Thesis submitted in fulfilment of the requirements of the degree of PhD from

Imperial College London

The Structure and Manipulation of Resources in Quantum Information Theory

Thomas Hebdige

Department of Physics

Supervised by

Dr. David Jennings

December 12, 2020

Declaration

The content of this thesis is the result of my work over the past three years under the supervision of Dr. David Jennings. I declare that the work is my own, and that the work of others has been appropriately acknowledged and cited.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

The content of this thesis is based on the following papers:

1. Thomas Hebdige & David Jennings, "On the classification of two-qubit group orbits and the use of coarse-grained 'shape' as a superselection property", Quantum 3, 119 (2019) [1].

2. Nikolaos Koukoulekidis, Rhea Alexander, Thomas Hebdige & David Jennings, "The geometry of passivity for quantum systems and a novel elementary derivation of the Gibbs state", arXiv:1912.07968 [quant-ph] [2].

3. Thomas Hebdige & David Jennings, "The role of parameter estimation in classical and quantum Bernoulli factories", in preparation.

Acknowledgements

Firstly, I would like to thank my supervisor, David Jennings, for all his support, advice, debates and patience during the project. I am grateful for the new perspectives on physics and academia that he has provided throughout my PhD, as well as important insights into Ireland and its people. I was also like to thank my collaborators Rhea Alexander and Nick Koukoulekidis, and also Carlo Sparaciari, Krys Latuszynski, Suzie Brown, Ewan Cameron, Iman Marvian and Cristina Cirstoiu for useful discussions during this work.

I am lucky to have been part of the CDT in Controlled Quantum Dynamics, not least for the funding from the EPSRC. My thanks go to the rest of Cohort 7 – Markus Frembs, Dougie Garrett, Claudia Gollner, Erick Hinds Mingo, Chris Ho, Zoë Holmes, Tom Parker, Mike Trigatzis & Ben Walker – for many cohort lunches and discussions in the Whiteley kitchen. Special thanks go to Miranda Smith, Veena Dhulipala and Naho Ollason for their help with stationary, expenses and general organisation. I would also like to thank the many inhabitants of Level 12 that I have met during my time at Imperial. Lunchtimes would have been far less interesting without Doug Plato, Ben Dive, Alex Paige, Adam Callison, Jack Clarke, Selwyn Simsek, Kiran Khosla, and the countless others that have joined over the years.

My thanks also go to all the volunteers and staff that I have been fortunate enough to work with at the Science Museum. Thank you to Dave Inman, Johanna Stegner, Mandy Ho, Cheryl Choo, Haira Gandolfi and Julia Barrett for support and encouragement throughout my studies, and always asking interesting questions about my research, even if my answers didn't really make sense. Last but not least, thank you to my parents for their support (especially the food), and the rest of my family. In particular, I would like to thank my grandad for making science interesting in the first place.

Abstract

Resources arise whenever there is some constraint, either practical or fundamental. Studying the structures that arise in connection to these resources can enhance our understanding of important features in quantum information theory. Furthermore, the manipulation of these resources can improve our ability to perform computational tasks, in certain cases providing a quantum advantage over the classical analogue. In this thesis we study the structure of resources in thermodynamics, in particular the structure of passivity, with a novel geometric approach. Then we provide a classification scheme for quantum resources under symmetry constraints, considering both the abstract structures and their consequences. Finally, we study computational resources for resampling algorithms on quantum computers, and how quantum resources enable a proven quantum advantage for a class of algorithms called Bernoulli factories.

Contents

1 Introduction **9**

1.1 The Geometry of Passivity 10

1.2 Resources from Symmetry Constraints 11

1.3 Resampling Algorithms for Quantum Computers 11

I The Geometry of Passivity **12**

2 A planar and simple derivation of the Gibbs state **13**

2.1 Motivation – Why Passivity? 14

2.2 Passivity 16

2.2.1 Definition of Passivity 17

2.2.2 Conditions for Passivity 18

2.3 Extensivity 20

2.4 A Geometric Perspective on Passivity 22

2.4.1 The ϵ-s Ensemble 22

2.4.2 Passivity as a Total Order on Vectors 23

2.5 A Geometric Perspective on Complete Passivity 24

2.5.1 k-Passivity and the Regularised ϵ-s Ensemble 24

2.5.2 Asymptotic ϵ-s Ensemble 26

2.5.3 Some Properties of the Asymptotic Ensemble 28

2.6 A geometric derivation of the Gibbs state . . . 30
2.6.1 Inverse Temperature as Gradient . . . 32
2.7 Degeneracy, Structural Stability and Ground States . . . 33
2.8 Summary and Outlook . . . 35

3 A Geometric Measure of Athermality 38

3.1 Macroscopic Non-equilibrium Thermodynamics . . . 39
3.2 The Macroscopic Regime and the Energy-Entropy Diagram . . . 40
3.3 Geometric Athermality . . . 43
3.4 Geometric Athermality as a Measure of Athermality . . . 44
3.5 Connecting to other measures of athermality . . . 46
3.5.1 Asymptotic Ergotropy $W_{\max}$ . . . 47
3.5.2 β-Athermality a_β . . . 47
3.5.3 Asymptotic Maximal Entropic Gain $\Delta S_{\max}$ . . . 48
3.5.4 Connecting $W_{\max}$ and $\Delta S_{\max}$ to $\mathcal{A}$ . . . 49
3.5.5 Comparison of athermality measures . . . 51
3.6 Conclusions and Outlook . . . 51

II Resources from Symmetry Constraints 53

4 Classifying Symmetries of Quantum States & Channels 54

4.1 Motivation – Symmetry in Quantum Information Theory . . . 55
4.2 A Mathematical Description of Symmetry . . . 56
4.2.1 Group Theory – the Structure of Symmetry . . . 57
4.2.2 Representations and Group Actions . . . 59
4.3 Classifying the Symmetries of Quantum States & Channels . . . 60
4.4 Group Orbits and Homogeneous Spaces . . . 62

4.5 The Robustness of Shapes to Small Perturbations 65
4.6 Conclusions & Outlook 69

5 Illustrative Example: Two Qubits Under SU(2) Symmetry Constraints 70

5.1 Two-Qubit States 71
5.2 SU(2) Group Action on Two Qubits 73
5.3 Symmetries of Bell States under SU(2) 76
5.3.1 Aside: Semi-Direct Product Groups 79
5.4 T-States 80
5.5 Maximally Mixed Marginals 83
5.5.1 Equal Local Unitaries 83
5.5.2 Unequal Local Unitaries 84
5.6 Local Bloch Vectors 85
5.7 Conclusions & Outlook 86

6 'Shape' as a Superselection Property 88

6.1 Motivation – Dynamics under Symmetry Constraints 89
6.2 Mapping to the Subgroup Lattice 91
6.3 H-Twirling 93
6.4 A Resource Theory for Shapes? 96
6.5 Asymmetry as a Resource 98
6.6 Efficient Simulation of Channels 99
6.7 Conclusions & Outlook 103

III Resampling Algorithms for Quantum Computers 105

7 Parameter Estimation in Classical & Quantum Bernoulli Factories 106

7.1 Background and Motivation 107

7.2 The von Neumann Coin Trick . . . 109
7.3 Classical Bernoulli Factories . . . 110
7.4 Quantum Bernoulli Factories . . . 114
7.5 Bernoulli Factories for Constant Functions . . . 116
7.6 Digit Functions . . . 118
7.7 Fuzzy Digit Functions . . . 121
7.8 Fuzzy Digit Functions from Parameter Estimation . . . 123
7.8.1 Bounding $f(p)$ away from 0 . . . 124
7.8.2 Bounding $f(p)$ away from 1 . . . 126
7.8.3 Continuity of $f(p)$ . . . 127
7.9 Quantum Speed-up for Bernoulli Factories . . . 130
7.9.1 Quantum Amplitude Estimation . . . 130
7.9.2 Amplitude Estimation and Bernoulli Factories . . . 131
7.10 Destructive Interference and the Quantum Advantage . . . 134
7.11 Outlook . . . 137

8 Kitaev-Webb Resampling **139**

8.1 Resampling Continuous Probability Distributions . . . 140
8.2 The Core KW Algorithm . . . 143
8.2.1 Fidelity of the KW Algorithm . . . 146
8.3 Implementing the KW Protocol . . . 148
8.4 Extending the KW Protocol to Non-linear Rescaling . . . 151
8.5 Conclusions & Outlook . . . 155

9 Concluding Remarks **156**

IV Appendices **161**

A Proof of geometric athermality monotonicity **162**

A.1 Virtual Qutrit Deformations 162

A.2 Monotonically Non-Increasing Deformations 165

A.3 Proving monotonicity 168

B Group Orbit 'Shapes' for Two Qubits under SU(2) **170**

B.1 Enumeration of the Isotropy Subgroups 171

B.1.1 $\mathbb{Z}_2 \prec \mathrm{Iso}(\rho)$ 171

B.1.2 $\mathbb{Z}_4 \prec \mathrm{Iso}(\rho)$ 172

B.1.3 $\mathrm{Iso}(\rho) = \mathbb{Z}_2$ 172

B.1.4 $U(1) \prec \mathrm{Iso}(\rho)$ 173

B.1.5 $K_2 \prec \mathrm{Iso}(\rho)$ 174

B.1.6 $\mathrm{Iso}(\rho) = \mathbb{Z}_4$ 175

B.1.7 $K_\infty \prec \mathrm{Iso}(\rho)$ 176

B.1.8 $\mathrm{Iso}(\rho) = \mathrm{K}_2$ 176

B.1.9 $\mathrm{Iso}(\rho) = \mathrm{U}(1)$ 176

B.1.10 $\mathrm{Iso}(\rho) = \mathrm{SU}(2)$ 176

B.1.11 $\mathrm{Iso}(\rho) = \mathrm{K}_\infty$ 177

Chapter 1

Introduction

It is becoming increasingly difficult to define what the field of quantum information is about. For much of its early development, quantum information theory focused on the consequences of quantum mechanics for the pre-existing notions of computation and communication [3–5]. While these remain central areas of research, the field has expanded into many other areas, from thermodynamics [6, 7] to quantum many-body physics [8] to high energy physics [9].

Nonetheless, a common theme or philosophy that permeates through many of these branches of quantum information theory is a focus on resources. These resources come in a number of flavours, ranging from practical resources that allow a certain physical operation to be performed in the laboratory, to the computational resources required to run a given quantum algorithm. This highlights the simple, operational approach that drives much of the field; it is about what we can do with quantum technologies given both practical and fundamental constraints, and what the minimal requirements are to achieve it.

This focus on resources has also been fruitful in understanding the fundamental features of quantum information and how it can be used in different scenarios. Entanglement was the first facet of quantum mechanics to be viewed from this perspective, characterised as that which could not be created through local operations and classical communication [10]. Since then, resource-led approaches have enhanced our understanding of quantum coherence [11], quantum

reference frames [12], and magic in quantum computation [13] to name a few.

This proliferation of resources in quantum information theory has motivated the study of the common structures and abstract features of resources. These have been termed 'Resource Theories' [14,15]. An abstract resource is defined by constraints; only certain operations are easy to implement, and these are called the *free operations*. This includes states that can be easily prepared, likewise called *free states*. While these mathematical structures are interesting in their own right, resource-theoretic approaches provide a standard framework to analyse practically important scenarios.

This thesis looks at several different aspects of resources in quantum information theory. We consider the abstract structures that arise in connection to resources in a variety of settings, and we also consider how they can be manipulated to implement important operations and algorithms.

1.1 The Geometry of Passivity

The first part of the thesis considers thermodynamic resources. Passivity is the property of not being able to extract work from a system. In other words, it describes whether a system is a useful resource. Furthermore, this property characterises Gibbs states – the states associated with systems in equilibrium, and the cornerstone of equilibrium statistical mechanics.

Our focus here is the structure of these thermodynamic resources. By recasting this physical property in geometric terms, the derivation of the Gibbs state becomes simple and clear, with many important consequences made obvious. Furthermore, this geometric framework introduces a natural measure of deviation from thermal equilibrium, offering a new non-equilibrium thermodynamic variable. This highlights the usefulness of the resource-led approach, highlighting new structures in well-established frameworks and enhancing our understanding of different areas of physics.

1.2 Resources from Symmetry Constraints

The second part of this thesis looks at resources arising from symmetry constraints. Symmetry principles are vital to diverse areas of physics, and they have important consequences for quantum information. This is a different type of resource to that considered in the first section, but one that has applications in quantum reference frames, thermodynamics and quantum computation.

Again the abstract structures of resources are our primary focus in this section, and we introduce a framework that classifies the different types of resources that can exist for a given symmetry constraint. However, we also consider how this classification framework provides coarse-grained information about dynamics in this context, and more importantly about the resources required to lift these constraints.

1.3 Resampling Algorithms for Quantum Computers

This moves us onto the final part of this thesis, which is concerned with the practical manipulation of computational resources, in particular for resampling algorithms on quantum computers. First we consider a family of algorithms called Bernoulli factories, studying the computational limitations of these schemes, and how quantum resources can lift those constraints. We also quantify the resources needed to perform a given Bernoulli factory algorithm. Our aim is to provide an intuitive understanding of how quantum resources enable a previously known quantum advantage.

Following on, we study a different type of resampling algorithm devised by Kitaev and Webb. Our aim here is to promote this as a useful primitive from which more complex algorithms can be built, making this algorithm a potentially useful resource for future quantum programmers. We analyse the fidelity of this algorithm for near-term implementations, and we explore how the scheme can be extended to non-linear resampling.

Part I

The Geometry of Passivity

Chapter 2

A planar and simple derivation of the Gibbs state

Since its inception in the development of steam engines, thermodynamics has been concerned with efficiently extracting useful work from systems, putting the analysis of resources at the heart of the subject. More recently, resource-led approaches have clarified many conceptual issues with quantum phenomena in the context of thermodynamics, most notably coherence. This chapter is concerned with one way to analyse work extraction from quantum systems. Centering on the notion of passivity, it characterises whether a state is a useful resource.

- After motivating passivity as a means to derive the Gibbs state, a new proof of the conditions for passivity in finite-dimensional quantum systems is given, based on bistochasticity.
- The bulk of the chapter builds up a novel geometric perspective on passivity and complete passivity for finite-dimensional quantum systems, viewing passivity as a total ordering of vectors in $\mathbb{R}^2$. Furthermore, many copies of a state are naturally represented by a convex shape called the asymptotic ϵ-s ensemble.
- Using this geometric framework, we present a simple and transparent derivation of the Gibbs state, and discuss some of its immediate consequences.

While the end result that the Gibbs states are the completely passive states is well known in the literature, the approach here is a novel one. The new contributions of this chapter are:

- In Section 2.2.2, a new proof of the necessary and sufficient conditions for a quantum state to be passive is given, using properties of majorization.
- A new description of passive states called the ϵ-s ensemble is introduced and passivity is recast as an ordering of vectors in $\mathbb{R}^2$, while also accommodating multiple copies of states in this novel framework.
- The asymptotic ϵ-s ensemble is defined and used to describe the many-copy limit and some basic properties of these objects are defined.
- Section 2.6 sets out a new geometric derivation of the Gibbs states as the only completely passive states, providing the transparency and intuition that was lacking from previous derivations. This reveals new results such as the fact that all passive qubit states are completely passive.

This chapter and the next are based on work presented in [2], done in collaboration with Rhea Alexander, Nick Koukoulekidis and David Jennings.

2.1 Motivation – Why Passivity?

Equilibrium statistical mechanics is a well-developed branch of physics with numerous applications. The lead character in this drama is the Gibbs state. One could simply posit this from the beginning, however it is desirable to have a firm foundation on which to justify its starring role. Several different approaches have been used to justify the Gibbs state, including:

- Micro-canonical ensemble derivation [16] – assuming equal a priori probabilities over energies for some large system, the marginal state of a small subsystem is found to be the Gibbs state.

- Maximum Entropy derivation [17, 18] – this explicitly approaches statistical mechanics as an inference problem. The Gibbs state is the maximally unbiased description of the system subject to an energy constraint.

- Dynamical derivations [19–21] – the dynamical evolution of a system is studied as it equilibrates, with a focus on thermalisation and typicality, studying how a notion of temperature emerges.

Each of these approaches emphasises different aspects of the physics. In this work, we focus on another path that takes a simple, operational approach to justifying the Gibbs state – passivity [22–24]. The Gibbs state is characterised by an inability to extract work from it, no matter how many copies of the system we have. In essence, the approach considers whether the state of the system is useful to us as a thermodynamic resource.

The simplicity of this approach is well-suited to exploring thermodynamics in quantum systems. It highlights the structural aspects of the theory, in particular the conceptual ingredients needed to establish a notion of equilibrium. Furthermore, this approach was originally developed in the context of statistical mechanics for infinite-dimensional quantum systems, therefore it is able to handle more complex scenarios where algebraic techniques are needed [25]. The utility of passivity is evident from the attention it has received in recent years, applied in studying work-extraction from non-equilibrium states [26–29], generalised Gibbs states [30–34] and quantum information-theoretic approaches to thermodynamics [7, 35–38].

Despite the simplicity of the core idea, early formulations of passivity in the infinite-dimensional setting were presented in the language of C^*-algebras and suffered from being abstract and inaccessible. Specialising to finite-dimensional quantum systems, Lenard provided a more user-friendly presentation of these ideas [23]. More recent work has framed passivity in terms of 'virtual temperatures' [24], however we desire a more intuitive picture of how the Gibbs state emerges from this simple physical principle.

This work follows the spirit of Lenard, emphasising clarity over abstraction. Using a novel geometric framework, we recast passivity as a total ordering of vectors. A simple geometric derivation of the Gibbs state follows naturally from this framework. We proceed from two core ingredients: passivity and extensivity.

2.2 Passivity

Passivity is a physical and operationally-motivated concept that is best illustrated through simple examples. Throughout this chapter, the average energy of a state ρ is given by $\langle H \rangle = \mathrm{tr}\big(H\rho\big)$.

Example 2.1: Passifying a state

Consider a three-level quantum system with Hamiltonian $H = |1\rangle\langle 1| + 2\ |2\rangle\langle 2| + 3\ |3\rangle\langle 3|$.

The average energy of the state

$$\rho = \frac{1}{4}\ |1\rangle\langle 1| + \frac{1}{4}\ |2\rangle\langle 2| + \frac{1}{2}\ |3\rangle\langle 3| \tag{2.1}$$

can be deterministically lowered by swapping levels 1 and 3, achieved by applying the unitary $U_1 = |1\rangle\langle 3| + |3\rangle\langle 1| + |2\rangle\langle 2|$. The average energy of the new state $\rho_p = U_1 \rho U_1^\dagger$ cannot be further lowered by any unitary.

Now consider the pure state

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle + |3\rangle), \tag{2.2}$$

which can be transformed to the ground state $|1\rangle$ by the unitary

$$U_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}, \tag{2.3}$$

where $\omega \coloneqq e^{\frac{2\pi i}{3}}$. Clearly, the average energy cannot be lowered any further.

Here, ρ_p and $|1\rangle$ are examples of passive states, their energy impossible to lower by unitary transformations.

2.2.1 Definition of Passivity

We now put this intuition on a formal footing, as originally defined in [22].

Definition 2.1: Passivity

A state ρ of a quantum system with Hamiltonian H is passive if $\mathrm{tr}\big(HU\rho U^\dagger\big) \geq \mathrm{tr}(H\rho)$ for all unitaries U acting upon the system.

To maintain mathematical simplicity and clarity, we consider only finite-dimensional quantum systems whose energies are bounded, but the definition extends to infinite-dimensional systems [22].

With this definition in mind, it is natural to partition the state space into unitary orbits, $\mathcal{M}(\rho) := \{U\rho U^\dagger : U \in U(d)\}$, sets of mutually accessible states under unitary operations. Passive states are the energetic minima of these orbits, and the definition guarantees at least one passive state for each orbit. Furthermore, passive states provide a notion of 'equilibrium' states, acting as natural representatives for each orbit. The orbit construction is a common way to analyse resources in quantum information theory, and we consider them extensively in the second part of this thesis.

At first sight, the definition of passivity has a similar flavour to the Kelvin-Planck formulation of the Second Law [39], which forbids work extraction from thermalised, adiabatically isolated systems. However, passivity does not invoke the full machinery of thermodynamics. In a sense prior to the thermodynamic framework, passivity is a basic physical statement about deterministic extraction of energy from a system. It remains valid beyond the usual equilibrium domain, making it well-suited to analysing the thermodynamics of individual quantum systems.

2.2.2 Conditions for Passivity

When can work be extracted from a quantum system? Necessary and sufficient conditions for a finite-dimensional quantum state to be passive were first derived in [23].

Theorem 2.1: Conditions for Passivity

Consider a d-dimensional quantum system with Hamiltonian H, with eigenvalue decomposition $H = \sum_{i=1}^{d} \epsilon_i \ |e_i\rangle \langle e_i|$. A state ρ of that system is passive iff

- $[\rho, H] = 0$ and,
- for eigenvalue decomposition $\rho = \sum_{i=1}^{d} p_i \ |e_i\rangle \langle e_i|$, then $\epsilon_i \leq \epsilon_j$ implies $p_i \geq p_j$ for all i and j in $\{1, \ldots, d\}$.

Firstly, note that a passive state is block-diagonal in the energy basis, eliminating the possibility of any coherence between different energy subspaces. The second condition pertains to the relative ordering between eigs(ρ) and eigs(H), and they should be 'anti-ordered' in a passive state. It is useful to introduce the notation $\boldsymbol{x}^{\uparrow}$ to denote that the entries of the vector $\boldsymbol{x}$ are in non-decreasing order. Similarly the entries of $\boldsymbol{x}^{\downarrow}$ are in non-increasing order.

