids.

Ponce-Ortega, J. M., V. Pham, M. M. El-Halwagi and A. A. El-Baz (2012). "A Disjunctive Programming Formulation for the Optimal Design of Biorefinery Configurations." Industrial & Engineering Chemistry Research **51**(8): 3381-3400.

Quirante, N., J. A. Caballero and I. E. Grossmann (2017). "A novel disjunctive model for the simultaneous optimization and heat integration." Computers & Chemical Engineering **96**: 149-168.

Quirante, N., I. E. Grossmann and J. A. Caballero (2018). "Disjunctive model for the simultaneous optimization and heat integration with unclassified streams and area estimation." Computers & Chemical Engineering **108**: 217-231.

Ragauskas, A. J., C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, J. P. Hallett, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer and T. Tschaplinski (2006). "The Path Forward for Biofuels and Biomaterials." Science **311**(5760): 484.

Rana, K. and R. G. Vickson (1991). "Routing Container Ships Using Lagrangean Relaxation and Decomposition." Transportation Science **25**(3): 201-214.

Rehm, B. H. A. (2010). "Bacterial polymers: biosynthesis, modifications and applications." Nature Reviews Microbiology **8**: 578.

Rotstein, E., D. Resasco and G. Stephanopoulos (1982). "Studies on the synthesis of chemical reaction paths—I: Reaction characteristics in the (ΔG, T) space and a primitive synthesis procedure." Chemical Engineering Science **37**(9): 1337-1352.

Ryan, P. G. (2015). A Brief History of Marine Litter Research. Marine Anthropogenic Litter. M. Bergmann, L. Gutow and M. Klages. Cham, Springer International Publishing**:** 1-25.

Santoso, T., S. Ahmed, M. Goetschalckx and A. Shapiro (2005). "A stochastic programming approach for supply chain network design under uncertainty." European Journal of Operational Research **167**(1): 96-115.

Schweiger, C. A. and C. A. Floudas (1999). "Optimization Framework for the Synthesis of Chemical Reactor Networks." Industrial & Engineering Chemistry Research **38**(3): 744-766.

Sharma, S. and A. K. Srivastava (2003). "Free radical copolymerization of limonene with butyl methacrylate: Synthesis and characterization." Indian Journal of Chemical Technology **12**(1): 62-67.

Sharma, S. and A. K. Srivastava (2004). "Synthesis and characterization of copolymers of limonene with styrene initiated by azobisisobutyronitrile." European Polymer Journal **40**(9): 2235-2240.

Siirola, J. J. and D. F. Rudd (1971). "Computer-Aided Synthesis of Chemical Process Designs. From Reaction Path Data to the Process Task Network." Industrial & Engineering Chemistry Fundamentals **10**(3): 353-362.

Singh, A. and M. Kamal (2012). "Synthesis and characterization of polylimonene: Polymer of an optically active terpene." Journal of Applied Polymer Science **125**(2): 1456-1459.

Somerville, C., H. Youngs, C. Taylor, S. C. Davis and S. P. Long (2010). "Feedstocks for Lignocellulosic Biofuels." Science **329**(5993): 790.

Stefanis, S. K., A. Livingston and E. N. Pistikopoulos (2001). Life Cycle Optimization. Wiley Encyclopedia of Electrical and Electronics Engineering, John Wiley & Sons, Inc.

Sun, N., H. Rodríguez, M. Rahman and R. D. Rogers (2011). "Where are ionic liquid strategies most suited in the pursuit of chemicals and energy from lignocellulosic biomass?" Chemical Communications **47**(5): 1405-1421.

Sun, Y. and J. Cheng (2002). "Hydrolysis of lignocellulosic materials for ethanol production: a review." Bioresource Technology **83**(1): 1-11.

Tang, M., M. S. Chin, K. Lim, Y. Mun, R. L. Ng, D. S. Tay and D. S. Ng (2013). "Systematic approach for conceptual design of an integrated biorefinery with uncertainties." Clean Technologies and Environmental Policy **15**(5): 783-799.

Tapp, M., S. T. Holland, D. Hildebrandt and D. Glasser (2004). "Column Profile Maps. 1. Derivation and Interpretation." Industrial & Engineering Chemistry Research **43**(2): 364-374.

Terrazas-Moreno, S., P. A. Trotter and I. E. Grossmann (2011). "Temporal and spatial Lagrangean decompositions in multi-site, multi-period production planning problems with sequence-dependent changeovers." Computers & Chemical Engineering **35**(12): 2913-2928.

Thiele, E. W. and R. L. Geddes (1933). "Computation of Distillation Apparatus for Hydrocarbon Mixtures." Industrial & Engineering Chemistry **25**(3): 289-295.

Thong, D. Y. C., G. Liu, M. Jobson and R. Smith (2004). "Synthesis of distillation sequences for separating multicomponent azeotropic mixtures." Chemical Engineering and Processing: Process Intensification **43**(3): 239-250.

Thurston George, D., T. Burnett Richard, C. Turner Michelle, Y. Shi, D. Krewski, R. Lall, K. Ito, M. Jerrett, M. Gapstur Susan, W. R. Diver and C. A. Pope (2016). "Ischemic Heart Disease Mortality and Long-Term Exposure to Source-Related Components of U.S. Fine Particle Air Pollution." Environmental Health Perspectives **124**(6): 785-794.