Given that unitary transformations can be related to bistochastic operations on the diagonal entries of ρ, the following result is therefore important in the context of passivity.

Lemma 2.1: Dot products under ordering and bistochastic mixing

Given two real finite-dimensional vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, and a bistochastic matrix B,

$$\boldsymbol{x}^{\uparrow} \cdot \boldsymbol{y}^{\uparrow} \geq \boldsymbol{x}^{\uparrow} \cdot (B\boldsymbol{y})^{\uparrow} \geq \boldsymbol{x} \cdot (B\boldsymbol{y}) \geq \boldsymbol{x}^{\uparrow} \cdot (B\boldsymbol{y})^{\downarrow} \geq \boldsymbol{x}^{\uparrow} \cdot \boldsymbol{y}^{\downarrow} \ . \tag{2.4}$$

The outer inequalities in this lemma are proved in [40], while the inner inequalities derive from the rearrangement inequality [41]. We can now provide a new proof of the conditions for passivity using bistochasticity.

Proof. Without loss of generality, we can assume that $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\uparrow}$, since this amounts to a relabelling

of the energy basis.

First, assume that ρ satisfies $[\rho, H] = 0$ with eigenvalues anti-ordered with $\boldsymbol{\epsilon}$ (i.e. $\epsilon_i \leq \epsilon_j$ implies $p_i \geq p_j$). By assumption, ρ is diagonal in the energy basis, such that

$$\rho = \sum_{i=1}^{d} p_i \ |e_i\rangle \langle e_i| \tag{2.5}$$

with $p_1 \geq \cdots \geq p_d$, and the average energy is $\mathrm{tr}[H\rho] = \boldsymbol{\epsilon}^{\uparrow} \cdot \boldsymbol{p}^{\downarrow}$. The average energy of $U\rho U^{\dagger}$ is

$$\mathrm{tr}[HU\rho U^{\dagger}] = \sum_{i,j} \epsilon_i \ p_j \ |\langle e_i| U |e_j\rangle|^2 = \boldsymbol{\epsilon}^{\uparrow} \cdot (B\boldsymbol{p}^{\downarrow}) \ , \tag{2.6}$$

where B is a bistochastic matrix with elements $B_{ij} = |\langle e_i| U |e_j\rangle|^2$. However, Lemma 2.1 implies that this is greater than or equal to the average energy of ρ for all unitaries U acting upon the system, hence ρ is passive.

Conversely, assume that ρ is passive. Expressing the state in the energy basis,

$$\rho = \sum_{i,j} \rho_{ij} \ |e_i\rangle \langle e_j| \ , \tag{2.7}$$

which gives an average energy

$$\mathrm{tr}[H\rho] = \sum_{i} \epsilon_i \rho_{ii} = \boldsymbol{\epsilon}^{\uparrow} \cdot \mathrm{diag}(\rho) \ . \tag{2.8}$$

By the Schur-Horn theorem [42], $\mathrm{diag}(\rho) \prec \mathrm{eigs}(\rho) = \boldsymbol{p}^{\downarrow}$. This implies that there exist a series of 'Robin Hood' transformations T_i (bistochastic transformations that only act non-trivially on a two-dimensional subspace) [42], such that

$$\mathrm{diag}(\rho) = T_{d-1} \ldots T_1 \boldsymbol{p}^{\downarrow}. \tag{2.9}$$

Lemma 2.1 implies that $\mathrm{tr}[H\rho] = \boldsymbol{\epsilon}^{\uparrow} \cdot \mathrm{diag}(\rho) \geq \boldsymbol{\epsilon}^{\uparrow} \cdot \boldsymbol{p}^{\downarrow}$. Moreover, there exists a unitary transformation that diagonalises ρ, therefore $\boldsymbol{\epsilon}^{\uparrow} \cdot \mathrm{diag}(\rho) \leq \boldsymbol{\epsilon}^{\uparrow} \cdot \boldsymbol{p}^{\downarrow}$ because ρ is passive, hence $\boldsymbol{\epsilon}^{\uparrow} \cdot \mathrm{diag}(\rho) = \boldsymbol{\epsilon}^{\uparrow} \cdot \boldsymbol{p}^{\downarrow}$. Furthermore, this must be true regardless of the basis that diagonalises ρ, therefore

$$\boldsymbol{\epsilon}^{\uparrow} \cdot (\mathbb{I} - T)\boldsymbol{p}^{\downarrow} = 0 \ . \tag{2.10}$$

for all Robin Hood transformations T. For T acting on the i and j components,

$$(1 - T_{ii})\epsilon_i p_i + (1 - T_{jj})\epsilon_j p_j - T_{ij}\epsilon_i p_j - T_{ji}\epsilon_j p_i = 0 \ , \tag{2.11}$$

which simplifies to

$$T_{ij}(\epsilon_j - \epsilon_i)(p_j - p_i) = 0 \tag{2.12}$$

because T is bistochastic. There are two scenarios to consider:

- $\epsilon_i \neq \epsilon_j$. If $p_i = p_j$, then there is no constraint on T, but if $p_i \neq p_j$, then $T_{ij} = 0$, which implies $T = \mathbb{I}$.

- $\epsilon_i = \epsilon_j$. There are no further constraints on T other than it is bistochastic.

The corresponding unitaries cannot mix $\boldsymbol{p}^{\downarrow}$ across different energy subspaces. This proves that ρ is block diagonal in the energy basis, hence $[\rho, H] = 0$, and that $\epsilon_i \leq \epsilon_j$ implies $p_i \geq p_j$. □

Using bistochasticity in this proof emphasises the connection between passivity and ordering. This link will be expanded upon later in the chapter when we recast passivity as a total ordering of vectors. We also note in passing that Lemma 2.1 also implies that the most active states have $[\rho, H] = 0$ and $\boldsymbol{p}$ ordered with respect to the energy. Therefore, the maximal work (ergotropy) that can be extracted from the state ρ is $\boldsymbol{\epsilon}^{\uparrow} \cdot (\boldsymbol{p}^{\uparrow} - \boldsymbol{p}^{\downarrow})$.

2.3 Extensivity

Passivity alone does not specify a Gibbs state. A further ingredient can be motivated by returning to the definition of passivity, which was given relative to the Hamiltonian of the system. Why not choose a different observable to H? This arises in the context of generalised Gibbs states, where non-commuting observables play a non-trivial role in the thermodynamics of the system [30–34]. Different observables relate to other types of resources, not just those from which work can be reversibly extracted.

Even restricting to energetic considerations, why not define passivity with respect to the second moment, $\langle H^2 \rangle$, or higher moments? If we do, the same set of states as the original definition are again picked out as passive. In fact, the same passive states are found whenever passivity is defined relative to $\langle f(H) \rangle$ for any monotonically increasing function f, because the conditions in Theorem 2.1 only depend on the relative ordering of the eigenvalues, as emphasised in our majorization-based proof.

An analogous issue arises in other derivations of the Gibbs state, including the MaxEnt approach [17,18], where entropy is maximised given a fixed $\langle H \rangle$ to obtain a state in Gibbs form,

$$\gamma_\beta = \frac{1}{Z_\beta} e^{-\beta H}. \tag{2.13}$$

Maximising the entropy with respect to a fixed value of $\langle H^2 \rangle$ instead gives

$$\gamma'_\beta \propto e^{-\beta H^2}, \tag{2.14}$$

certainly not a Gibbs state.

The extra physical ingredient that is required to single out $\langle H \rangle$ and derive the Gibbs state is *extensivity*. This relates to how observables behave when systems combine. It is clear that H is an extensive observable, while other powers of H are not. Physically, this is because energy is additively conserved microscopically.

The issues become apparent when considering multiple systems. Consider two systems, each in a passive state for their respective Hamiltonians. In general, the composite system is not guaranteed to be passive under the combined Hamiltonian. The number of copies needed to extract work from a state quantifies how useful it is as a work resource, motivating the following definition [23].

Definition 2.2: k-Passivity and Complete Passivity

A quantum state ρ of a system with Hamiltonian H is k-passive if $\rho^{\otimes k}$ is passive under $\sum_{i=1}^{k} H_i$. If a state is k-passive for all integers $k \geq 1$, it is completely passive.

For finite-dimensional systems, the completely passive states are essentially the Gibbs states [22, 23]. We present a novel geometric proof of this well-known result that makes this conclusion transparent and inevitable, using the framework introduced in the next few sections.

2.4 A Geometric Perspective on Passivity

At the heart of our framework is the ϵ-s ensemble, which we now define. Crucially, passivity becomes a simple geometric criterion.

2.4.1 The ϵ-s Ensemble

Consider a finite-dimensional quantum system with Hamiltonian H, and a passive state ρ of that system. For simplicity, we assume that H is non-degenerate, leaving discussion of degenerate energy spectra until Section 2.7. In light of Theorem 2.1, we need only consider the eigenvalues $(p_i) = \text{eigs}(\rho)$ and $(\epsilon_i) = \text{eigs(H)}$. All relevant information about the system and its state is encoded in these two sets of eigenvalues. A state can then be represented as a set of points in $\mathbb{R}^2$, using the pairs $\{(\epsilon_i, p_i) : 1 \leq i \leq d\}$.

Unfortunately, this representation is ill-suited to handling the composition of systems. The central problem is that the energies and probabilities are treated differently when systems combine. Energy is additive, whereas probabilities multiply. Taking additivity seriously, a more natural representation substitutes p_i with $s_i := -\log p_i$, a "single instance" entropy. No information about the system is lost, however manipulations become much simpler, harmonising with vector addition in $\mathbb{R}^2$. This new representation is formally defined as follows.

Definition 2.3: ϵ-s Ensemble

Given a finite-dimensional quantum system with Hamiltonian $H = \sum_{i=1}^{d} \epsilon_i |e_i\rangle \langle e_i|$, for any full-rank state $\rho = \sum_{i=1}^{d} p_i |e_i\rangle \langle e_i|$, the associated ϵ-s ensemble is $\mathcal{V}(\rho) := \{\boldsymbol{v}_i : 1 \leq i \leq d\}$, where $\boldsymbol{v}_i := (\epsilon_i, s_i)$.

Technically, this definition only applies to full-rank states, preventing $s_i \to \infty$ as $p_i \to 0$. Nonetheless, s_i can be made arbitrarily large ($p_i > 0$ can be arbitrarily small). The restriction to full-rank states does not constrain the physics at all. Any lower rank state σ is physically indistinguishable from the full rank state $(1-\delta)\,\sigma + \delta\,\mathbb{I}/d$ for sufficiently small $\delta > 0$.

The name "ϵ-s ensemble" alludes to the physical interpretation of this set. Weighting the elements of $\mathcal{V}(\rho)$ by the corresponding probabilities p_i, the expectation values of the two components are the average energy and von Neumann entropy of ρ. We will explore this further in Section 3.1.

2.4.2 Passivity as a Total Order on Vectors

The importance of this geometric representation is the simple form of the passivity conditions, which become a purely geometric property in $\mathbb{R}^2$, stated as a total ordering for vectors.

Definition 2.4: Total Ordering of Vectors

Two vectors $\boldsymbol{r}_i \coloneqq (x_i, y_i)$ and $\boldsymbol{r}_j \coloneqq (x_j, y_j)$ in $\mathbb{R}^2$ are ordered $\boldsymbol{r}_i \leq \boldsymbol{r}_j$ if $x_i \leq x_j$ and $y_i \leq y_j$. Moreover, a set $R = \{\boldsymbol{r}_1, \boldsymbol{r}_2, \dots\}$ of vectors in $\mathbb{R}^2$ is totally ordered if, for any two $\boldsymbol{r}_i, \boldsymbol{r}_j \in R$, then either $\boldsymbol{r}_i \leq \boldsymbol{r}_j$ or $\boldsymbol{r}_j \leq \boldsymbol{r}_i$.

This allows us to reformulate the necessary and sufficient conditions for a state to be passive in a particularly simple form.

Lemma 2.2: Passivity as a Total Order

For a d-dimensional quantum system with Hamiltonian H, a state ρ satisfying $[\rho, H] = 0$ is passive iff its ϵ-s ensemble $\mathcal{V}(\rho)$ is totally ordered.

Proof. This is a direct consequence of Theorem 2.1, since $-\log p_i \leq -\log p_j$ iff $p_i \geq p_j$. □

Although this lemma merely recasts Theorem 2.1 in a new framework, this result is at the heart of our derivation of the Gibbs state and its consequences. The simplicity of this geometric

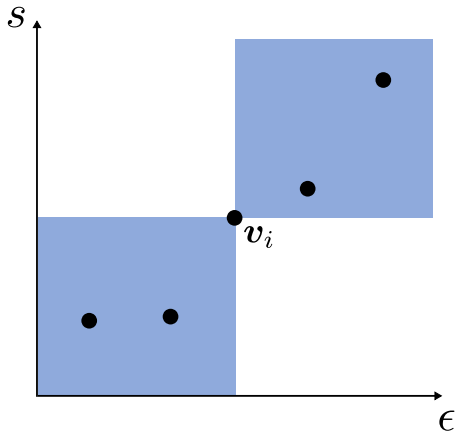


Figure 2.1: Passivity $\iff$ totally ordered ϵ-s ensemble. *For any $\boldsymbol{v}_i \in \mathcal{V}(\rho)$, the remaining points of the ϵ-s ensemble must lie in the shaded regions for the state to be passive. Reproduced from [2].*

condition is illustrated in Figure 2.1, with passivity determined by the relative positions of the ϵ-s ensemble elements.

2.5 A Geometric Perspective on Complete Passivity

Passivity concerns an individual system, while k-passivity and complete passivity shift the focus towards multiple independent, identically prepared systems. Our geometric framework again provides a novel perspective on these properties.

2.5.1 k-Passivity and the Regularised ϵ-s Ensemble

The ϵ-s ensemble naturally handles multiple systems because it respects additivity. Specifically, given states ρ and σ with ϵ-s ensembles $\mathcal{V}(\rho)$ and $\mathcal{V}(\sigma)$ respectively, the state $\rho \otimes \sigma$ has ϵ-s ensemble

$$\mathcal{V}(\rho \otimes \sigma) = \{\boldsymbol{v}_i + \boldsymbol{w}_j : \boldsymbol{v}_i \in \mathcal{V}(\rho), \boldsymbol{w}_j \in \mathcal{V}(\sigma)\}, \tag{2.15}$$

where $+$ denotes usual vector addition in $\mathbb{R}^2$. This mathematical structure is called the Minkowski sum [43], which we denote $\oplus$, therefore $\mathcal{V}(\rho \otimes \sigma) = \mathcal{V}(\rho) \oplus \mathcal{V}(\sigma)$.

To prevent the graphical representation becoming unwieldy, it is beneficial to scale the elements of $\mathcal{V}(\rho)$ by the number of copies of the system, regularising the ϵ-s ensemble.

Definition 2.5: Regularised ϵ-s ensemble

For a state ρ of a d-dimensional system with ϵ-s ensemble $\mathcal{V}(\rho) = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d\}$, the regularised ϵ-s ensemble for k independent identically prepared copies of the system is

$$\mathcal{V}_k(\rho) := \frac{1}{k}\mathcal{V}(\rho^{\otimes k}) = \left\{ \frac{1}{k} \sum_{i=1}^{d} c_i \, \boldsymbol{v}_i \;:\; c_i \in \mathbb{N},\ \sum_{i=1}^{d} c_i = k \right\}. \tag{2.16}$$

Scaling by the number of copies means that $\mathcal{V}_k(\rho)$ is constructed by taking convex combinations of the original ensemble, as shown by the following example.

Example 2.2: Additivity for ϵ-s ensembles

A qubit system ($d = 2$) has ϵ-s ensemble $\mathcal{V}(\rho) = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. The regularised ϵ-s ensemble for two copies of the state is

$$\frac{1}{2}\mathcal{V}(\rho^{\otimes 2}) = \left\{ \boldsymbol{v}_1, \frac{1}{2}(\boldsymbol{v}_1 + \boldsymbol{v}_2), \boldsymbol{v}_2 \right\}, \tag{2.17}$$

and for three copies,

$$\frac{1}{3}\mathcal{V}(\rho^{\otimes 3}) = \left\{ \boldsymbol{v}_1, \frac{1}{3}(\boldsymbol{v}_1 + 2\boldsymbol{v}_2), \frac{1}{3}(2\boldsymbol{v}_1 + \boldsymbol{v}_2), \boldsymbol{v}_2 \right\}. \tag{2.18}$$

Note how the scaling ensures the points are convex combinations of the original elements in $\mathcal{V}(\rho)$.

We give further examples for a qutrit in Figure 2.2. Importantly, k-passivity for quantum states becomes a simple geometric criterion.

Lemma 2.3: k-passivity as a total order

A state ρ of a d-dimensional quantum system with Hamiltonian H is k-passive iff $[\rho, H] = 0$ and $\mathcal{V}_k(\rho)$ is totally ordered.

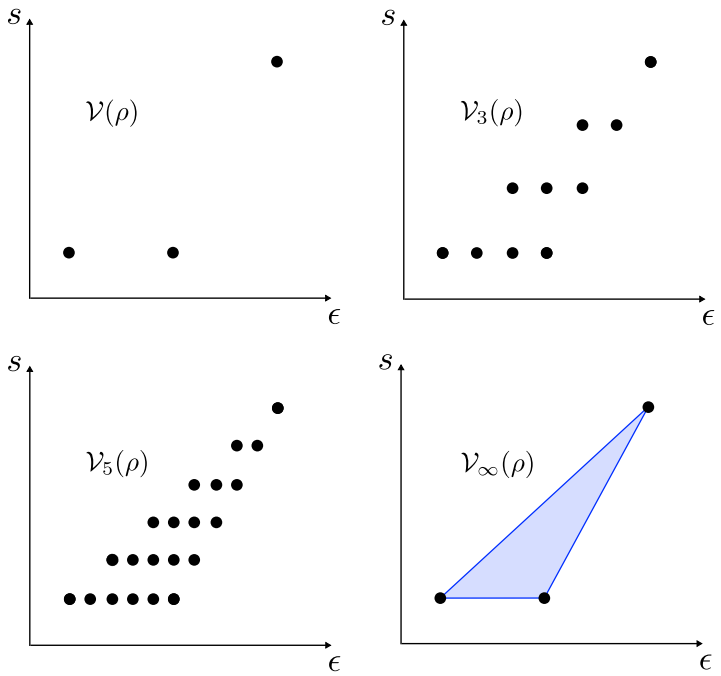


Figure 2.2: Regularised ϵ-s ensembles for multiple copies of a passive qutrit state ρ. *Increasing the number of copies of the system increases the number of points in $\mathcal{V}_k(\rho)$, taking convex combinations of the single copy ensemble $\mathcal{V}(\rho)$ (top left) and tending towards the asymptotic version $\mathcal{V}_\infty(\rho)$ (bottom right, and introduced in the next section). Reproduced from [2].*

Proof. This is Lemma 2.2 applied to the ϵ-s ensemble of k copies of the state. □

This new compact characterisation of k-passivity opens up scope for further study. The mathematical construction of the regularised ϵ-s ensemble $\mathcal{V}_k(\rho)$ features in additive combinatorics, where it is termed the k'th sumset [44]. It remains an open question whether results in additive combinatorics can be applied to the study of k-passivity.

2.5.2 Asymptotic ϵ-s Ensemble

Complete passivity concerns arbitrary numbers of copies, extending k-passivity over all $k \in \mathbb{N}$. We therefore define an asymptotic version of the ϵ-s ensemble, taking the limit of $\mathcal{V}_k(\rho)$ as $k \to \infty$.

Definition 2.6: Asymptotic ϵ-s Ensemble

The asymptotic ϵ-s ensemble $\mathcal{V}_\infty(\rho)$ of a system in the state ρ, with Hamiltonian H, is the set of limit points $\boldsymbol{x}$ of sequences $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots)$, where $\boldsymbol{x}_k \in \mathcal{V}_k(\rho)$, with convergence to $\boldsymbol{x}$ under the standard vector norm on $\mathbb{R}^2$.

This technical definition can be viewed more intuitively by noting that points in the regularised ensemble are convex combinations of the single-copy ensemble $\mathcal{V}(\rho)$, made clear by expressing $\mathcal{V}_k(\rho)$ as

$$\mathcal{V}_k(\rho) = \left\{ \sum_{i=1}^{d} q_i \, \boldsymbol{v}_i \ : \ q_i k \in \mathbb{Z}_k \ , \ \sum_{i=1}^{d} q_i = 1 \right\}, \tag{2.19}$$

where $\boldsymbol{v}_i \in \mathcal{V}(\rho)$. This set ranges over all probability distributions $\{q_i\}$ with d rational components $q_i \in \mathbb{Z}_k/k$. As $k \to \infty$, these probabilities tend towards non-negative real numbers. The following lemma makes the intuition of Figure 2.2 rigorous, offering a simple characterisation of the asymptotic ensemble.

Lemma 2.4: Asymptotic ensemble as the convex hull of $\mathcal{V}(\rho)$

For a d-dimensional system with Hamiltonian H, and state ρ with $[\rho, H] = 0$ and ϵ-s ensemble $\mathcal{V}(\rho)$,

$$\mathcal{V}_\infty(\rho) = \mathrm{conv}\big[\mathcal{V}(\rho)\big], \tag{2.20}$$

where $\mathrm{conv}\big[\mathcal{V}(\rho)\big]$ denotes the convex hull of $\mathcal{V}(\rho)$.

Proof. First, we show that $V_\infty(\rho) \subseteq \mathrm{conv}\big[\mathcal{V}(\rho)\big]$. Definition 2.6 specifies that for a given $\boldsymbol{x} \in \mathcal{V}_\infty(\rho)$, there exists a sequence of vectors $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots)$ where $\boldsymbol{x}_k \in \mathcal{V}_k(\rho)$, such that for any $\delta > 0$ there is an $N \in \mathbb{N}$ such that $||\boldsymbol{x} - \boldsymbol{x}_k|| \leq \delta$ for all $k \geq N$. Definition 2.5 implies that $\boldsymbol{x}_k \in \mathrm{conv}[\mathcal{V}(\rho)]$ for all k. For a finite-dimensional system, $|\mathcal{V}(\rho)|$ is finite, therefore $\mathrm{conv}[\mathcal{V}(\rho)]$ is a closed compact set in $\mathbb{R}^2$, which implies that the limit point $\boldsymbol{x} \in \mathrm{conv}[\mathcal{V}(\rho)]$, hence $\mathcal{V}_\infty(\rho) \subseteq \mathrm{conv}[\mathcal{V}(\rho)]$.

Now we prove the converse, namely that $\mathrm{conv}[\mathcal{V}(\rho)] \subseteq \mathcal{V}_\infty(\rho)$. Given $\boldsymbol{x} \in \mathrm{conv}[\mathcal{V}(\rho)]$, such

that $\boldsymbol{x} = \sum_{i=1}^{d} q_i \boldsymbol{v}_i$, where $q_i \in [0,1]$ for all i and $\sum_{i=1}^{d} q_i = 1$, we construct a sequence of vectors $\boldsymbol{x}_k = \sum_{i=1}^{d} c_i \boldsymbol{v}_i$ with coefficients

$$c_i = \begin{cases} \frac{\lfloor q_i k \rfloor}{k}, & i = 1, \ldots, d-1 \\ 1 - \sum_{j=1}^{d-1} \frac{\lfloor q_j k \rfloor}{k}, & i = d, \end{cases} \tag{2.21}$$

where $\lfloor \cdot \rfloor$ is the floor function. Since $\sum_{i=1}^{d} c_i = 1$ and $c_i k \in \mathbb{Z}_k$ for all $i = 1, \ldots, d$, we know that $\boldsymbol{x}_k \in \mathcal{V}_k(\rho)$ for all $k \in \mathbb{N}$. To prove that the sequence converges to $\boldsymbol{x}$, let $q_i = \sum_{j=1}^{\infty} d_{i,j} 10^{-j}$ be the decimal expansion of the i-th coefficient of $\boldsymbol{x}$. For a $\delta > 0$, choose an integer $N = 10^n$, where $n \in \mathbb{N}$ and $n > \max\{-\log \delta, 0\}$. Then, for any $k \geq N$,

$$\left| \frac{\lfloor q_i k \rfloor}{k} - q_i \right| = q_i - \frac{\lfloor q_i k \rfloor}{k} \qquad \left(\lfloor q_i k \rfloor \leq q_i k \right) \tag{2.22}$$

$$\leq q_i - \sum_{j=1}^{n} d_{i,j} 10^{-j} \qquad \left(k \geq N \right) \tag{2.23}$$

$$= \sum_{j=n+1}^{\infty} d_{i,j} 10^{-j} \qquad \left(\text{decimal expansion of } q_i \right) \tag{2.24}$$

$$< 10^{-n} < \delta \qquad \left(-n < \min\{\log \delta, 0\} \leq \log \delta \right). \tag{2.25}$$

The coefficient c_i converges to q_i, including $c_d \to 1 - \sum_{j=1}^{d-1} q_j = q_d$, therefore $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots)$ converges to $\boldsymbol{x}$ under the standard vector norm, hence $\mathrm{conv}[\mathcal{V}(\rho)] \subseteq \mathcal{V}_\infty(\rho)$.

We have shown that $V_\infty(\rho) \subseteq \mathrm{conv}\big[\mathcal{V}(\rho)\big] \subseteq V_\infty(\rho)$, therefore $V_\infty(\rho) = \mathrm{conv}\big[\mathcal{V}(\rho)\big]$. □

2.5.3 Some Properties of the Asymptotic Ensemble

Figure 2.3 provides some simple examples of asymptotic ϵ-s ensembles, emphasising the connection between $\mathcal{V}_\infty(\rho)$ and $\mathrm{conv}\big[\mathcal{V}(\rho)\big]$. The structure of a passive state is therefore described by a convex polygon in the $\mathbb{R}^2$ plane.

The geometric properties of the asymptotic ensemble are encoded in the elements of $\mathcal{V}(\rho)$. Consider a d–dimensional system with Hamiltonian H, with maximum and minimum energies $\epsilon_{\max} := \max \mathrm{eigs}(H)$ and $\epsilon_{\min} := \min \mathrm{eigs}(H)$ respectively. We note the following features of asymptotic ensembles.

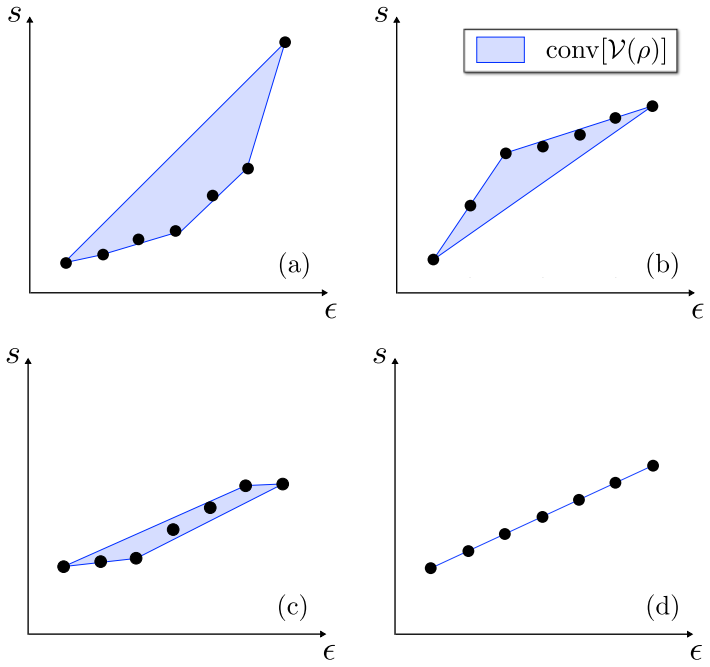


Figure 2.3: Examples of passive asymptotic ϵ-s ensembles. *These totally ordered ϵ-s ensembles $\mathcal{V}(\rho) = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_7\}$ correspond to four possible passive states of a 7-d quantum system with Hamiltonian $H = \mathrm{diag}\{1, 2, \ldots, 7\}$. The blue region is the convex hull of $\mathcal{V}(\rho)$. We shall prove shortly that if the area of the convex hull is zero (as in (d)), then ρ is completely passive. Reproduced from [2].*

1. $\mathcal{V}_1(\rho) \subseteq \mathcal{V}_k(\rho) \subset \mathcal{V}_\infty(\rho)$ for all $k \geq 1$, and the closure of $\cup_{k=1}^{\infty} \mathcal{V}_k(\rho)$ equals $\mathcal{V}_\infty(\rho)$. However, it is not generally true that $\mathcal{V}_j(\rho) \subset \mathcal{V}_k(\rho)$ for $j < k$, e.g. $\mathcal{V}_2(\rho) \nsubseteq \mathcal{V}_3(\rho)$.

2. The asymptotic ensemble $\mathcal{V}_\infty(\rho)$ of a passive state ρ has:

   - upper boundary: piece-wise linear concave function $f_{\mathrm{upper}}(\epsilon)$ on $[\epsilon_{\min}, \epsilon_{\max}]$,
   - lower boundary: piece-wise linear convex function $f_{\mathrm{lower}}(\epsilon)$ on $[\epsilon_{\min}, \epsilon_{\max}]$.

   Although any piece-wise linear convex or concave function on a region can be used to (non-uniquely) define a passive state, not all convex shapes constructed by such linear piece-wise functions will be the asymptotic ensemble for some passive state – a simple counter-example is a square.

3. When composing systems, the asymptotic ensembles combine according to

$$\mathcal{V}_\infty(\rho \otimes \sigma) = \text{conv}\left[\mathcal{V}(\rho \otimes \sigma)\right] = \text{conv}\left[\mathcal{V}(\rho) \oplus \mathcal{V}(\sigma)\right] \tag{2.26}$$

$$= \text{conv}[\mathcal{V}(\rho)] \oplus \text{conv}[\mathcal{V}(\sigma)] = \mathcal{V}_\infty(\rho) \oplus \mathcal{V}_\infty(\sigma), \tag{2.27}$$

respecting the Minkowski sum.

2.6 A geometric derivation of the Gibbs state

The geometric framework that we have built up is the setting for our derivation of the Gibbs state, at the heart of which is the following result.

Lemma 2.5: Total ordering for a convex set implies colinearity

A convex set $\mathcal{C}$ is totally ordered in $\mathbb{R}^2$ iff its points are colinear with non-negative slope.

Proof. By inspection, a convex set of colinear points in $\mathbb{R}^2$ with non-negative slope is totally ordered. Conversely, consider a totally ordered convex set $\mathcal{C}$ in $\mathbb{R}^2$. Suppose there exist three points $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \mathcal{C}$ that are not colinear, just as in Figure 2.4. By convexity, the triangular region formed by their convex hull also lies within $\mathcal{C}$. Furthermore, any circle inscribed in this triangle is also in $\mathcal{C}$. As illustrated in Figure 2.4, a circle in the plane cannot be totally ordered, thereby contradicting the initial assumption. To avoid this contradiction, all points of $\mathcal{C}$ must be colinear with non-negative gradient. □

With this result in hand, the proof of the following well-known theorem [22, 23] becomes straightforward.

Theorem 2.2: Complete passivity of Gibbs states

A state ρ of a d-dimensional quantum system with Hamiltonian H is completely passive iff it is a Gibbs state $\rho = e^{-\beta H}/Z_\beta$ for some $\beta \geq 0$.

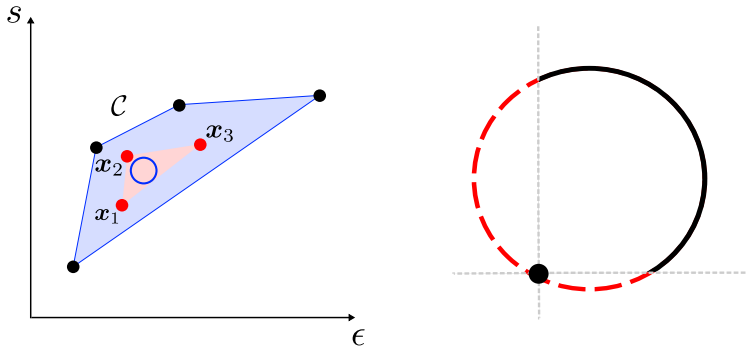


Figure 2.4: Proof that points in a totally ordered convex set are colinear. *The three points in $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \mathcal{C}$ form a triangle, inside which a circle is inscribed. The circle cannot be totally ordered in $\mathbb{R}^2$, because the dashed red points are not totally ordered with respect to the black point (as seen on the right). Reproduced from [2].*

Proof. It is easily verified that a Gibbs state with $\beta \geq 0$ has an ϵ-s ensemble that is colinear with non-negative slope. Lemma 2.5 implies this state is completely passive.

Conversely, if a state ρ of a d-dimensional system is completely passive, it is passive for all $k \geq 1$, thus $\mathcal{V}_k(\rho)$ must be totally ordered for all $k \geq 1$, which implies that $\mathcal{V}_\infty(\rho)$ is also totally ordered. However, $\mathcal{V}_\infty(\rho)$ is convex, therefore Lemma 2.5 implies its elements must be colinear with non-negative slope. Using the standard equation for a straight line, each $\boldsymbol{v}_i \in \mathcal{V}(\rho)$ can be written $\boldsymbol{v}_i = (\epsilon_i \ , \ \beta\epsilon_i - \log Z_\beta)$ for constants Z_β and $\beta \geq 0$. This implies that $p_i = e^{-\beta\epsilon_i}/Z$, thus ρ is a Gibbs state. □

This geometric derivation makes obvious some simple consequences.

Corollary 2.1: Qubit passivity

A passive qubit state is also completely passive.

This is because the convex hull of two totally ordered points is trivially colinear. Only systems with $d \geq 3$ have states that are passive but not completely passive. A further corollary is also obvious in our framework.

Corollary 2.2: Completely passive $\iff$ Zero Area

A passive state ρ of a d-dimensional system with Hamiltonian H is completely passive iff $\mathcal{V}(\rho)$ is totally ordered and the area of $\mathcal{V}_\infty(\rho)$ is zero.

The area of $\mathcal{V}_\infty(\rho)$ is a witness of athermality, and we shall explore this further in the next chapter.

2.6.1 Inverse Temperature as Gradient

For a completely passive state, the points of the ϵ-s ensemble reside on a straight line characterised by its gradient (the intercept with the s-axis is determined by normalisation). This gradient has a natural physical interpretation – the inverse temperature β. This connection extends to all passive states in our framework via virtual temperatures [24].

Definition 2.7: Virtual Temperature

The virtual temperature between the elements $\boldsymbol{v}_i, \boldsymbol{v}_j \in \mathcal{V}(\rho)$ is defined as

$$\beta_{i,j} := \frac{s_j - s_i}{\epsilon_j - \epsilon_i} = \frac{1}{\epsilon_j - \epsilon_i} \log\left(\frac{p_i}{p_j}\right), \tag{2.28}$$

the gradient of the line joining those points.

Virtual temperatures $\beta_{i,j}$ specify a passive state, while a completely passive state is specified by a single temperature β. This suggests that if the state has several different virtual temperatures, a heat engine can be operated between them and work can be extracted [27]. In contrast, all pairs of energy levels in a completely passive state have the same virtual temperature, therefore no work can be extracted.

Thinking of temperature as a gradient in our geometric framework agrees with the usual thermodynamic temperature. The joint state of two systems, each in completely passive states characterised by inverse temperature β, is also completely passive at the same temperature, because points on the same line are closed under vector addition. As expected, two systems at

the same temperature are in thermal equilibrium.

However, the notion of temperature as gradient is advantageous when considering individual quantum systems. Single finite-dimensional quantum systems in completely passive states can be sensibly ascribed a temperature, even an individual qubit. This bypasses many conceptual challenges faced by other approaches.

2.7 Degeneracy, Structural Stability and Ground States

Throughout this chapter, we have assumed non-degenerate energy spectra. However, our geometric framework can easily accommodate degeneracies, and the derivation of the Gibbs state remains practically unchanged. This section discusses the subtleties that arise when degeneracies are included.

Earlier, the state space was partitioned into unitary orbits, each with a representative passive state. With degeneracies, an orbit can contain multiple passive states – accessed by permuting the eigenvalues in the degenerate subspace. However, we are free to choose a particular passive state as the representative of that orbit (for example, we take eigs(ρ) in non-increasing order).

For the ϵ-s ensemble, there is some freedom in how to represent degeneracies. Degenerate levels with equal occupations will lead to repeats in the ϵ-s ensemble. Defined as a set, $\mathcal{V}(\rho)$ cannot encode this. Instead, we redefine the ϵ-s ensemble as a multiset – a simple extension of the set concept that allows repeated elements [45]. The rest of the framework remains unchanged, with Minkowski addition working as before to handle the composition of systems.

Although the geometric derivation of the Gibbs state is unchanged, there are a number of reasons to modify our definition of the ϵ-s ensemble to retain such information. Firstly, when considering unitary orbits as equivalence classes, the representative passive state characterises the eigenvalues eigs(ρ). Multisets preserve the information about the orbit fully, even if it has little consequence for the geometric perspective on passivity.

A more physical motivation for switching to multisets is their ability to adapt to perturba-

tions that break the degeneracy of the energy spectrum. Small changes to the system should not greatly affect the properties of the state, motivating the following property that captures the physical robustness of the mathematical framework [46].

Definition 2.8: Structural stability

Consider a system with Hamiltonian H in a passive state ρ. The state ρ is structurally stable if, given any small perturbation δH to the Hamiltonian, there exists a state $\rho' = \rho + \delta\rho$, in an arbitrarily small neighbourhood of ρ, that is passive under the new Hamiltonian $H' = H + \delta H$.

Structural stability imposes the additional requirement on passive states that the eigenvalues of ρ should be equal within a degenerate subspace, necessary to guarantee another passive state in the neighbourhood of ρ for the perturbed Hamiltonian. Pleasingly, structural stability follows as a consequence of complete passivity [23], and our geometric framework makes it clear why this is the case.

Lemma 2.6: Complete passivity $\implies$ structural stability

A completely passive state that is not a ground state is structurally stable.

Proof. Consider a completely passive state ρ that is not a ground state (ρ has non-zero eigenvalues outside the lowest energy subspace). Assume that ρ has unequal eigenvalues within a degenerate energy subspace, which leads to a non-zero area for $\mathcal{V}_\infty(\rho)$, as illustrated in Figure 2.5. However, this implies that $\mathcal{V}_\infty(\rho)$ cannot be totally ordered, hence it cannot be completely passive. This contradiction proves that the eigenvalues of a completely passive state ρ must be equal within a degenerate energy subspace, which in turn implies structural stability. □

There is a caveat to this for ground states, namely that all ground states are clearly completely passive, even if they have unequal eigenvalues. The ϵ-s ensemble for a ground state is always totally ordered. Therefore *complete passivity does not imply structural stability for*

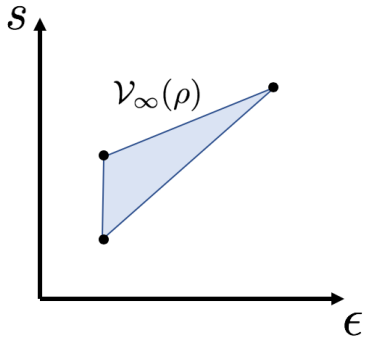


Figure 2.5: Completely passive states have uniform eigenvalues within a energy subspace. *The unequal eigenvalues in the degenerate energy subspace of this simple qutrit example break the total ordering of $\mathcal{V}_\infty(\rho)$, contradicting the assumption of complete passivity. Reproduced from [2].*

ground states. Since we should like equilibrium states to be structurally stable, we must additionally impose it for ground states, while it comes for free with complete passivity for all other states.

2.8 Summary and Outlook

In this chapter, we introduced a simple geometric framework for the study of passivity and complete passivity, recasting physical notions as simple planar geometric properties. This enabled a simple and transparent derivation of the Gibbs state, with several consequences made apparent by the geometric perspective. Furthermore, we highlighted that the asymptotic ϵ-s ensemble associated to a completely passive state had zero area – its elements were colinear.

Although many other derivations of the Gibbs state exist in the literature, the novel contribution of this chapter is to provide a new perspective, a geometric framework that makes reasoning about passivity and complete passivity clear and simple. Unlike previous derivations based upon passivity, the identification of Gibbs states as the completely passive states follows almost immediately, where before the path was not straightforward.

Another novel contribution of this chapter has been to emphasise the connection between passivity and ordering. The necessary and sufficient conditions for passivity of quantum states were already known in the literature, but we provide a new proof based on majorization and ordering. Furthermore, the geometric framework we develop here recasts passivity as a total ordering of vectors in the plane. The theme of ordering is deeply intertwined with the study of resources and reappears in later chapters of this thesis.

The new geometric framework makes obvious other new results along the way, such as the fact that all passive qubit states are completely passive. It also has the potential to greatly simplify the analysis of 'activating' states, showing how many copies are needed before work can be extracted. When compared with more algebraic approaches to this problem, such as [28], the ϵ-s ensemble handles this problem more elegantly and transparently. Although not addressed here, it would be of interest to explore whether techniques from additive combinatorics shed further light on the structure of k-passivity when applied to our geometric framework.

Here we restricted attention to simple systems described by energies and populations, but more general systems will be described by several variables. If these variables are commuting we expect the framework to naturally extend to higher dimensional settings, where $\mathcal{V}_\infty(\rho)$ will consist of a *volume* in $\mathbb{R}^n$. However non-commuting observables complicate matters and it remains to be seen whether our geometric account extends to this setting. Furthermore, it is unclear whether non-passive states can be included in the same framework, in particular states with coherences between energy eigenspaces. Nonetheless, this geometric perspective offers new directions to attack important questions in the thermodynamics of quantum systems with multiple non-commuting observables [31].

This chapter has laid out a framework to study passivity, but many of its consequences are yet to be explored. For example, complete passivity was first studied in the infinite-dimensional setting [22]. Do problems arise if we extend our geometric perspective on passivity into this setting, considering a continuous convex region as opposed to the convex polygons of the finite-

dimensional regime? To illustrate some important consequences of the framework, the next chapter moves to the asymptotic regime, considering the connection between area and athermality in detail.

Chapter 3

A Geometric Measure of Athermality

The previous chapter introduced a geometric framework for analysing passivity, and presented a simple derivation of the Gibbs state. We associated to each state a convex shape in the 2-d plane, and showed that this shape must have zero area for a completely passive state. In this chapter:

- We connect the geometric framework from the previous chapter to thermodynamics in the asymptotic, many-copy regime. The notion of asymptotic equivalence between states is reframed as the asymptotic ϵ-s ensembles having the same 'centre-of-mass'.
- Building on the connection between thermality and colinearity of the ϵ-s ensemble, we use the area of the asymptotic ensemble to define a new geometric measure of athermality. In particular, we prove monotonicity of the geometric athermality under a class of physically motivated transformations.
- Finally, we connect with other measures of athermality in the literature, and discuss the pros and cons of our new geometric measure.