Tullo, A. (2017). "New route planned to biobased ethylene glycol." C&EN Global Enterprise **95**(46): 10-10.

Uguina, M. A., J. A. Delgado and J. Carretero (2009). "Regiospecific Epoxidation of Carvone α,β-Unsaturated Ketone Group with a Basic Resin." Industrial & Engineering Chemistry Research **48**(23): 10217-10221.

Underwood, A. J. V. (1949). "Fractional Distillation of Multicomponent Mixtures." Industrial & Engineering Chemistry **41**(12): 2844-2847.

van den Heever, S. A., I. E. Grossmann, S. Vasantharajan and K. Edwards (2001). "A Lagrangean Decomposition Heuristic for the Design and Planning of Offshore Hydrocarbon Field Infrastructures with Complex Economic Objectives." Industrial & Engineering Chemistry Research **40**(13): 2857-2875.

Verevkin, S. P., Emel, apos, V. N. yanenko, V. Diky, C. D. Muzny, R. D. Chirico and M. Frenkel (2013). "New Group-Contribution Approach to Thermochemical Properties of Organic Compounds: Hydrocarbons and Oxygen-Containing Compounds." Journal of Physical and Chemical Reference Data **42**(3): 033102.

Voll, A. and W. Marquardt (2012). "Reaction network flux analysis: Optimization-based evaluation of reaction pathways for biorenewables processing." AIChE Journal **58**(6): 1788-1801.

Weastra (2013). Determination of market potential for selected platform chemicals: Itaconic acid, Succinic acid, 2,5-Furandicarboxylic acid.

Wehe, R. R. and A. W. Westerberg (1987). "An algorithmic procedure for the synthesis of distillation sequences with bypass." Computers & Chemical Engineering **11**(6): 619-627.

Weiss, M., J. Haufe, M. Carus, M. Brandão, S. Bringezu, B. Hermann and M. K. Patel (2012). "A Review of the Environmental Impacts of Biobased Materials." Journal of Industrial Ecology **16**(s1): S169-S181.

Westerberg, A. W. (2004). "A retrospective on design and process synthesis." Computers & Chemical Engineering **28**(4): 447-458.

Wiemann Lars, O., C. Faltl and V. Sieber (2012). Lipase-mediated Epoxidation of the Cyclic Monoterpene Limonene to Limonene Oxide and Limonene Dioxide. Zeitschrift für Naturforschung B. **67:** 1056.

Xu, J. and B.-H. Guo (2010). Microbial Succinic Acid, Its Polymer Poly(butylene succinate), and Applications. Plastics from Bacteria: Natural Functions and Applications. G. G.-Q. Chen. Berlin, Heidelberg, Springer Berlin Heidelberg**:** 347-388.

Yee, T. F. and I. E. Grossmann (1990). "Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis." Computers & Chemical Engineering **14**(10): 1165-1184.

Yee, T. F., I. E. Grossmann and Z. Kravanja (1990). "Simultaneous optimization models for heat integration—I. Area and energy targeting and modeling of multi-stream exchangers." Computers & Chemical Engineering **14**(10): 1151-1164.

Yee, T. F., I. E. Grossmann and Z. Kravanja (1990). "Simultaneous optimization models for heat integration—III. Process and heat exchanger network optimization." Computers & Chemical Engineering **14**(11): 1185-1200.

Yerramsetty, K. M. and C. V. S. Murty (2008). "Synthesis of cost-optimal heat exchanger networks using differential evolution." Computers & Chemical Engineering **32**(8): 1861-1876.

You, F. and I. E. Grossmann (2013). "Multicut Benders decomposition algorithm for process supply chain planning under uncertainty." Annals of Operations Research **210**(1): 191-211.

You, F., I. E. Grossmann and J. M. Wassick (2011). "Multisite Capacity, Production, and Distribution Planning with Reactor Modifications: MILP Model, Bilevel Decomposition Algorithm versus Lagrangean Decomposition Scheme." Industrial & Engineering Chemistry Research **50**(9): 4831-4849.

Yuan, Z., B. Chen and R. Gani (2013). "Applications of process synthesis: Moving from conventional chemical processes towards biorefinery processes." Computers & Chemical Engineering **49**: 217-229.

Zábranský, M. and V. Růžička (2004). "Estimation of the Heat Capacities of Organic Liquids as a Function of Temperature Using Group Additivity: An Amendment." Journal of Physical and Chemical Reference Data **33**(4): 1071-1081.

Zhou, C.-H., J. N. Beltramini, Y.-X. Fan and G. Q. Lu (2008). "Chemoselective catalytic conversion of glycerol as a biorenewable source to valuable commodity chemicals." Chemical Society Reviews **37**(3): 527-549.

Zhou, C.-H., X. Xia, C.-X. Lin, D.-S. Tong and J. Beltramini (2011). "Catalytic conversion of lignocellulosic biomass to fine chemicals and fuels." Chemical Society Reviews **40**(11): 5588-5617.