The main novel contributions of this chapter are the following:

- We make the connection between our geometric description and the E-S diagrams of [47], identifying the pair (E, S) as the centre-of-mass of the ϵ-s ensemble.

- We define a new measure of athermality – the geometric athermality – based on the area of the asymptotic ϵ-s ensemble and motivate its use as a macroscopic non-equilibrium thermodynamic variable.

- The monotonicity of the geometric athermality under appropriate transformations is proved in Section 3.4.

- Theorem 3.2 proves a connection between the β-athermality and the maximal entropic gain not previously noted in the literature. The geometric athermality is then related to other measures of athermality in Theorem 3.3, showing that both the asymptotic ergotropy and maximal entropic gain vary monotonically with the geometric athermality.

3.1 Macroscopic Non-equilibrium Thermodynamics

Thermodynamics assumes that a system in equilibrium is well described by a small number of macroscopic variables, such as volume, pressure, internal energy and entropy [48]. Can this be extended beyond the equilibrium setting?

In the previous chapter, we introduced a geometric perspective on passive states, providing a fine-grained description of the state, and encoding the full complement of energies and occupations. In the many-copy macroscopic limit, we introduced the asymptotic ϵ-s ensemble $\mathcal{V}_\infty(\rho)$, which provides fine-grained information about both equilibrium and non-equilibrium macroscopic states, in contrast to the usual coarse-grained thermodynamic description. Can we supplement the traditional macroscopic equilibrium variables with a macroscopic non-equilibrium variable that quantifies deviation from equilibrium? Does our geometric framework provide a natural extension of the usual macroscopic thermodynamic framework that captures some non-equilibrium features?

Several measures of athermality already exist in the literature, including the asymptotic ergotropy – the optimal rate of work extraction from many copies of the state [49]. However, this is defined by a process, whereas we should ideally like a measure that is intrinsic to the state.

In this chapter, we exploit the connection between athermality and the asymptotic ensemble's non-zero area to introduce a new measure – the geometric athermality. This measure depends only upon the state, and furthermore is not defined relative to a particular thermal state. Finally, we connect the geometric athermality with other measures in the literature.

3.2 The Macroscopic Regime and the Energy-Entropy Diagram

Throughout this chapter, we consider systems whose primary thermodynamic variables are internal energy E and entropy S. The macroscopic many-copy thermodynamics can be represented on an energy-entropy (E-S) diagram, where points in the 2-d plane correspond to pairings (E, S).

This graphical description is closely related to the framework introduced in the previous chapter. In fact, it as an averaged version of the ϵ-s ensemble. Specifically, we can think of $\mathcal{V}(\rho)$ as a two-component random variable that takes values (ϵ_i, s_i), with expectation value

$$\mathbb{E}\big[\mathcal{V}(\rho)\big] = \big(E(\rho), S(\rho)\big), \tag{3.1}$$

where $E(\rho) \coloneqq \mathrm{tr}[\rho H]$ is the average energy and $S(\rho) \coloneqq -\mathrm{tr}[\rho \log \rho]$ is the von Neumann entropy of the state ρ. A point on the E-S diagram is the 'centre-of-mass' description of an ϵ-s ensemble.

We illustrate key features of the E-S diagram in Figure 3.1. The region of allowed points is upper-bounded by the curve of thermal states, whose position is determined by the Hamiltonian of the system. States that are not completely passive lie strictly below this curve.

How are the two macroscopic variables of energy and entropy picked out as the sole variables in the macroscopic regime? One answer is asymptotic equivalence [47].

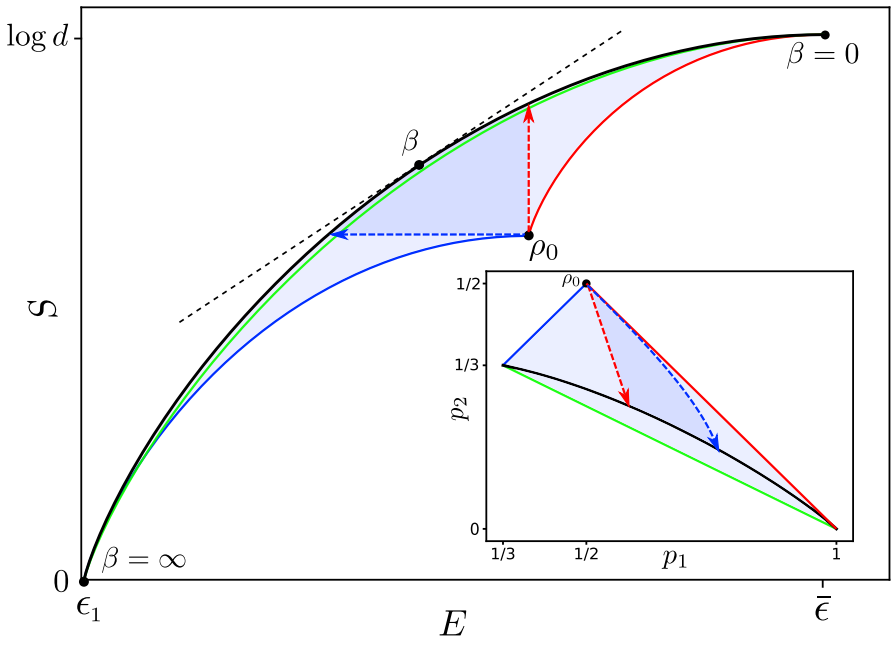


***Figure 3.1: Important features of the** E**-**S **diagram**. For concreteness, consider a qutrit system ($d = 3$) with $H = \sum_{i=1}^{3} \epsilon_i |e_i\rangle\langle e_i|$. A state of this system is represented by a point (E, S) lying on or below the curve of equilibrium states (black), which runs from $(\epsilon_1, 0)$ to $(\bar{\epsilon}, \log 3)$, where $\bar{\epsilon}$ is the average energy of the maximally mixed state $\frac{1}{3}\mathbb{1}_3$. The gradient of this curve is the inverse temperature β. The shaded region identifies the set of passive states. In the qutrit case, the E-S diagram maps continuously onto the qutrit parameter space in the inset graph (bottom right), with key features retained. This is special to a qutrit because of a bijective mapping between states and allowed pairs (E, S). The solid red, green and blue lines identify the boundaries of the set of passive qutrit states in the parameter space and the corresponding lines are highlighted in the main diagram. For the passive state $\rho_0 = \frac{1}{2}\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big)$ (black point), the isentropic (blue dashed line) and isoenergetic (red dashed line) trajectories are depicted, as well as all states attainable via activation trajectories (darker shaded region) – all discussed in later sections. Reproduced from [2].*

Definition 3.1: Asymptotic equivalence

For a d-dimensional system with Hamiltonian H, states ρ and σ are asymptotically equivalent, denoted $\rho \asymp \sigma$, if there exists an auxiliary system A of $\mathcal{O}(\sqrt{n \log n})$ qudits in a state η_A, with Hamiltonian H_A satisfying $\|H_A\| \leq \mathcal{O}\left(n^{\frac{2}{3}}\right)$, and an energy conserving unitary U, such that

$$\lim_{n\to\infty} \left\| \mathrm{tr}_A \left[U(\rho^{\otimes n} \otimes \eta_A) U^\dagger \right] - \sigma^{\otimes n} \right\| = 0 \ , \tag{3.2}$$

where $\|M\| := \mathrm{tr}\left(\sqrt{M^\dagger M}\right)$ denotes the trace norm.

States ρ and σ are asymptotically equivalent if they can be interconverted in the asymptotic many-copy limit under unitaries that globally conserve energy, while allowing the use of auxiliary systems that are sublinear in size. The sublinear size of the auxiliary systems ensures that the energy and entropy transferred to this system can be neglected in the asymptotic limit. Further technical details are found in [47], where the following condition for asymptotic equivalence is proved.

Lemma 3.1: Asymptotic Equivalence of ϵ-s ensembles

For a system with Hamiltonian H, $\rho \asymp \sigma$ iff $\mathbb{E}\big[\mathcal{V}(\rho)\big] = \mathbb{E}\big[\mathcal{V}(\sigma)\big]$.

To be asymptotically equivalent, ϵ-s ensembles must have the same statistical average or 'centre of mass'. This strengthens the connection between the geometric framework, introduced in the analysis of passivity, and the usual description of thermodynamics in the macroscopic regime.

Asymptotic equivalence picks out average energy and entropy as the relevant quantities for describing thermodynamics in the asymptotic many-copy limit. Therefore, we partition the state space into equivalence classes according to these variables.

Definition 3.2: (E, S) partitions

Given a system with Hamiltonian H, we define $\mathcal{D}_{(E,S)}$ as the set of states

$$\mathcal{D}_{(E,S)} \coloneqq \left\{\rho \ : \left(\mathrm{tr}[H\rho], -\mathrm{tr}[\rho \log \rho]\right) = (E, S)\right\} \tag{3.3}$$

with average energy E and entropy S.

In the E-S diagram, the area under the thermal curve is the union of all $\mathcal{D}_{(E,S)}$.

3.3 Geometric Athermality

Asymptotic equivalence does not bring with it a notion of thermal equilibrium. It picks out the centre of mass of the ϵ-s ensemble, however further information is required to describe the shape of the ensemble in the plane. Having noted that a non-zero area for the asymptotic ensemble $\mathcal{V}_\infty(\rho)$ signals athermality, we ask whether this area has a physical interpretation, and whether it can be used as an additional non-equilibrium variable.

To assess this, we require an explicit expression for the area of the asymptotic ensemble, most easily expressed in terms of the vertices of $\mathcal{V}_\infty(\rho)$.

Definition 3.3: Vertices of $\mathcal{V}_\infty(\rho)$

For a d-dimensional system in a passive state ρ, the set of $n \leq d$ vertices of the convex set $\mathcal{V}_\infty(\rho)$ is

$$\mathcal{V}_{\mathrm{vert}}(\rho) \coloneqq \mathrm{ext}\left[\mathcal{V}_\infty(\rho)\right] = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}, \tag{3.4}$$

where $\mathrm{ext}[\mathcal{C}]$ denotes the set of extremal points of the convex set $\mathcal{C}$. By convention, the vertices $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are labelled clockwise modulo n around the convex hull $\mathcal{V}_\infty(\rho)$.

Lemma 3.2: Area of $\mathcal{V}_\infty(\rho)$

The area of $\mathcal{V}_\infty(\rho)$ is

$$A(\rho) = \frac{1}{2} \sum_{k \bmod n} s_k \ \Delta_{k-1,k+1} \ , \tag{3.5}$$

where the sum runs over the n elements of $\mathcal{V}_{\text{vert}}(\rho)$, and $\Delta_{i,j} \coloneqq \epsilon_i - \epsilon_j$ is the energy level spacing between the corresponding elements.

This expression can be derived from the trapezium rule [50]. For a qutrit ($d = 3$), each state uniquely specifies (E, S), meaning we can straightforwardly supplement them with the area $A(\rho)$. However, for systems with $d > 3$, the situation is more complex because multiple states can have the same (E, S), and $A(\rho)$ can vary over the different states in $\mathcal{D}_{(E,S)}$. We exploit asymptotic equivalence to single out the minimal area for a given (E, S) pairing. This is the geometric athermality.

Definition 3.4: Geometric athermality

Given a d-dimensional system with Hamiltonian H, the geometric athermality $\mathcal{A}$ for (E, S) is

$$\mathcal{A}(E, S) \coloneqq \inf_{\rho \in \mathcal{D}_{(E,S)}} A(\rho) \ . \tag{3.6}$$

States are described by the triple $(E, S, \mathcal{A})$. The geometric athermality encodes whether the state is in thermal equilibrium, without providing the fine-grained description of the ϵ-s ensemble.

3.4 Geometric Athermality as a Measure of Athermality

In the previous chapter, we noted that the area of $\mathcal{V}_\infty(\rho)$ *witnessed* athermality, being zero iff ρ is completely passive. Derived from this area, can the geometric athermality be ascribed a physical interpretation? Is $\mathcal{A}$ a valid measure of athermality? Specifically, is $\mathcal{A}$ monotonically non-increasing under a set of physically motivated trajectories in the E-S plane that move the

state towards the Gibbs states?

Consequently, we must specify the allowed transformations on passive states, operations that lead to work extraction without the addition of ordered energy to the system. To prevent work resources being implicitly smuggled in, we restrict to transformations that neither increase the average energy nor decrease the entropy of the system. This is encapsulated in the concept of an activation trajectory.

Definition 3.5: Activation trajectory

An activation trajectory on the E-S diagram has tangent unit vector $\boldsymbol{u} \coloneqq (u_E, u_S)$ satisfying $u_E \leq 0$ and $u_S \geq 0$ at all points along the trajectory directed towards the manifold of equilibrium states.

In Figure 3.1, these maps take the state ρ_0 only within the shaded region, moving within its upper-left quadrant on the E-S diagram. We can now justify the claim that $\mathcal{A}$ is a true measure of athermality.

Theorem 3.1: Monotonicity of $\mathcal{A}$

For a d-dimensional system with non-degenerate Hamiltonian H, the geometric athermality $\mathcal{A}$ is monotonically non-increasing along all activation trajectories.

The complete proof is presented in Appendix A, involving several technicalities, however it essentially generalises the proof for the qutrit case. The qutrit case is special because there is a bijective mapping between states and pairs (E, S), therefore minimisation over $\mathcal{D}_{(E,S)}$ is not required to obtain $\mathcal{A}$.

Proof. (For $d = 3$ systems only) The differential of the geometric athermality $\mathcal{A}$ for any passive qutrit state ρ is

$$\mathrm{d}\mathcal{A} = \frac{1}{4\mathcal{A}} \sum_{k \bmod 3} \frac{\Delta^2_{k+2,k+1}}{p_k} \left(\beta_{i,j}\mathrm{d}E - \mathrm{d}S\right), \tag{3.7}$$

derived in Appendix A. Since $\partial_E \mathcal{A} > 0$ and $\partial_S \mathcal{A} < 0$, the geometric athermality is monotonically

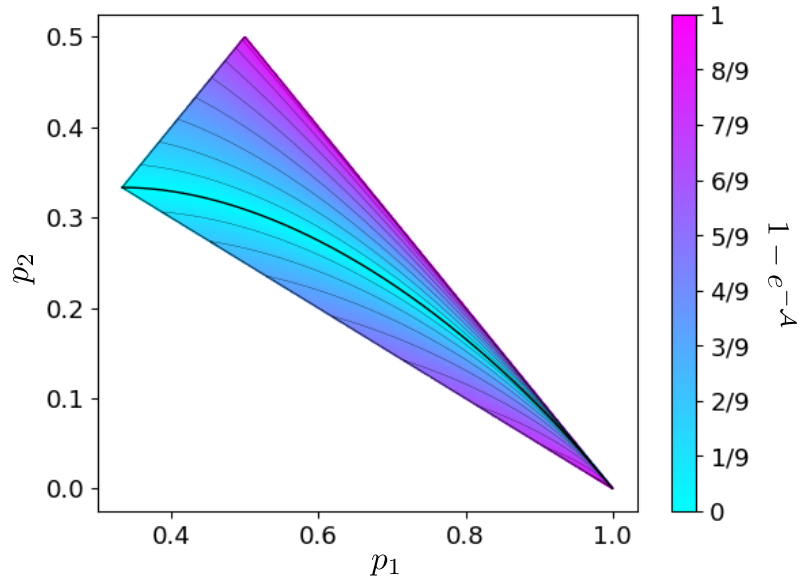


Figure 3.2: Contour plot of $\mathcal{A}$ for passive states of a qutrit. *The system has energy spectrum $(0, 1, 2)$. The thick black contour in the lighter region corresponds to the set of thermal equilibrium states. We rescale the value of $\mathcal{A}$ to $1 - e^{-\mathcal{A}}$ for clarity to highlight the structure of the contours close to the equilibrium curve. Reproduced from [2].*

non-increasing along any activation trajectory, because the directional derivative is

$$\nabla_{\boldsymbol{u}}\mathcal{A} = u_E\ \partial_E\mathcal{A} + u_S\ \partial_S\mathcal{A} < 0, \tag{3.8}$$

as required. □

The monotonicity of $\mathcal{A}$ is seen in Figure 3.2 for qutrits, attaining a minimum on the curve of thermal states.

3.5 Connecting to other measures of athermality

Geometric athermality is not the only measure of athermality for quantum states. In this section, we will explore how $\mathcal{A}$ relates to some of these other measures, highlighting their simple geometric interpretations in our framework.

3.5.1 Asymptotic Ergotropy $W_{\max}$

Given that Gibbs states are completely passive, deviation from equilibrium can be quantified by the rate of work extraction from many copies of the state [49, 51].

Definition 3.6: Asymptotic ergotropy

The asymptotic ergotropy of a state ρ is

$$W_{\max} \coloneqq E(\rho) - E(\gamma_{\beta_{\max}}), \tag{3.9}$$

where $\beta_{\max}$ is chosen such that $S(\gamma_{\beta_{\max}}) = S(\rho)$.

$W_{\max}$ is only 0 for completely passive states. The asymptotic work extraction process follows an isentropic trajectory, moving (E, S) horizontally in the E-S diagram. In fact, $W_{\max}$ is simply the horizontal distance to the thermal curve. This measure is physically motivated, but it is defined according to a process rather than intrinsic to the state.

3.5.2 β-Athermality a_β

β–athermality quantifies the distance to a Gibbs state using a relative entropy measure [47].

Definition 3.7: β-athermality

The β-athermality relative to a particular Gibbs state γ_β is

$$a_\beta(\rho) \coloneqq S_{\text{rel}}(\rho||\gamma_\beta) = \beta(F(\rho) - F(\gamma_\beta)), \tag{3.10}$$

where $F(\sigma) \coloneqq E(\sigma) - \beta^{-1}S(\sigma)$ is the free energy of σ.

Like the geometric athermality and asymptotic ergotropy, a_β can only attain the value 0 on the thermal curve. However, $a_\beta = 0$ is only attained at a single point on the thermal curve, since a Gibbs state $\gamma_{\beta'}$ with $\beta' \neq \beta$ has non-zero β–athermality. It explicitly depends on a particular Gibbs state, unlike the geometric athermality, which does not have this temperature dependence. In certain circumstances, this β-dependence is desirable, for example in the usual

resource-theoretic approach to thermodynamics, where systems can freely equilibrate with some background thermal reservoir at inverse temperature β [52]. On the other hand, geometric athermality quantifies deviation from equilibrium without reference to a particular temperature, being zero for *all* Gibbs states.

3.5.3 Asymptotic Maximal Entropic Gain $\Delta S_{\max}$

A further measure can be derived analogously to the asymptotic ergotropy, instead considering the maximal entropic gain along isoenergetic (energy-conserving) trajectories in the E-S diagram. In a sense, this is dual to the asymptotic ergotropy.

Definition 3.8: Asymptotic maximal entropic gain

The asymptotic maximal entropic gain for a state ρ is

$$\Delta S_{\max} \coloneqq S(\gamma_{\beta_{\min}}) - S(\rho), \tag{3.11}$$

where $\beta_{\min}$ is chosen such that $E(\gamma_{\beta_{\min}}) = E(\rho)$.

A disadvantage of the β-athermality was its freedom in the choice of β. The maximal entropic gain naturally selects β, picking out the minimal β-athermality.

Theorem 3.2: Minimal β-athermality

The maximal entropic gain for a system in a state ρ along all possible activation trajectories is given by its minimal β–athermality,

$$\Delta S_{\max} = a_{\beta_{\min}}(\rho) = \inf_{\beta \in [0,\infty)} a_\beta(\rho). \tag{3.12}$$

Proof. By definition of β-athermality,

$$a_\beta(\rho) = \beta\big(F(\rho) - F(\gamma_\beta)\big) = \beta\big(E(\rho) - E(\gamma_\beta)\big) + \big(S(\gamma_\beta) - S(\rho)\big). \tag{3.13}$$

This has an extremal point when

$$\frac{\mathrm{d}a_\beta}{\mathrm{d}\beta} = E(\rho) - E(\gamma_\beta) - \beta\frac{\mathrm{d}E(\gamma_\beta)}{\mathrm{d}\beta} + \frac{\mathrm{d}S(\gamma_\beta)}{\mathrm{d}\beta} = E(\rho) - E(\gamma_\beta) = 0, \tag{3.14}$$

which is a minimum because

$$\frac{\mathrm{d}^2a_\beta}{\mathrm{d}\beta^2} = -\frac{\mathrm{d}E(\gamma_\beta)}{\mathrm{d}\beta} > 0, \tag{3.15}$$

seen by considering the gradient β of the thermal curve on the E-S diagram. The minimum β-athermality is attained when $E(\rho) = E(\gamma_{\beta_{\min}})$, which implies that

$$\min_{\beta\in[0,\infty)} a_\beta(\rho) = a_{\beta_{\min}}(\rho) = \beta_{\min}\big(E(\rho) - E(\gamma_{\beta_{\min}})\big) + \big(S(\gamma_{\beta_{\min}}) - S(\rho)\big) \tag{3.16}$$

$$= S(\gamma_{\beta_{\min}}) - S(\rho) = \Delta S_{\max}. \tag{3.17}$$

□

The minimal β-athermality is the vertical distance on the E-S diagram from the point corresponding to ρ to the thermal curve. Physically, it is the entropy gain of the state ρ as it thermalises at temperature $\beta_{\min}$ with zero work output. This selects a particular β for the β-athermality, and ensures that all thermal states are assigned a value of zero.

3.5.4 Connecting $W_{\max}$ and $\Delta S_{\max}$ to $\mathcal{A}$

The simple geometric interpretations of $W_{\max}$ and $\Delta S_{\max}$ suggest that there should be some connection with the geometric athermality. This is confirmed in numerical calculations for the qutrit case, where we calculate how $\mathcal{A}$ (simply $A(\rho)$ in this case) varies with respect to the asymptotic ergotropy $W_{\max}$, as pictured in Figure 3.3.

Theorem 3.3: $W_{\max}$ and $\Delta S_{\max}$ vary monotonically with $\mathcal{A}$

$W_{\max}$ and $\Delta S_{\max}$ are both monotonically non-decreasing functions of $\mathcal{A}$ along any differentiable activation trajectory.

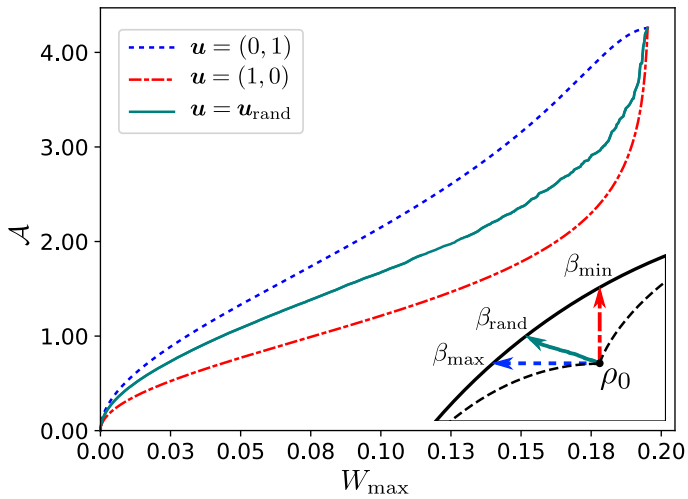


Figure 3.3: $\mathcal{A}$ ***vs.*** $W_{\max}$ ***for a passive qutrit state along three activation trajectories.*** *The energy spectrum is* $(0, 1, 2)$*. The initial state is the maximally energetic passive state* $\rho_0 = \frac{1}{2}(|e_0\rangle\langle e_0| + |e_1\rangle\langle e_1|) + \frac{\delta}{3}\mathbb{1}$ *[26], where* $\delta \sim 10^{-4}$ *to keep the initial* $\mathcal{A}$ *finite. The trajectories are labelled by their tangent vector* $\boldsymbol{u} \coloneqq (u_E, u_S)$ *with* $\boldsymbol{u}_{\text{rand}}$ *corresponding to a random choice of* $\boldsymbol{u}$*. Reproduced from [2].*

Proof. The differentials of $W_{\max}$ and $\Delta S_{\max}$ are

$$\mathrm{d}W_{\max} = \mathrm{d}E - \beta_{\max}^{-1}\mathrm{d}S \quad \text{and} \quad \mathrm{d}\Delta S_{\max} = \beta_{\min}\mathrm{d}E - \mathrm{d}S. \tag{3.18}$$

For a particular trajectory $S(E)$, we have

$$\mathrm{d}W_{\max} = \left(1 - \beta_{\max}^{-1}\frac{\mathrm{d}S}{\mathrm{d}E}\right)\mathrm{d}E \quad \text{and} \quad \mathrm{d}\Delta S_{\max} = \left(\beta_{\min} - \frac{\mathrm{d}S}{\mathrm{d}E}\right)\mathrm{d}E. \tag{3.19}$$

An activation trajectory has $\mathrm{d}S/\mathrm{d}E \leq 0$ at all points, therefore

$$\frac{\mathrm{d}W_{\max}}{\mathrm{d}E} > 0 \quad \text{and} \quad \frac{\mathrm{d}\Delta S_{\max}}{\mathrm{d}E} \geq 0. \tag{3.20}$$

Furthermore, $\partial_E\mathcal{A} \geq 0$ and $\partial_S\mathcal{A} \leq 0$ for activation trajectories,

$$\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}W_{\max}} = \left(\partial_E\mathcal{A} + \partial_S\mathcal{A}\,\frac{\mathrm{d}S}{\mathrm{d}E}\right)\frac{\mathrm{d}E}{\mathrm{d}W_{\max}} \geq 0 \tag{3.21}$$

and similarly $\mathrm{d}\mathcal{A}/\mathrm{d}\Delta S_{\max} \geq 0$. □

3.5.5 Comparison of athermality measures

Although $W_{\max}$ and $\Delta S_{\max}$ have natural operational interpretations, their definitions are based on processes rather than the state itself, and particular temperatures such as $\beta_{\min}$ or $\beta_{\max}$ are introduced. On the other hand, the geometric athermality is defined solely with reference to the state of the system, and treats all points on the thermal curve equally.

The monotonicity of the geometric athermality is only guaranteed over the passive states on the E-S diagram, while monotonicity of $W_{\max}$ and $\Delta S_{\max}$ holds for all pairs (E, S). However, equivalence classes extend this framework to all quantum states, taking the geometric athermality of the representative passive state as the value for all states in that unitary orbit.

Although the definition of $\mathcal{A}$ is simple and intuitive, it is not necessarily easy to compute, requiring minimisation of $A(\rho)$ over $\mathcal{D}_{(E,S)}$. Further work is required to determine if the minimal area ensemble has a natural shape.

3.6 Conclusions and Outlook

In this chapter, we built on the geometric perspective of passivity, using the area of the asymptotic ensemble to define a valid measure of athermality, and related it to other measures in the literature. We view the asymptotic ϵ-s ensemble as an extension of the coarse-grained E-S description of thermodynamics in the asymptotic regime, with the single point (E, S) replaced by a convex area. The minimal area for a given (E, S) indicates how close the state could be to equilibrium, motivating the geometric athermality as a measure of athermality.

Despite the simplicity of its definition, the geometric athermality is difficult to compute exactly, although approximate values can easily be obtained. For a d–dimensional system with Hamiltonian H, it remains an open question what shape of $\mathcal{V}_\infty(\rho)$ has minimal area for given values of (E, S), and whether one can find an explicit form of the quantum state ρ such that $A(\rho) = \mathcal{A}\big(E(\rho), S(\rho)\big)$. Perhaps this question can be answered more easily in an infinite-

dimensional analogue of our framework, where calculus of variations may indicate the minimal area shape in this setting. Alternatively, concepts from information geometry may provide a metric structure that more naturally expresses the geometric athermality and how it varies over the E-S plane.

Consideration of the geometric athermality is justified by its monotonicity along activation trajectories, being non-increasing as the state moves towards the curve of equilibrium states in the E-S plane. Unlike other common sets of operations, such as thermal operations, these activation maps do not make reference to a particular temperature, whilst also ensuring that resources are not smuggled in. Geometric athermality is therefore a valid resource measure under these physically motivated constraints.

Although the proof of the monotonicity of $\mathcal{A}$ given in Appendix A is somewhat technical, it is worth noting the special role played by boundary points of the asymptotic ϵ-s ensemble. These are also key in the derivation of the Gibbs state in the previous chapter, in contrast to the derivation using virtual temperatures in [24], which does not make this distinction between boundary and internal virtual temperatures. The significance of boundary points suggest that transformations might be described purely in terms of boundary virtual temperatures. This could simplify and enhance the use of virtual temperatures in the analysis of thermodynamic transformations, opening an interesting avenue for further study.

Together with average energy and entropy, the geometric athermality acts like an additional non-equilibrium thermodynamic variable, intermediate between the coarse-grained (E, S) description and the full fine-grained microscopic description of a state in terms of its energy levels and populations. We do not claim that the geometric athermality is a unique description, only that it naturally follows from the geometric description of the previous chapter, and it would be interesting to explore further coarse-grained thermodynamic descriptions in the asymptotic regime. Moreover, having multiple levels of description for resources is a theme we return to in the next part of the thesis.

Part II

Resources from Symmetry Constraints

Chapter 4

Classifying Symmetries of Quantum States & Channels

Like many fields of physics, symmetry is at the heart of quantum information theory. In recent years, the consideration of symmetry principles has borne fruit in areas such as thermodynamics, metrology and quantum computation. However, structures arising from symmetry constraints are often highly non-trivial and difficult to analyse.

In this part of the thesis, we introduce a framework for classifying quantum states and channels in terms of their symmetries, allowing us to study the structure and manipulation of resources in the context of symmetry constraints.

1. After motivating the importance of symmetry in quantum information theory, we review the mathematical formalism of symmetry – group theory and representation theory.

2. We introduce a classification of quantum states and channels based upon their residual symmetries. This is encoded in the 'shape' of the corresponding group orbit.

3. Finally, we discuss the robustness of the framework to small perturbations, confirming that our framework is physically meaningful.

Much of this chapter is a brief introduction of the necessary mathematics, however it still

contains some novel contributions:

- We propose the 'shape' of the group orbit as a way to classify quantum states and channels by their residual symmetries. Although group orbits are well-studied in the literature, their use as a coarse-grained classification in a resource-theoretic context is new.

- In Section 4.5 we discuss the extent to which a distance measure can be defined between these 'shapes', proving that the minimum distance between sets of quantum states with given isotropy subgroups is zero. Nonetheless, we argue that a smoothed version of the isotropy subgroup can still be defined in a meaningful and useful way.

The next three chapters are all based on work presented in [1], carried out in collaboration with David Jennings.

4.1 Motivation – Symmetry in Quantum Information Theory

Modern physics could be glibly characterised as the study of symmetry principles in both the classical and quantum realms. Particle physics is a glorious vindication of this programme, from Noether's theorem – linking symmetries to conservation laws [53] – to gauge theories and the formulation of the standard model [54]. In materials science, crystallography highlights how bulk properties of materials are largely determined by the symmetry group of the crystal structure [55]. Fundamentally, symmetry constrains dynamics and simplifies computation. It is therefore unsurprising that symmetry considerations have driven research throughout quantum information theory.

- Symmetry groups provide an invaluable toolkit in the context of quantum computing, from practical applications of stabilizer codes in quantum error correction and benchmarking quantum computers [3, 56–58], to understanding the computational power of quantum computers [59–61].

- Quantum speed limits provide bounds on how quickly an operation can be performed, resulting in rigorous versions of the time-energy uncertainty relation [62].

- The WAY theorem implies that conservation laws limit the precision of measurements, however quantum states that break the symmetry allow us to circumvent these limitations [63–68]. Moreover, limitations in metrology can be interpreted in terms of symmetries [69]. For example, phase measurements rely upon broken symmetries to distinguish different transformations. Fundamentally, a clock uses asymmetry under time translation.

- Moving beyond Noether's theorem, one line of research has addressed the regime in which symmetry principles disconnect from conservation laws [70, 71], with applications in quantum thermodynamics [72–74].

These applications have benefited from the resource-centric philosophy adopted in this thesis, focusing on symmetry constraints and analysing the structures that arise. This structure is often prohibitively complex, therefore we can profit from coarse-grained information about quantum states and operations under such constraints.

This chapter describes one such tool based on the group orbit, a mathematical object encoding the symmetries of quantum states and channels. Its 'shape' provides a means of classification. The following chapters explore the structures that arise from this classification and to what extent this tool can be exploited.

4.2 A Mathematical Description of Symmetry

The study of symmetry is a vast and elegant subject. See [54, 55, 75] for introductions to this topic from a physics perspective. We are concerned with only a few basic aspects, which we now introduce.

4.2.1 Group Theory – the Structure of Symmetry

The abstract language of symmetry is group theory; its basic objects are groups. The definition of groups generalises the properties of bijective transformations, otherwise known as permutations in the finite case.

Definition 4.1: Groups

A group is a set G of objects that combine according to the following axioms:

1. *(Closure)* If $g, h \in G$, then $gh \in G$.
2. *(Associativity)* For all $f, g, h \in G$, we have that $f(gh) = (fg)h$.
3. *(Identity)* There exists a unique element $e \in G$ such that $eg = ge = g$ for all $g \in G$.
4. *(Inverse)* For every $g \in G$, there exists an element $g^{-1} \in G$ such that $gg^{-1} = g^{-1}g = e$.

Throughout these chapters, we denote the composition of elements g and h as gh. Crucially, operations that leave an object unchanged together form a group. Here are some examples of groups that feature in the discussion that follows.

Example 4.1: Some useful groups

- $\mathbb{Z}_n$ – integers modulo n under addition (or n-fold rotations about a single axis) .
- SO(n) – $n \times n$ orthogonal matrices with determinant 1, composed by matrix multiplication.
- SU(n) – $n \times n$ unitary matrices with determinant 1, composed by matrix multiplication.

SO(n) and SU(n) are associated with rotations, while O(n) and U(n) drop the requirement that the determinant of the matrices is +1, incorporating reflections and inversions. Groups can contain either finite or infinitely many elements, and we shall encounter both in the course

of these chapters.

Although the above groups are given in terms of particular examples, we often consider them as abstract mathematical structures – sets of elements that compose in a particular way. If two groups have the same abstract structure, they are isomorphic.

Definition 4.2: Group Isomorphism

Groups G and H are isomorphic, denoted $G \cong H$, if there is a bijective mapping $f : G \to H$ such that, for all $g_1, g_2 \in G$,

$$f(g_1 g_2) = f(g_1)\ f(g_2)\ , \tag{4.1}$$

where the composition rule for H holds on the right-hand side.

Example 4.2: SO(2) ≅ U(1)

There is an invertible mapping between SO(2) and U(1), given by

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}(2) \quad \to \quad e^{i\theta} \in \mathrm{U}(1)\ , \tag{4.2}$$

which preserves the group structure. Rotations in the two-dimensional plane share the structure of unit complex numbers under multiplication.

Groups can also be contained within other groups.

Definition 4.3: Subgroups

The group H is a subgroup of another group G if $H \subseteq G$ (subset inclusion) with the same composition law.

Example 4.3: SO(2) $\subseteq$ SO(3)

It is clear that SO(2) is a subgroup of SO(3), since we can embed the SO(2) matrices as

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{SO}(3) \tag{4.3}$$

while preserving the group composition structure. Physically, rotations about a single axis form a subgroup of rotations in three dimensions.

Subgroups form the foundation for our classification scheme, since we consider subgroups of transformations that leave quantum states and channels unchanged.

4.2.2 Representations and Group Actions

While group theory sees abstract mathematical entities, we more often think of group elements as transformations on some physical object. In quantum mechanics, our description of a system and its dynamics plays out in the world of vector spaces, therefore we consider how group transformations act on such spaces.

Definition 4.4: Representation of a group

A representation of a group G on a vector space V is a mapping $U : G \to GL(V)$, such that for all $g, h \in G$,

$$U(gh) = U(g)\, U(h), \tag{4.4}$$

where $GL(V)$ denotes the space of linear operators on V.

This mapping need not be invertible, only needing to respect the group composition law. Therefore, a representation does not have to capture the full group structure.

Example 4.4: SO(2) Representations

The sets of matrices given in Example 4.2 are both representations for SO(2). This two-dimensional defining representation is embedded alongside a trivial one-dimensional representation (where all group elements are mapped to $\mathbb{I}$) to form the three-dimensional representation in Example 4.3.

We consider only unitary representations, such that $U(g^{-1}) = U^\dagger(g)$. This can be done without loss of generality for finite and compact groups because all representations are equivalent (up to some similarity transformation) to a unitary representation [54].

We use a particular representation (or group action) to describe how a symmetry group acts on quantum systems.

Definition 4.5: Group actions for quantum systems

For any quantum system, we denote its associated Hilbert space by $\mathcal{H}$, and the set of linear operators on $\mathcal{H}$ as $\mathcal{B}(\mathcal{H})$. The group element $g \in G$ acts upon a quantum system in the following ways:

- The group action on a quantum state $\rho \in \mathcal{B}(\mathcal{H})$ is given by the adjoint action $\mathcal{U}_g(\rho) := U(g)\ \rho\ U^\dagger(g)$.
- The group action on a quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is $\mathfrak{U}_g(\mathcal{E}) := \mathcal{U}_g \circ \mathcal{E} \circ \mathcal{U}_{g^{-1}}$, where $\circ$ denotes the concatenation of channels.

Note that the group action on states is a special case of the action on channels, because state preparation is a particular class of channel.

4.3 Classifying the Symmetries of Quantum States & Channels

Group transformations either change a quantum channel (or state) or leave it unaltered. Some states and channels are invariant under all transformations in the symmetry group.

Definition 4.6: Symmetric states and operations

A symmetric state satisfies $\mathcal{U}_g(\rho) = \rho$ for all $g \in G$. Likewise, a symmetric operation is a quantum channel such that $\mathfrak{U}_g(\mathcal{E}) = \mathcal{E}$ for all $g \in G$.

Example 4.5: Some simple symmetric states and channels

- The maximally mixed state $\mathbb{I}/d$ is a symmetric state for all groups that act on a finite-dimensional system.
- A 'classical' qubit state $\rho = (1-p)\left|0\right\rangle\left\langle 0\right| + p\left|1\right\rangle\left\langle 1\right|$ is symmetric under the U(1) group generated by Z, i.e. $\mathrm{U}(1) = \{e^{-i\theta Z} : \theta \in [0, 2\pi)\}$.
- A channel that discards the system and prepares a symmetric state is a symmetric operation.

These states and operations are special; they can be prepared or implemented under symmetry constraints. However, most states and channels are invariant under a proper subset of the group transformations. These residual symmetries classify quantum states and channels, and they belong to the isotropy subgroup (or stabilizer).

Definition 4.7: Isotropy subgroup

The isotropy subgroup of a quantum channel $\mathcal{E}$ under the action of a group G is

$$\mathrm{Iso}(\mathcal{E}) := \{g \in G \ : \ \mathfrak{U}_g(\mathcal{E}) = \mathcal{E}\}. \tag{4.5}$$

This notion applies to quantum states, where $\mathrm{Iso}(\rho) := \{g \in G \ : \ \mathcal{U}_g(\rho) = \rho\}$.

Symmetric channels are invariant under all group transformations, and have $\mathrm{Iso}(\mathcal{E}) = G$. Furthermore, the isotropy subgroup will always be a subgroup of G. Quantum states and channels can be partitioned according to their residual symmetries under the group action.

Definition 4.8: Partition by isotropy subgroups

Under the action of a group G, there is an equivalence relation $\sim$ for quantum channels, such that

$$\mathcal{E}_1 \sim \mathcal{E}_2 \iff \mathrm{Iso}(\mathcal{E}_1) = \mathrm{Iso}(\mathcal{E}_2). \tag{4.6}$$

This partitions the space of quantum channels into sets $C(H) := \{\mathcal{E} : \mathrm{Iso}(\mathcal{E}) = H\}$. The same construction applies for states.

The sets $C(H)$ are not generally convex. A simple counter-example exists for a single qubit under an SU(2) group action, where $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ belong to $C\big(\mathrm{U}(1)\big)$, but the convex mixture $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\mathbb{I} \in C\big(\mathrm{SU}(2)\big)$. However, $C(G)$ is convex, because there is no larger isotropy subgroup than G itself. These are instances of the following general result.

Lemma 4.1: Symmetries of convex combinations

Given two quantum channels $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ and $\mathcal{F} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ under the group action $\mathfrak{U}_g$, then $\mathrm{Iso}(p\,\mathcal{E} + (1-p)\,\mathcal{F}) \supseteq \mathrm{Iso}(\mathcal{E}) \cap \mathrm{Iso}(\mathcal{F})$, where $0 \leq p \leq 1$.

Proof. If $g \in \mathrm{Iso}(\mathcal{E}) \cap \mathrm{Iso}(\mathcal{F})$, then $\mathfrak{U}_g(p\,\mathcal{E} + (1-p)\,\mathcal{F}) = p\,\mathfrak{U}_g(\mathcal{E}) + (1-p)\,\mathfrak{U}_g(\mathcal{F}) = p\,\mathcal{E} + (1-p)\,\mathcal{F}$ by linearity, implying $\mathrm{Iso}(\mathcal{E}) \cap \mathrm{Iso}(\mathcal{F}) \subseteq \mathrm{Iso}(p\,\mathcal{E} + (1-p)\,\mathcal{F})$. □

The utility of the isotropy subgroup is clear from the widespread use of stabilizers in quantum error correction and quantum computing [3], where the focus is on the isotropy subgroup under a Pauli group action. We are applying the concept to general groups. We shall later employ the structure of these subgroups to analyse resources in the context of symmetry constraints.

4.4 Group Orbits and Homogeneous Spaces

The isotropy subgroup is intimately connected with the group orbit – the set that can be reached by applying all group transformations.

Definition 4.9: Group orbit

The group orbit of a quantum channel is $\mathcal{M}(\mathcal{E}) := \{\mathfrak{U}_g(\mathcal{E}) : \ g \in G\}$. Similarly, the group orbit of a quantum state is $\mathcal{M}(\rho) := \{\mathcal{U}_g(\rho) : \ g \in G\}$.

In the thermodynamics section of this thesis, we saw how orbits were used to partition states and analyse resources for work extraction. Once again, the group orbits partition the space of states or channels, but they also implicitly encode which group transformations leave a state or channel invariant.

The structure of the group orbit implicitly depends on how the group acts on the state or channel. Group transformations that belong to the isotropy subgroup effectively do not contribute to the group orbit. Instead of the full group G, the structure of the group orbit is described by the cosets of the isotropy subgroup.

Definition 4.10: Cosets

The coset gH of a subgroup $H \subseteq G$ is $gH := \{gh \ : \ h \in H\}$. The cosets of H partition the group G, and we denote the set of these cosets as G/H.

Cosets collect together the group transformations that have the same result on the state or channel in question. For any $h \in \text{Iso}(\rho)$, $\mathcal{U}_g(\rho) = \mathcal{U}_{gh}(\rho)$ because elements of $\text{Iso}(\rho)$ leave the state unchanged. All elements of the coset $g\text{Iso}(\rho)$ have the same effect on ρ and contribute the same element to $\mathcal{M}(\rho)$, therefore the group orbit can be fully described by the set G/H of cosets instead of the full group action of G. This eliminates from the description the redundancy that was caused by residual symmetries. In many physically important scenarios, we focus on continuous Lie groups[1] such as U(1) or SU(2). For these cases, the group orbit is a special kind of continuous manifold called a homogeneous space, specified by just the group G and the isotropy subgroup $\text{Iso}(\mathcal{E})$ (up to conjugacy class) [76–78].

[1] Lie groups are continuous groups whose elements form a smooth differentiable manifold [75].

Theorem 4.1: Group orbits are homogeneous spaces

The group orbit $\mathcal{M}(\mathcal{E})$ of a quantum channel $\mathcal{E}$ coincides with the quotient space $G/\mathrm{Iso}(\mathcal{E})$ up to diffeomorphisms, which we write $\mathcal{M}(\mathcal{E}) \cong G/\mathrm{Iso}(\mathcal{E})$ (and likewise for quantum states).

One could intuitively view $G/\mathrm{Iso}(\mathcal{E})$ as the group manifold of G with the action of $\mathrm{Iso}(\mathcal{E})$ eliminated, encoding that it has no effect on $\mathcal{E}$. We refer to $G/\mathrm{Iso}(\mathcal{E})$ as the 'shape' of the group orbit, although in the mathematical literature it is called the orbit-type [78]. Partitioning by shape collects together quantum states or channels with isomorphic isotropy subgroups, which amounts to taking the union of the sets $C(H)$ with isomorphic subgroups H.

Example 4.6: One-qubit shapes

The possible shapes for one-qubit states under an SU(2) group action are

$$\mathcal{M}\left(\frac{1}{2}[\mathbb{I} + \boldsymbol{r}\cdot\sigma]\right) = \begin{cases} SU(2)/U(1) \cong S^2 & \boldsymbol{r} \neq 0 \\ SU(2)/SU(2) \cong \{e\} & \boldsymbol{r} = 0\ , \end{cases} \tag{4.7}$$

as illustrated in Figure 4.1.

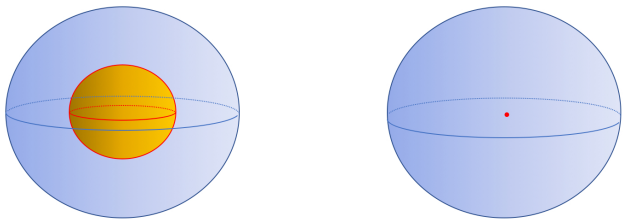


Figure 4.1: Orbits of one-qubit states under $SU(2)$***.*** *The only possible shapes in this case are shown on the Bloch sphere: (left) a sphere S^2 for non-zero Bloch vectors, or (right) a point $\{e\}$ for the symmetric maximally mixed state. Reproduced from [1].*

The same framework incorporates quantum states and channels, since the state ρ can be viewed as the channel $1 \to \rho$. In the next chapter, we consider SU(2) acting on two-qubit states. Even for that simple case, the classification is complex, nonetheless this coarse-grained

description of symmetry properties offers a handle on the full structure imposed by symmetry constraints. We explore this further in the final chapter of this part of the thesis.

4.5 The Robustness of Shapes to Small Perturbations

In reality, any operations will involve some level of noise, so we should like a framework that is robust to small perturbations, such that tiny errors do not cause it to fail or become trivial. As presented thus far, our framework appears to fail in this regard, as the following intuitive example illustrates.

Example 4.7: The "speck of dust" argument

Arbitrarily small perturbations to any quantum channel $\mathcal{E}$ can completely break the symmetry, leaving no residual symmetries in $\mathrm{Iso}(\mathcal{E})$. As illustrated in Fig. 4.2, a rotationally symmetric ball can be made completely asymmetric by sticking arbitrarily small pieces of dust onto it.

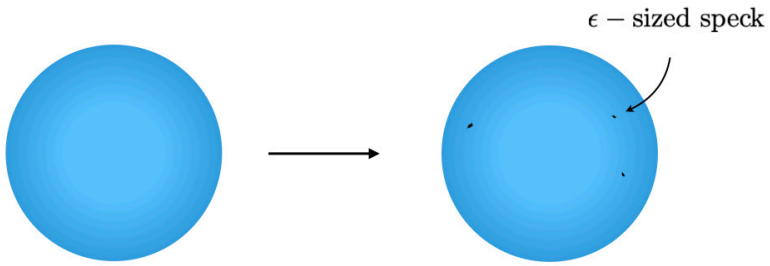


Figure 4.2: The speck of dust argument. *From a rotationally symmetric object (left) we can add arbitrarily small perturbations that completely break the rotational symmetry (right). Reproduced from [1].*

Problematically, the perturbed and unperturbed states (or channels) are practically indistinguishable, however their isotropy subgroups (and therefore their symmetry properties) can be very different. A similar argument still applies when the speck does not break the symmetry

completely. As a result, one cannot use a metric on the set of quantum states to talk of the 'distance' between the sets $C(H)$ – such a distance will be zero.

Lemma 4.2: The distance between $C(H)$ sets is zero

Let $d(\cdot,\cdot)$ be any metric on the space of quantum states, from which we define

$$\mathbf{d}(C(H_1), C(H_2)) := \inf_{\substack{\sigma_1 \in C(H_1) \\ \sigma_2 \in C(H_2)}} d(\sigma_1, \sigma_2). \tag{4.8}$$

Then $\mathbf{d}(C(H_1), C(H_2)) = 0$ for all $H_1, H_2 \subseteq G$.

Proof. Consider a symmetric state $\sigma \in C(G)$. Then we can construct a state $\sigma_1 = \epsilon\, \rho_1 + (1-\epsilon)\, \sigma$ with $\mathrm{Iso}(\rho_1) = H_1$, therefore $\mathrm{Iso}(\sigma_1) = H_1$ and hence $\sigma_1 \in C(H_1)$. Likewise, for $\rho_2 \in C(H_2)$, the state $\sigma_2 = \epsilon\, \rho_2 + (1-\epsilon)\, \sigma$ is also in $C(H_2)$. As $\epsilon \to 0$, $d(\epsilon\rho_1 + (1-\epsilon)\sigma, \epsilon\rho_2 + (1-\epsilon)\sigma) \to d(\sigma, \sigma) = 0$, therefore $\mathbf{d}(C(H_1), C(H_2)) = 0$ for all $H_1, H_2 \subseteq G$. □

This problem is related to the principal orbit-type theorem [78].

Lemma 4.3: Principal orbit-type theorem for quantum channels

The set of quantum channels that have principal isotropy, meaning that they have the smallest isotropy subgroup, forms a dense subset of all quantum channels.

Almost all quantum states or channels break the symmetry in a maximal way, with the remainder forming a measure zero subset.

Example 4.8: Principal orbit-type theorem for one-qubit states

The principal orbit-type theorem is obvious for a single qubit under SU(2) in Example 4.6. The principal orbit-type (the most common shape) is $\mathrm{SU}(2)/\mathrm{U}(1) \cong S^2$. The only state to not have this shape is the symmetric state $\mathbb{I}/2$. Small perturbations to $\mathbb{I}/2$ introduce a non-zero Bloch vector, recovering the principal orbit-type.

This example shows that the principal orbit-type does not have to correspond to the trivial isotropy subgroup $\{e\}$ – here the smallest isotropy subgroup was U(1). In the next chapter, the

smallest isotropy subgroup for two-qubit states (and by channel-state duality, also for single-qubit channels) under an SU(2) group action is $\mathbb{Z}_2 = \{\pm\mathbb{I}\}$.

Given these arguments, is the 'shape' physically meaningful? The answer is yes, and here are several reasons why the shape of the group orbit is both useful and meaningful.

(Minimal Resources) We are interested in the *minimal* resources required to perform a given operation in the presence of symmetry constraints. In this context, minimisation of the isotropy subgroup by small perturbations does not pose a problem for the analysis of resources.

(Channels & States of Interest) Although the states and channels that do not have principal orbit-type (minimal isotropy subgroup) form a measure zero subset, this nonetheless includes many practically important examples.

Example 4.9: Channels under SU(2) group action

Almost all channels on one qubit have the minimal isotropy subgroup $\mathbb{Z}_2 = \{\mathbb{I}, -\mathbb{I}\}$ under the action of SU(2), however many important channels have non-trivial residual symmetries. This includes symmetric channels, such as the depolarising channel: $\mathcal{E}(\rho) = p\rho + (1-p)(\frac{1}{2}\mathbb{I})$, with $0 \leq p \leq 1$, while examples of channels that are symmetric under a U(1) subgroup of SU(2) includes:

1. Unitary rotation about the $\boldsymbol{n}$ axis: $U_{\boldsymbol{n}}(\theta) = \exp[-i\boldsymbol{n}\cdot\boldsymbol{\sigma}]$.

2. Projective measurement onto the $\boldsymbol{n}$ axis: $\mathcal{E}(\rho) = \sum_k \mathrm{tr}(\Pi_k\rho)\Pi_k$, where $\Pi_0 = |\boldsymbol{n}\rangle\langle\boldsymbol{n}|$ and $\Pi_1 = \mathbb{I} - \Pi_0 = |-\boldsymbol{n}\rangle\langle-\boldsymbol{n}|$.

3. Partial dephasing onto the $\boldsymbol{n}$ axis: $\mathcal{E}(\rho) = p\ \rho + (1-p)\sum_k \mathrm{tr}(\Pi_k\rho)\ \Pi_k$, where $\Pi_0 = |\boldsymbol{n}\rangle\langle\boldsymbol{n}|$ and $\Pi_1 = \mathbb{I} - \Pi_0 = |-\boldsymbol{n}\rangle\langle-\boldsymbol{n}|$.

4. State preparation for $\rho = \frac{1}{2}(\mathbb{I} + p\ \boldsymbol{n}\cdot\boldsymbol{\sigma})$.

(Measure zero sets are not irrelevant) There are important examples throughout physics of phenomena restricted to measure zero subsets. For example, the set of rotationally-symmetric dynamics is measure zero, being unstable under small perturbations, however this is the domain of applicability for Noether's theorem, a result with many important *physical* consequences.

To make the framework robust to problematic perturbations, we only consider symmetries down to some finite resolution scale. This motivates a smoothed version of the isotropy subgroup, using the largest isotropy group found in the neighbourhood of states or channels, where the neighbourhood is defined by some distance measure (such as the L_1-norm).

Definition 4.11: Smoothed isotropy subgroup

Given a state ρ, there is a neighbourhood of states, $\mathcal{B}_\epsilon(\rho) = \{\sigma : d(\rho, \sigma) \leq \epsilon\}$, measured by a distance measure d. The smoothed isotropy subgroup of ρ is

$$\mathrm{Iso}_\epsilon(\rho) := \max_{\sigma \in \mathcal{B}_\epsilon(\rho)} \mathrm{Iso}(\sigma) \tag{4.9}$$

where the maximisation is with respect to the size of the isotropy subgroups. An analogous definition holds for channels.

States sufficiently close to symmetric states are treated as symmetric, negating the unwanted symmetry-breaking effects of small perturbations up to the smoothing scale. The maximisation in this definition gives the most coarse-grained description of the residual symmetries. Consider an $SU(2)$ group action, where there are many different $U(1)$ subgroups, each with different generators. If one of these subgroups is picked as the smoothed isotropy subgroup, the choice of generator is not the primary concern because these different $U(1)$ subgroups induce the same group orbit 'shape'. One could define a smoothing that takes into account these details, but here we are focusing on the simplest, most coarse-grained description of the residual symmetries. In the next chapter, we provide examples of this smoothing within a complete classification of shapes for two-qubit states.

4.6 Conclusions & Outlook

In this chapter, we have introduced a new coarse-grained description of the symmetry properties of quantum states and channels. Furthermore, we have argued that this description captures the basic behaviour of a state or channel under a given symmetry group and is physically meaningful, in the sense that it is robust to perturbations. We have given a couple of simple examples of the 'shapes' that these group orbits can take, but in the next chapter we provide further examples, enumerating all possible shapes for the orbits of two-qubit states under an $SU(2)$ group action.

Despite the coarseness of the 'shape' description, the complexity of the structures imposed by symmetry constraints means that even this basic description holds merit. The remainder of this part of the thesis explores some consequences of this framework and how it can be used as a new type of resource measure, mapping quantum states and channels to a partial order rather than the totally ordered set of real numbers. This is the first step towards a new approach to describing the structure and manipulation of resources, one that aims to capture certain basic features of their relationships without taking into account the full structure.

Some motivations for studying symmetry constraints were given at the beginning of this chapter. It is hoped that the new framework may offer tractable approaches to study symmetries in practically-important settings. For example, the quantum thermodynamics of non-commuting charges involves more complex symmetry constraints, such as an $SU(2)$ constraint, making the analysis and manipulation of quantum resources in this setting difficult. This framework could provide insight into these problems without delving into the mire of complexity.

Chapter 5

Illustrative Example: Two Qubits Under SU(2) Symmetry Constraints

The previous chapter introduced a framework to classify and analyse the structure of quantum states and channels with respect to some symmetry group. To provide intuition and highlight the complexity of the problem, this chapter provides a complete classification of group orbits for two-qubit states under a rigid SU(2) group action. The novel work in this chapter is as follows:

- We show how normal subgroups permit a simpler description of the group orbit. For example, the shapes of SU(2) group orbits can be described in terms of more familiar SO(3) subgroups – rotations in three-dimensional space.
- The isotropy subgroups of T-states under rigid SU(2) transformations are calculated, revealing new structure in this important set of states, pictured in Figure 5.2.
- Building on the T-state classification, the isotropy subgroups of all two-qubit states under the SU(2) group action are described. This is the first full classification of its kind for a widely-studied class of quantum states, albeit at a highly coarse-grained level.

5.1 Two-Qubit States

Two qubits provide a tractable example that still exhibits many non-trivial features under the action of SU(2). Furthermore, this complements an analogous study of quantum channels on two qubits using the structure of irreducible representations [75, 79, 80]. Two-qubit states have been studied in a variety of contexts; their orbits have been analysed in relation to thermodynamics and the study of correlations between subsystems [81, 82].

To simplify our treatment, we utilise the structure of two-qubit states, considering more general states step-by-step. These classes are as follows.

(Bell states) Maximally entangled Bell states,

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2}}\big(\,|01\rangle \pm |10\rangle\,\big) \quad \text{and} \quad |\phi_\pm\rangle = \frac{1}{\sqrt{2}}\big(\,|00\rangle \pm |11\rangle\,\big), \tag{5.1}$$

are a starting point for the study of two qubits, providing a basis for the two-qubit Hilbert space.

(T-states) Convex mixtures of Bell states can be expressed as

$$\rho_T = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_j \tau_j\ \sigma_j \otimes \sigma_j\right). \tag{5.2}$$

Coefficients (τ_1, τ_2, τ_3) specify a position within a tetrahedron in $\mathbb{R}^3$, with vertices $(-1,-1,-1)$, $(1,1,-1)$, $(1,-1,1)$ and $(-1,1,1)$ corresponding to the Bell states [83].

(Maximally mixed marginals) Local unitaries $U_1 \otimes U_2$ act upon T-states to give the full range of two-qubit states with maximally mixed marginals, expressed as

$$\tilde{\rho}_{AB} = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_{i=1}^{3} \tau_i\ \boldsymbol{c}_i\cdot\sigma \otimes \boldsymbol{d}_i\cdot\sigma\right)\ , \tag{5.3}$$

where $\{\boldsymbol{c}_i\}$ and $\{\boldsymbol{d}_i\}$ are orthonormal bases of $\mathbb{R}^3$. The coefficients (τ_1, τ_2, τ_3) remain within the same tetrahedron as for the T-states.

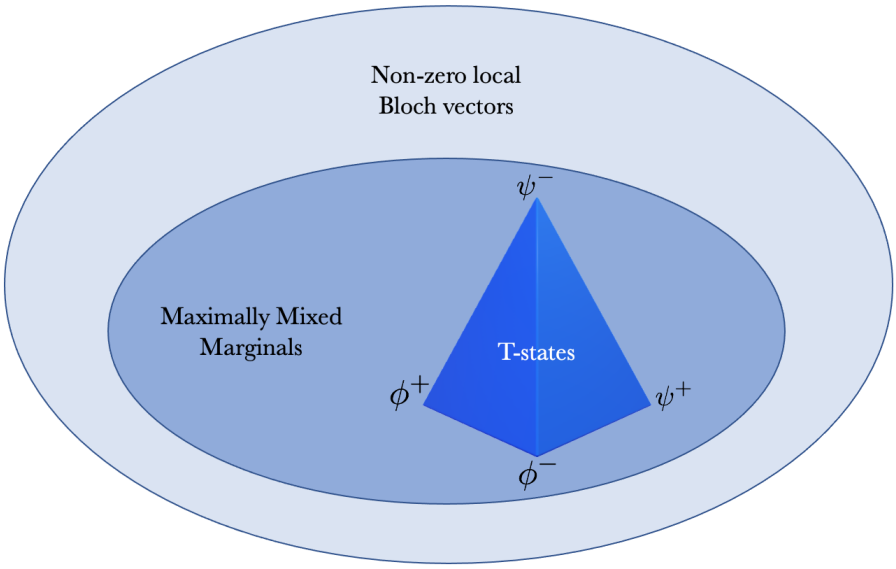


Figure 5.1: Classes of two-qubit states. *In this chapter, we classify the allowed group orbits for two-qubit states under a tensor product SU(2) group action. Starting with the Bell states, we then consider their convex combinations – the T-states. We next move onto the more general class of states with maximally mixed marginals, before adding in local Bloch vectors to include all two-qubit states. This schematic diagram is a roadmap for the complete classification given in this chapter.*

(General two-qubit state) A general two-qubit state can have local Bloch vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, such that

$$\rho_{AB} = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes\boldsymbol{b}\cdot\sigma + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right) \tag{5.4}$$

where $\{\boldsymbol{c}_i\}$ and $\{\boldsymbol{d}_i\}$ are orthonormal bases of $\mathbb{R}^3$, and $|\boldsymbol{a}| \leq 1$ and $|\boldsymbol{b}| \leq 1$.

These classes (shown in Figure 5.1) simplify our description of symmetry properties of two-qubit states, making clear how different components of a given state break the SU(2) symmetry considered in this chapter.

5.2 SU(2) Group Action on Two Qubits

To discuss symmetries of two-qubit states, we must specify a particular group action, in other words how SU(2) acts upon the quantum state space. In this chapter, we restrict our attention to a tensor product representation for the SU(2) group action,

$$\mathcal{U}_g(\rho) = U(g) \otimes U(g)\ \rho\ U^\dagger(g) \otimes U^\dagger(g), \tag{5.5}$$

where $U(g)$ is the defining two dimensional representation of SU(2). This choice is the most relevant for applications of the mathematical framework, particularly in the context of conservation laws and quantum reference frames.

This chapter is concerned with the different sets of residual symmetries that two-qubit states can exhibit, using the framework of isotropy subgroups and group orbits introduced previously. The shape of the group orbit, $G/\mathrm{Iso}(\rho)$, eliminates redundancies from the group manifold G, quotienting out the residual symmetries of the state. Sometimes this description admits further simplification. Our SU(2) action has a subgroup $\mathbb{Z}_2 = \{\pm\mathbb{I}\}$ that is always a symmetry of a two-qubit state, and which is associated with SU(2) being the double cover of SO(3). Since this symmetry is always present, we eliminate this subgroup from our description. Mathematically, this is framed in terms of normal subgroups [55].

Definition 5.1: Normal subgroup

N is a normal subgroup of G, denoted $N \triangleleft G$, if $gNg^{-1} = N$ for all $g \in G$.

This leads to a well-known result in group theory, where the set of cosets themselves constitute a group [55].

Lemma 5.1: Quotient group

Given a normal subgroup $N \triangleleft G$, the group G can be partitioned into equivalence classes $gN := \{gn_i : \ n_i \in N\}$, which are elements of the quotient group G/N under the composition law $(g_1N)(g_2N) = (g_1g_2)N$.

In other words, elements of G are uniquely specified by the pairing (gN, n) of a coset gN and an element n of the normal subgroup N. Moreover, the cosets gN themselves constitute a new group called G/N. In contrast, for a $H \subseteq G$ that is not a normal subgroup, the coset space G/H will not be a group. Therefore the group orbit for a state or channel will only be diffeomorphic to a group manifold if the corresponding isotropy subgroup is a normal subgroup of the full group action. We can often further simplify the description of the group orbits, eliminating the additional redundancy of a normal subgroup.

Lemma 5.2: Quotienting out a normal subgroup

If $N \triangleleft H$ and $N \triangleleft G$ for a subgroup $H \subseteq G$, there is a bijective mapping $f : G/H \to (G/N)/(H/N)$ that respects the group action of G on G/H, such that if $g \in G$ maps $g_i H \to (gg_i)H$ then $f(g_i H) \to f(gg_i H)$. In terms of describing the effect of the group action, $G/H \cong (G/N)/(H/N)$.

Proof. Consider the map $f : G/H \to (G/N)/(H/N)$ such that $f(gH) = (gN)(H/N)$. This map is surjective because ranging over $g \in G$ reaches all possible $(gN)(H/N) \in (G/N)/(H/N)$. To show that f is injective, consider $g_1 H \neq g_2 H$. These are mapped to

$$f(g_1 H) = \{(g_1 h_1)N, (g_1 h_2)N, (g_1 h_3)N, \dots\} \tag{5.6}$$

$$f(g_2 H) = \{(g_2 h_1)N, (g_2 h_2)N, (g_2 h_3)N, \dots\} \, . \tag{5.7}$$

If $f(g_1 H) = f(g_2 H)$, then $(g_1 h_1)N = (g_2 h_i)N$ for some $h_i \in H$. This implies that $h_i^{-1} g_2^{-1} g_1 h_1 N = N$, therefore $h_i^{-1} g_2^{-1} g_1 h_1 \in N$. Furthermore, $N \triangleleft H$, therefore $h_i^{-1} g_2^{-1} g_1 h_1 \in H$, which implies that $g_2^{-1} g_1 \in H$ and then $g_1 \in g_2 H$. However, this means that $g_1 H = g_2 H$ because cosets partition the group G, contradicting the assumption that $g_1 H \neq g_2 H$, hence f must be injective. The map f is a bijection that respects the group action, since G acts on G/H such that $g_i H \to (gg_i)H$, while G/N acts on $(G/N)/(H/N)$ such that $f(g_i H) \to (gN)(g_i N)(H/N) = (gg_i N)(H/N) = f(gg_i H)$. We therefore conclude that $G/H \cong (G/N)/(H/N)$. □

Intuitively, the normal subgroup 'cancels out' from the description of the group orbits. In this situation, the action of the group G on a quantum state or channel can be more simply described in terms of the quotient group G/N. A concrete example of this is provided by the SU(2) case considered in this chapter.

Example 5.1: SU(2) to SO(3)

All two-qubit quantum states are invariant under the normal subgroup $\mathbb{Z}_2 = \{\pm\mathbb{I}\}$, therefore $\mathcal{M}(\rho) \cong \mathrm{SU}(2)/\mathrm{Iso}(\rho) \cong \mathrm{SO}(3)/(\mathrm{Iso}(\rho)/\mathbb{Z}_2)$. The SU(2) group orbits can be more intuitively described in terms of quotients derived from SO(3), and we can instead refer to the more familiar *chiral point groups* – subgroups of SO(3), listed in Table 5.1, that play a central role in crystallography [55].

Group	Description
C_n	Cyclic group of order n
C_∞	Symmetries of Cone
D_n	Symmetries of a regular n-sided Polygon
D_∞	Symmetries of Cylinder
T	Symmetries of Tetrahedron
O	Symmetries of Cube/Octahedron
I	Symmetries of Dodecahedron/Icosahedron

Table 5.1: Chiral point groups: subgroups of SO(3). *Their relation to simple three-dimensional solid bodies makes them more familiar than their SU(2) counterparts.*

In the remainder of this chapter, we enumerate the possible isotropy subgroups for two-qubit states under an SU(2) group action. We shall begin with the Bell states and T-states, then work towards a full classification.

5.3 Symmetries of Bell States under SU(2)

Rather than diving straight into a classification of general two-qubit states, we first consider familiar subsets, then gradually increase the generality of the states being considered. Bell states appear in many areas of quantum information theory, useful for studying the properties and dynamics of two qubits. Furthermore, they are the extremal points in the tetrahedron of T-states, which we discuss later in this chapter.

Identifying the isotropy subgroups of the Bell states is aided by vectorization [84], which transforms a matrix into a vector, such that $\text{vec}\big(|i\rangle\langle j|\big) := |i\rangle|j\rangle$ for computational basis vectors $|i\rangle$ and $|j\rangle$. Bell states are the vectorization of Pauli matrices, with

$$\left|\phi^+\right\rangle \propto |\text{vec}(\mathbb{I})\rangle \qquad \left|\phi^-\right\rangle \propto |\text{vec}(\text{Z})\rangle \qquad \left|\psi^+\right\rangle \propto |\text{vec}(\text{X})\rangle \qquad \left|\psi^-\right\rangle \propto |\text{vec}(\text{Y})\rangle\,. \tag{5.8}$$

To find the isotropy subgroups, we must identify which group elements give $\mathcal{U}_g(\rho) = \rho$. For Bell states, this occurs when $\mathcal{U}_g\big(|\text{vec}(\sigma_i)\rangle\langle\text{vec}(\sigma_i)|\big) = |\text{vec}(\sigma_i)\rangle\langle\text{vec}(\sigma_i)|$, or equivalently

$$U(g)\otimes U(g)\,|\text{vec}(\sigma_i)\rangle = e^{i\phi}\,|\text{vec}(\sigma_i)\rangle\ , \tag{5.9}$$

where $\phi \in \mathbb{R}$ is some arbitrary phase. Since $A\otimes B\,|\text{vec}(M)\rangle = \left|\text{vec}(AMB^T)\right\rangle$ [84], elements of the isotropy subgroup satisfy

$$U(g)\sigma_i U^T(g) = e^{i\phi}\sigma_i, \tag{5.10}$$

which we use to find the isotropy subgroups for the corresponding Bell state.

Under the tensor product SU(2) group action, ψ^- has the simplest symmetry properties – it is a symmetric state.

Lemma 5.3: ψ^- is symmetric

Under the tensor product SU(2) group action, $\text{Iso}(\psi^-) = \text{SU}(2)$.

Proof. For SU(2), $U(g) = \exp(i\varphi[aX + bY + cZ])$ with $a^2 + b^2 + c^2 = 1$ $(a, b, c \in \mathbb{R})$ and $0 \le \varphi < 2\pi$, therefore for $g \in \text{Iso}(\psi^-)$, $U(g)YU^T(g) = e^{i\phi}Y$, or $U(g)YU^T(g)Y = e^{i\phi}\mathbb{I}$. On the

left hand side,

$$YU^T(g)Y = Y\big(\cos\varphi\ \mathbb{I} + i\sin\varphi\ (aX + bY + cZ)\big)^T Y \tag{5.11}$$

$$= Y\big(\cos\varphi\ \mathbb{I} + i\sin\varphi\ (aX - bY + cZ)\big)Y \tag{5.12}$$

$$= \cos\varphi\ \mathbb{I} + i\sin\varphi\ (aYXY - bYYY + cYZY) \tag{5.13}$$

$$= \cos\varphi\ \mathbb{I} - i\sin\varphi\ (aX + bY + cZ) = U^\dagger(g) \tag{5.14}$$

because $Y^T = -Y$, so the condition becomes $U(g)U^\dagger(g) = e^{i\phi}\mathbb{I}$, or $U(g) = e^{i\phi}U(g)$, which gives

$$e^{i\varphi(aX+bY+cZ)} = e^{i\phi}e^{i\varphi(aX+bY+cZ)}. \tag{5.15}$$

This is satisfied for all elements of SU(2) when $\phi = 0$, therefore $\mathrm{Iso}(\psi^-) = \mathrm{SU}(2)$. □

Since ψ^- is a symmetric state, its group orbit $\mathcal{M}(\psi^-) = \{e\}$ is a point. The triplet Bell states do break the SU(2) symmetry to some extent, and have group orbits with a non-trivial shape. The proof is somewhat more involved, but amounts to finding the SU(2) group parameters that satisfy Eq. (5.10). Those group parameters then correspond to the residual symmetries of the triplet Bell states.

Lemma 5.4: Isotropy subgroups for ϕ^+, ϕ^- & ψ^+

Under the tensor product SU(2) group action:

$$\mathrm{Iso}(\phi^+) = \left\{(iZ)^\alpha e^{i\theta Y} :\ 0 \le \theta < 2\pi,\ \alpha = 0,1,2,3\right\}, \tag{5.16a}$$

$$\mathrm{Iso}(\phi^-) = \left\{(iY)^\alpha e^{i\theta X} :\ 0 \le \theta < 2\pi,\ \alpha = 0,1,2,3\right\}, \tag{5.16b}$$

$$\mathrm{Iso}(\psi^+) = \left\{(iX)^\alpha e^{i\theta Z} :\ 0 \le \theta < 2\pi,\ \alpha = 0,1,2,3\right\}. \tag{5.16c}$$

Proof. Given that $|\phi^+\rangle \propto |\mathrm{vec}(\mathbb{I})\rangle$, an SU(2) group element belongs to $\mathrm{Iso}(\phi^+)$ if

$$e^{i\varphi(aX+bY+cZ)}e^{i\varphi(aX-bY+cZ)} = e^{i\phi}\ \mathbb{I}, \tag{5.17}$$

or equivalently $e^{i\varphi(aX+bY+cZ)} = e^{i\phi}e^{i\varphi(-aX+bY-cZ)}$. Expanding the matrix exponentials,

$$\cos\varphi\mathbb{I} + i\sin\varphi(aX + bY + cZ) = e^{i\phi}[\cos\varphi\mathbb{I} + i\sin\varphi(-aX + bY - cZ)]\ , \tag{5.18}$$

which requires

$$(1-e^{i\phi})\cos\varphi = (1+e^{i\phi})\ ia\sin\varphi = (1-e^{i\phi})\ ib\sin\varphi = (1+e^{i\phi})\ ic\sin\varphi = 0. \tag{5.19}$$

The value of ϕ is heavily constrained. There are two cases:

- $(\cos\varphi \neq 0)$: Needing $(1-e^{i\phi})\cos\varphi = 0$ implies $\phi = 0$ (modulo 2π).
- $(\cos\varphi = 0)$: At least one of a, b or c is non-zero because $a^2+b^2+c^2 = 1$. Since $\sin\varphi \neq 0$, either $1+e^{i\phi} = 0$ or $1-e^{i\phi} = 0$, therefore $\phi = 0$ or π (modulo 2π).

We consider these two possibilities for ϕ in turn.

1. When $\phi = 0$, the conditions reduce to $2ia\sin\varphi = 2ic\sin\varphi = 0$. Two possibilities satisfy these conditions:
   - $a = c = 0$, leaving φ unconstrained and $b = \pm 1$, giving $U(g) = e^{i\varphi Y}$.
   - $\sin\varphi = 0$, so $\varphi = 0, \pi$, which gives $U(g) = \pm\mathbb{I}$.
2. In the $\phi = \pi$ case, the conditions become $2\cos\varphi = 2ib\sin\varphi = 0$. Clearly $\cos\varphi = 0$, which gives $\varphi = \pi/2$ or $3\pi/2$. If $b \neq 0$, then $\cos\varphi = \sin\varphi = 0$, which is not possible, therefore $b = 0$, while a and c satisfy $a^2+c^2 = 1$. When $\varphi = \pi/2$,

   $$U(g) = i(aX+cZ) = (iZ)(c\mathbb{I}+iaY) = (iZ)(\cos\theta\mathbb{I}+i\sin\theta Y) = (iZ)e^{i\theta Y} \tag{5.20}$$

   by parameterising in terms of θ. Similarly $U(g) = (-iZ)e^{i\theta Y}$ for the $\varphi = 3\pi/2$ case.

In summary,

$$\mathrm{Iso}(\phi^+) = \left\{ (\mathrm{iZ})^\alpha \mathrm{e}^{\mathrm{i}\theta\mathrm{Y}} :\ 0 \leq \theta < 2\pi,\ \alpha = 0,1,2,3 \right\}. \tag{5.21}$$

The proof for ϕ^- and ψ^+ runs analogously. □

Vectorisation aids finding the isotropy subgroups of certain states, and one can extend this proof to Bell states in alternative bases. These results also provide a starting point for classifying the residual symmetries of convex combinations of the Bell states – a wider class of two-qubit

states called T-states. However before moving on to that, we discuss the form of the non-trivial isotropy subgroups for the triplet Bell states.

5.3.1 Aside: Semi-Direct Product Groups

The isotropy subgroups of the triplet Bell states are isomorphic to one another and we denote this subgroup K_∞, consisting of a U(1) subgroup generated by one of the Pauli matrices, coupled with a discrete subgroup of rotations 'perpendicular' to that axis. The structure of K_∞ is an example of a semi-direct product group [55].

Example 5.2: Euclidean group

The Euclidean group – rotations and translations in three-dimensional space – has a semi-direct product structure. Each element can be decomposed as a rotation followed by a translation, but note that these two subgroups are not treated independently by the group composition rule.

$\mathrm{K}_\infty \cong \mathrm{U}(1) \rtimes \mathbb{Z}_4$ couples

$$\mathrm{U}(1) = \left\{ e^{i\theta Y} : \ 0 \leq \theta < 2\pi \right\} \tag{5.22}$$

and the discrete group

$$\mathbb{Z}_4 = \{ (iZ)^\alpha : \ \alpha = 0, 1, 2, 3 \} . \tag{5.23}$$

Each element of K_∞ can be thought of as a pairing (n, h) where $n \in N = \mathrm{U}(1)$ and $h \in H = \mathbb{Z}_4$, although they are not treated independently when composing group elements.

Definition 5.2: Semi-direct product group

The semi-direct product group $N \rtimes H$ is the set of pairs (h, n), where $n \in N$ and $h \in H$, with composition law

$$(n_1, h_1)(n_2, h_2) = \left(\ f_{h_2}(n_1) n_2 \ , \ h_1 h_2 \ \right) , \tag{5.24}$$

where $f_h : N \to N$ is a bijective mapping on N specified by an element of H.

The choice of mappings f_h defines the structure of the semi-direct product group. A special case is when $f_h(n) = n$ for all $h \in H$, a construction called the direct product group [55].

Example 5.3: Multiplication law for K_∞

For $\mathrm{K}_\infty = \mathrm{U}(1) \rtimes \mathbb{Z}_4$, the group multiplication law is

$$\left[(iZ)^{\alpha_1} e^{i\theta_1 Y}\right] \left[(iZ)^{\alpha_2} e^{i\theta_2 Y}\right] = (iZ)^{\alpha_1+\alpha_2} \; e^{i[(-1)^{\alpha_2}\theta_1+\theta_2]Y}, \tag{5.25}$$

clearly demonstrating the interdependence of the subgroups in the pairing.

The group orbits of the triplet Bell states have the same shape, namely $\mathcal{M}(\phi^+) \cong \mathcal{M}(\phi^-) \cong \mathcal{M}(\psi^+) \cong \mathrm{SU}(2)/\mathrm{K}_\infty$. As discussed earlier, this can be more intuitively expressed in terms of SO(3) subgroups, so that $\mathrm{SU}(2)/\mathrm{K}_\infty \cong \mathrm{SO}(3)/D_\infty$. Effectively, this corresponds to discarding the effect of $\alpha = 2, 3$ in the $\mathbb{Z}_4$ part of K_∞ when specifying the shape of the group orbit, eliminating the $\{\pm\mathbb{I}\}$ normal subgroup. In this sense, the $\mathbb{Z}_4$ subgroup of K_∞ acts like the 180° rotations perpendicular to the continuous rotations in D_∞. Therefore the triplet Bell states under a tensor product SU(2) group action behave much like cylinders under three-dimensional rotations.

5.4 T-States

The next stage in our classification of two-qubit states is to consider the T-states, consisting of convex combinations of the Bell states,

$$\rho_T = p_1\phi^+ + p_2\phi^- + p_3\psi^+ + p_4\psi^-, \tag{5.26}$$

where $p_i \geq 0$ and $p_1 + p_2 + p_3 + p_4 = 1$. These states can be represented as points within a regular tetrahedron, with the vertices corresponding to the Bell states.

Despite the T-states filling a three-dimensional tetrahedron, their symmetry properties are fully exhibited by a two-dimensional triangular slice. To see why, consider the triplet Bell states as the base of the tetrahedron, and the singlet state ψ^- as the peak. The vertical height of a

point in this tetrahedron is determined by the proportion p_4 contributed by the singlet state. However, ψ^- is a symmetric state, therefore

$$\mathcal{M}(p_1\phi^+ + p_2\phi^- + p_3\psi^+ + p_4\psi^-) = \mathcal{M}(p_1\phi^+ + p_2\phi^- + p_3\psi^+) \tag{5.27}$$

under the SU(2) group action. The full range of shapes for T-state group orbits is therefore exhibited within the convex hull of the triplet Bell states, as visualised in Figure 5.2

The triplet Bell states have K_∞ isotropy subgroups, and it is readily verified that states on the heights of the triangle defined by those Bell states have the same residual symmetries. Finally, all T-states are invariant under the subgroup $\mathrm{K}_2 = \{\pm\mathbb{I}, \pm iX, \pm iY, \pm iZ\}$. This exhausts the possible isotropy subgroups for the T-states, as confirmed in Appendix B. To summarise, the possible isotropy subgroups (and shapes of group orbit) for T-states under SU(2) are:

- $\mathrm{Iso}(\rho_T) = \mathrm{SU}(2) \Rightarrow \mathcal{M}(\rho) \cong \{e\}$: Werner states, positioned on the central axis of the T-state tetrahedron, with $\tau_1 = \tau_2 = \tau_3$, or equivalently $\rho = (1-p)\ \psi^- + \frac{p}{3}\ (\psi^+ + \phi^+ + \phi^-)$ for $0 \leq p \leq 1$.

- $\mathrm{Iso}(\rho_T) = \mathrm{K}_\infty \Rightarrow \mathcal{M}(\rho) \cong \mathrm{SO}(3)/\mathrm{D}_\infty$: T-states with coefficients $\tau_i = \tau_j \neq \tau_k$, sitting on or above the heights of the triplet Bell state triangle (excluding the Werner states on the central axis of the tetrahedron).

- $\mathrm{Iso}(\rho_T) = \mathrm{K}_2 \Rightarrow \mathcal{M}(\rho) \cong \mathrm{SO}(3)/\mathrm{D}_2$: T-states with coefficients $\tau_1 \neq \tau_2 \neq \tau_3$, that lie off the heights of the triplet Bell state triangle.

Furthermore, this classification can be made robust to small perturbations, because the smoothed isotropy subgroups follow the same pattern in the tetrahedron, as illustrated in Figure 5.2 for a smoothing scale of $\epsilon = 0.04$ under the trace distance. However, note that the horizontal slices are no longer equivalent under the smoothing, since near the top of the tetrahedron (ψ^-) all states are assigned an SU(2) isotropy subgroup, while at the base there remain states with $\mathrm{Iso}_\epsilon(\rho) = \mathrm{K}_2$.

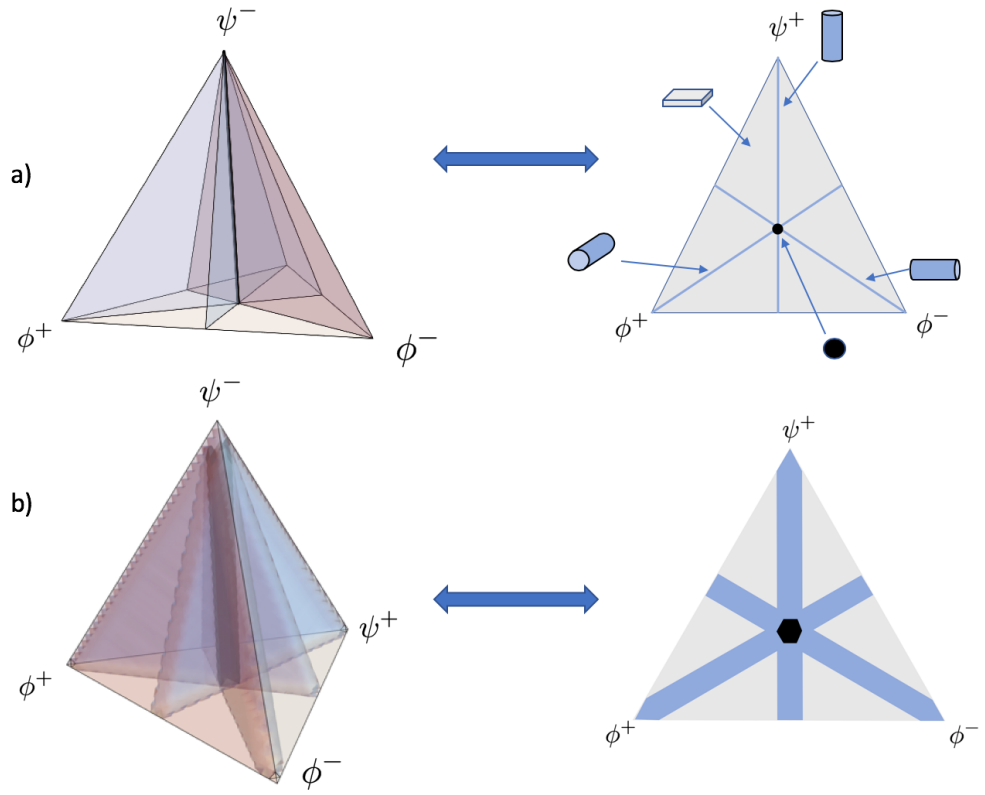


Figure 5.2: The partitioning of the T-states by isotropy subgroup. *a) Tetrahedron of T-states coloured to indicate different group orbit shapes: grey –* $\mathrm{SU}(2)/\mathrm{K}_2 \cong \mathrm{SO}(3)/\mathrm{D}_2$*, light blue –* $\mathrm{SO}(3)/\mathrm{D}_\infty$*, black –* $\{e\}$*. Horizontal slices through the tetrahedron, parallel to the plane of the triplet Bell states, exhibit the same structure, exhibited in the triangle to the right. To provide intuition, the regions are labelled by three-dimensional shapes with the same residual symmetries. b) Tetrahedron mapping out the smoothed isotropy subgroups with smoothing scale* $\epsilon = 0.04$ *under the trace distance. The bands along the heights indicate* $\mathrm{Iso}_\epsilon(\rho) \cong \mathrm{K}_\infty$*, while states in the central hexagon are assigned* $\mathrm{Iso}_\epsilon(\rho) \cong \mathrm{SU}(2)$*. The remaining states retain their* K_2 *isotropy subgroup under smoothing. Reproduced from [1].*

5.5 Maximally Mixed Marginals

The T-states are a natural starting point to consider a more general class of two-qubit states. Any two-qubit states with maximally mixed marginals can be reached from T-states via local SU(2) unitaries $U(g_1)\otimes U(g_2)$. There are two cases to consider here: (1) states reached by equal local unitaries on each qubit, and (2) states reached by unequal local unitaries on each qubit.

5.5.1 Equal Local Unitaries

Obtained from T-states via rigid local SU(2) unitaries $U(g)\otimes U(g)$, these are T-states in a rotated Pauli basis, and have the form

$$\tilde{\rho}_T = \mathcal{U}_g(\rho_T) = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma\right), \tag{5.28}$$

where the vectors $\{\boldsymbol{c}_i\}$ form an orthonormal basis of $\mathbb{R}^3$. As expected, these states have the same structure as the T-states, with isotropy subgroups that are isomorphic to those for the T-states. This is an example of the following lemma.

Lemma 5.5: Isotropy subgroups in 'rotated' frames

For a channel $\mathcal{E}$, $\mathrm{Iso}\big(\mathfrak{U}_g(\mathcal{E})\big) = g[\mathrm{Iso}(\mathcal{E})]g^{-1}$ for any $g\in G$. Similarly, for a state ρ, $\mathrm{Iso}\big(\mathcal{U}_g(\rho)\big) = g[\mathrm{Iso}(\rho)]g^{-1}$ for any $g\in G$.

Proof. Assume $h\in\mathrm{Iso}\big(\mathfrak{U}_g(\mathcal{E})\big)$. Then $\mathcal{U}_h\circ\mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}}\circ\mathcal{U}_{h^{-1}} = \mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}}$, or equivalently $\mathcal{U}_{(g^{-1}hg)}\circ\mathcal{E}\circ\mathcal{U}_{(g^{-1}hg)^{-1}} = \mathcal{E}$. Therefore $g^{-1}hg\in\mathrm{Iso}(\mathcal{E})$, or $h\in g[\mathrm{Iso}(\mathcal{E})]g^{-1}$. When $gh'g^{-1}\in g[\mathrm{Iso}(\mathcal{E})]g^{-1}$ acts upon $\mathfrak{U}_g(\mathcal{E})$, we have

$$(\mathcal{U}_g\circ\mathcal{U}_{h'}\circ\mathcal{U}_{g^{-1}})\circ(\mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}})\circ(\mathcal{U}_g\circ\mathcal{U}_{h'^{-1}}\circ\mathcal{U}_{g^{-1}}) = \mathcal{U}_g\circ\mathcal{U}_{h'}\circ\mathcal{E}\circ\mathcal{U}_{h'^{-1}}\circ\mathcal{U}_{g^{-1}} \tag{5.29}$$

$$= \mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}}, \tag{5.30}$$

implying $gh'g^{-1}\in\mathrm{Iso}(\mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}})$, hence $\mathrm{Iso}(\mathcal{U}_g\circ\mathcal{E}\circ\mathcal{U}_{g^{-1}}) = g[\mathrm{Iso}(\mathcal{E})]g^{-1}$. An analogous proof holds for states. □

Just as T-states are invariant under the subgroup $\mathrm{K}_2 = \{\pm\mathbb{I}, \pm iX, \pm iY, \pm iZ\}$, these rotated T-states are invariant under the isomorphic subgroup $\mathrm{K}_2 = \{\pm\mathbb{I},\ \pm i(\boldsymbol{c}_1\cdot\sigma),\ \pm i(\boldsymbol{c}_2\cdot\sigma),\ \pm i(\boldsymbol{c}_3\cdot\sigma)\}$. Therefore the group orbit shapes are the same as for the T-states.

5.5.2 Unequal Local Unitaries

There exist two-qubit states with maximally mixed marginals that cannot be accessed from the T-states under the tensor product SU(2) group action, requiring unequal local unitaries $U(g_1) \otimes U(g_2)$ on each qubit to reach them. Such states can be expressed as

$$\tilde{\rho}_{AB} = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_{i=1}^{3} \tau_i\ \boldsymbol{c}_i\cdot\sigma \otimes \boldsymbol{d}_i\cdot\sigma\right). \tag{5.31}$$

where $\{\boldsymbol{c}_i\}$ and $\{\boldsymbol{d}_i\}$ are two different orthonormal bases of $\mathbb{R}^3$. This opens the door to new isotropy subgroups.

Case 1: $U(g_1)$ and $U(g_2)$ share a generator. Suppose that $U(g_1)\otimes U(g_2) = e^{i\theta_1 \boldsymbol{r}\cdot\sigma} \otimes e^{i\theta_2 \boldsymbol{r}\cdot\sigma}$ with $\theta_1 \neq \theta_2$. Acting on a (rotated) T-state, this gives

$$\rho_M = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \tau_1\ \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \sum_{i=2}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right) \tag{5.32}$$

where $\{\boldsymbol{r}, \boldsymbol{c}_2, \boldsymbol{c}_3\}$ and $\{\boldsymbol{r}, \boldsymbol{d}_2, \boldsymbol{d}_3\}$ are two sets of orthonormal bases for $\mathbb{R}^3$. These states are invariant under a $\mathbb{Z}_4 = \{\pm\mathbb{I}, \pm i(\boldsymbol{r}\cdot\sigma)\}$ subgroup, and there are two possibilities for the isotropy subgroup:

- $\mathrm{Iso}(\rho_\mathrm{M}) = \mathrm{U}(1) \Rightarrow \mathcal{M}(\rho_\mathrm{M}) \cong \mathrm{S}^2$: States with $\tau_2 = \tau_3$, reached from (rotated) T-states with $\mathrm{Iso}(\rho_\mathrm{T}) = \mathrm{K}_\infty$.
- $\mathrm{Iso}(\rho_\mathrm{M}) = \mathbb{Z}_4 \Rightarrow \mathcal{M}(\rho_\mathrm{M}) \cong \mathrm{SO}(3)/\mathrm{C}_2$: States with $\tau_2 \neq \tau_3$, reached from (rotated) T-states with $\mathrm{Iso}(\rho_\mathrm{T}) = \mathrm{K}_2$.

Case 2: $U(g_1)$ and $U(g_2)$ have different generators. Acting on (rotated) T-states with local SU(2) unitaries $U(g_1)\otimes U(g_2) = e^{i\theta_1 \boldsymbol{r}_1\cdot\sigma} \otimes e^{i\theta_2 \boldsymbol{r}_2\cdot\sigma}$, where $|\boldsymbol{r}_1| = |\boldsymbol{r}_2| = 1$ and $\boldsymbol{r}_1 \neq \boldsymbol{r}_2$,

reaches states of the form

$$\tilde{\rho}_M = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_{i=1}^{3} \tau_i \; \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right), \tag{5.33}$$

where $\{\boldsymbol{c}_i\}$ and $\{\boldsymbol{d}_i\}$ are different orthonormal bases of $\mathbb{R}^3$ with no shared elements, i.e. $\boldsymbol{c}_i \neq \boldsymbol{d}_j$ for all i and j. This gives states that are only invariant under $\mathbb{Z}_2 = \{\pm\mathbb{I}\}$, so a new shape of group orbit is observed:

- $\mathrm{Iso}(\rho) = \mathbb{Z}_2 \Rightarrow \mathcal{M}(\rho) \cong \mathrm{SO}(3)$: States of the form $\tilde{\rho}_M$.

5.6 Local Bloch Vectors

To complete our classification of the residual symmetries of two-qubit states, and hence the allowed group orbit shapes, we must include the effect of local Bloch vectors. A general two-qubit state can be written

$$\rho = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes\boldsymbol{b}\cdot\sigma + \sum_{i=1}^{3} \tau_i \; \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right), \tag{5.34}$$

where $\{\boldsymbol{c}_i\}$ and $\{\boldsymbol{d}_i\}$ are orthonormal bases of $\mathbb{R}^3$, and the local Bloch vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ have length less than or equal to 1. No new shapes of group orbit are introduced, but for completeness we describe the possible isotropy subgroups attained with non-zero local Bloch vectors.

- $\mathrm{Iso}(\rho) = \mathbb{Z}_2 \Rightarrow \mathcal{M}(\rho) \cong \mathrm{SO}(3)$:

  1. $\boldsymbol{a}$ and $\boldsymbol{b}$ are not parallel, i.e. $\boldsymbol{a} \neq \gamma\boldsymbol{b}$ for any $\gamma \in \mathbb{R}$.

  2. The local Bloch vectors are parallel, but not with $\boldsymbol{c}_i$ or $\boldsymbol{d}_i$.

- $\mathrm{Iso}(\rho) = \mathbb{Z}_4 \Rightarrow \mathcal{M}(\rho) \cong \mathrm{SO}(3)/\mathrm{C}_2$: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}_1$ and $\boldsymbol{d}_1$ are parallel but $\tau_2 \neq \tau_3$.

- $\mathrm{Iso}(\rho) = \mathrm{U}(1) \Rightarrow \mathcal{M}(\rho) \cong \mathrm{S}^2$: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}_1$ and $\boldsymbol{d}_1$ are parallel and $\tau_2 = \tau_3$.

5.7 Conclusions & Outlook

This completes the classification of two-qubit group orbits under a tensor product SU(2) group action. A full derivation of this classification is presented in Appendix B, making extensive use of H-twirling, a mathematical tool introduced in the next chapter. The classification also provides a plethora of examples of new shapes for group orbits, going beyond the single-qubit examples.

The backbone of the classification builds on the calculation of isotropy subgroups for T-states under the SU(2) group action. In doing so, a new level of structure in this widely-studied set of states was revealed, as pictured in 5.2, and suggesting that these techniques may reveal new features in other practically important sets of quantum states. Furthermore, we demonstrate that this structure is robust to small perturbations, supporting the arguments given at the end of the previous chapter for why the shape of the group orbit is a physically meaningful description of the symmetry properties of quantum states and channels.

Although the consequences of the this classification are not explored here, since our focus is on the abstract structures that exist when considering different resources, two-qubit states are a common toy model to explore quantum-mechanical features. For example, two-qubit systems provide a useful test-bed to study quantum thermodynamics with non-commutative observables. The simplest non-commutative continuous group to apply in this toy model would be SU(2), and the relation to spin makes this a physically important scenario, therefore this classification will provide useful insights for those explorations.

The allowed isotropy subgroups for two qubits under the SU(2) group action can be summarised in a simple lattice, pictured in Figure 5.3. This Hasse diagram shows the relationship between the available subgroups, with links indicating the inclusion of one group in the other (up to isomorphism). The next chapter focuses on this in greater detail, discussing how the group orbit shape acts as a new type of coarse-grained resource measure that retains aspects of

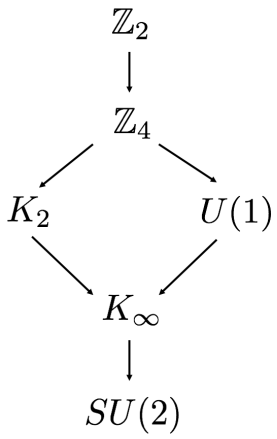


Figure 5.3: Lattice of isotropy subgroups for two-qubit states under SU(2) group action. *The arrows denote subgroup (subset) inclusion (up to isomorphism). This is a subset of the full SU(2) subgroup lattice. Reproduced from [1].*

the full structure of the resource theory without reducing it down to totally ordered sets.

The complexity of the classification and the effort needed to map out the complete set of possibilities further highlights the difficulty in understanding the structure and manipulation of resources based on symmetry constraints. In the next chapter, we discuss how this framework provides valuable coarse-grained information about the relevant resources.

Chapter 6

'Shape' as a Superselection Property

In the previous two chapters, we introduced a classification of quantum states and channels under the action of a given symmetry group, and demonstrated it for two-qubit states under an SU(2) action. So far, our treatment has merely provided a taxonomy of residual symmetries. However, our motivation was to better understand the dynamics of quantum systems constrained by symmetry principles, and furthermore to understand how certain states can be used to lift the superselection rules induced by that symmetry. In this chapter, we explore how our framework sheds light on the structure and manipulation of resources in this context.

- We motivate the 'shape' framework through the study of dynamics under symmetry constraints, relevant to important physical scenarios.
- We discuss to what extent the partial ordering of subgroups provides a coarse-grained resource theory. We also introduce H-twirling as a useful tool in the study of this coarse-grained structure.
- Finally, we show how the 'shape' of the group orbit characterises the resources required to lift the symmetry-induced superselection rules and implement new operations.

The novel contributions of this chapter are as follows.

- We show how group orbit shapes admit a partial ordering, lifted from the partial ordering

of the subgroup lattice, and show how the group orbit behaves under tensor product and composition of channels.

- H-twirling is defined as a generalisation of G-twirling where only a subgroup is averaged over. Some properties of the H-twirling map are proved in Section 6.3, including when H-twirling is a symmetric operation (Lemma 6.4).

- In Theorems 6.2 and 6.3 we prove that the group orbit shape of a channel constrains the resources able to implement that channel, and that this bound is tight in the sense that there exists a suitable resource with the same group orbit shape as that channel.

6.1 Motivation – Dynamics under Symmetry Constraints

Symmetry constrains the dynamics of systems in almost all areas of physics. Such constraints arise for a number of different reasons, both practical and fundamental, with important consequences.

Example 6.1: Lacking a shared reference frame

Alice and Bob use spin-1/2 particles to communicate between their labs, each measuring and preparing the particles with their own Stern-Gerlach apparatus. They each prepare states relative to the z-direction in their respective labs. Suppose Alice prepares and sends the single-particle state ρ to Bob. If Bob does not know the relative alignment of the labs, the state he receives from Alice is

$$\int_G dg\, \mathcal{U}_g(\rho) = \frac{\mathbb{I}}{2}\,, \tag{6.1}$$

averaged uniformly over different orientations [12]. This problem is pertinent to space-based quantum technologies, where one must take into account possible uncertainties in the alignment between distant satellites.

Example 6.2: Symmetry constraints in thermodynamics

Resource-theoretic approaches to thermodynamics often restrict to thermal operations [85],

$$\mathcal{E}(\rho_A) = \mathrm{tr}_B \left[U_{AB}(\rho_A \otimes \gamma_B) U_{AB}^\dagger \right] \ , \tag{6.2}$$

where γ_B is a thermal (Gibbs) state with inverse temperature β, and $[H_{AB}, U_{AB}] = 0$. This final condition relates to time-translation symmetry, globally conserving energy [72]. Attempts to include multiple non-commuting observables into the thermodynamic framework often follow this symmetry-led approach [86].

Resource theories offer a fruitful way to frame symmetry constraints, restricting to symmetric operations and the preparation of symmetric states. This is called the resource theory of reference frames, or more generally the resource theory of asymmetry [12, 80, 87–90].

The central problem within this framework is understanding when there exists a symmetric operation that transforms one state into another. The standard approach is to identify functions $M : \mathcal{B}(\mathcal{H}) \to \mathbb{R}$ such that $M(\rho) \geq M(\sigma)$ if $\sigma = \mathcal{E}(\rho)$ for some symmetric operation $\mathcal{E}$ [87, 88]. These functions are called resource measures or monotones. Usually a single measure will not encode the full interconversion structure, since in general the resource theory defines a pre-order $\succ$, such that $\rho \succ \sigma$ iff $\rho \to \sigma$ under symmetric operations [89, 90]. A complete set of measures is known for symmetry-constrained dynamics [73].

Theorem 6.1: Complete set of measures for symmetric dynamics

For a group G, there exists a symmetric operation $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ that takes a state $\rho \in \mathcal{B}(\mathcal{H}_A)$ to $\sigma \in \mathcal{B}(\mathcal{H}_B)$ if

$$H_{\min}(R|A)_{\mathcal{G}[\eta \otimes \rho]} \leq H_{\min}(R|B)_{\mathcal{G}[\eta \otimes \sigma]} \tag{6.3}$$

for all reference frame states $\eta \in \mathcal{B}(\mathcal{H}_R)$, where

$$H_{\min}(R|A)_\Omega := -\log \inf_{\tau_A \geq 0} \{ \mathrm{tr}[\tau_A] \ : \ \mathbb{I}_R \otimes \tau_A \geq \Omega_{RA} \} \tag{6.4}$$

is a single-shot conditional entropy, and

$$\mathcal{G}(\tau) = \int_G dg\ \bar{U}(g) \otimes U(g)\ \tau\ \bar{U}(g)^\dagger \otimes U(g)^\dagger\ , \tag{6.5}$$

with the overbar representing complex conjugation in the chosen basis.

Intuitively, the conditional entropies measure the information that the system encodes about an external reference frame R. Unfortunately, these measures are difficult to use in practice, and they are overcomplete, meaning that a smaller set of measures should fully specify the interconversion structure.

Group orbits offer a handle on this complex interconversion structure. Rather than fully solving the structure, they provide coarse-grained information about the structures involved. From this, we might hope to better understand the complete structure and the measures given in [73].

6.2 Mapping to the Subgroup Lattice

More than simply classifying, our framework brings a natural partial ordering. Just as symmetry constraints define a partial ordering on the state space, we can also define a partial ordering – a lattice, in fact – for subgroups [91, 92].

Definition 6.1: Subgroup lattice

The subgroup lattice of G is a partial order $\prec$ on its subgroups defined by subset inclusion, namely $H_1 \prec H_2$ if $H_1 \subseteq H_2$. In addition, every pair of subgroups have a unique supremum and unique infimum, which define binary operations called the meet and join. For $H_1, H_2 \subseteq G$:

1. their meet $H_1 \wedge H_2$ is $H_1 \cap H_2$,
2. their join $H_1 \vee H_2$ is the subgroup generated by $H_1 \cup H_2$.

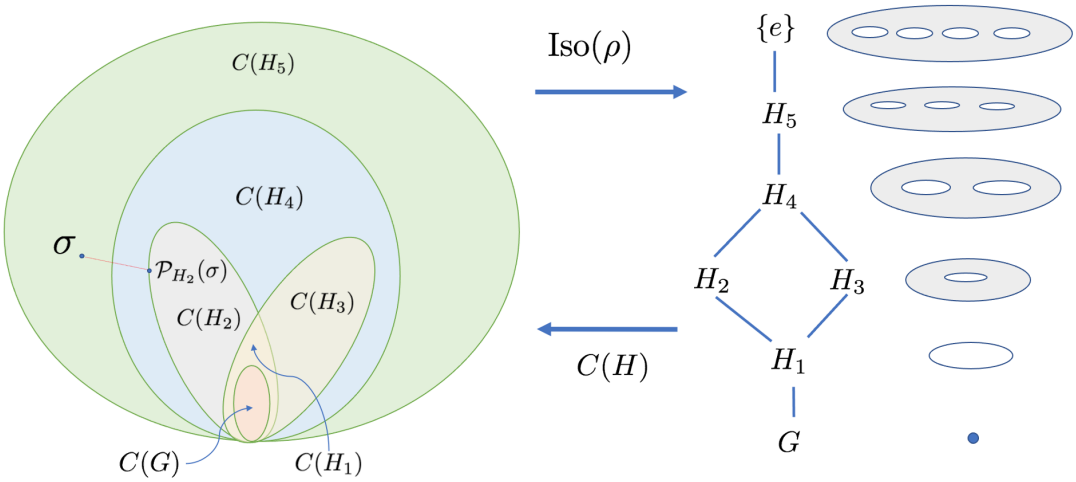


Figure 6.1: Mapping the state space onto the subgroup lattice. *The state space (left) is partitioned into sets $C(H)$ according to isotropy subgroups under the action of a group G. States map onto some subset of the subgroup lattice (middle) through* Iso(ρ)*, which connects to the 'shape' of the group orbit (right). $C(H)$ bridges back from the subgroup lattice onto the state space. In the next section, we introduce the H-twirling map $\mathcal{P}_H$, which projects onto the nearest state ρ with $H \subseteq$* Iso(ρ)*. Based on figure in [1].*

The subgroup lattice can be schematically illustrated by a Hasse diagram of the subgroups, as seen at the end of the previous chapter.

We have already introduced mappings that move between the fine-grained and coarse-grained descriptions of the dynamics. Iso(ρ) maps from the state space to the subgroup lattice, while $C(H)$ goes in the reverse direction, selecting the subset of states with a given isotropy subgroup. These connections are illustrated in Figure 6.1.

With this bridge, the partial ordering of the subgroups lifts onto the equivalence classes $C(H)$, such that $C(H_1) \succ C(H_2)$ whenever $H_1 \prec H_2$. Similarly, the partial ordering of the subgroups exports to the shapes of group orbits.

Definition 6.2: Partial ordering for group orbits

Given the partial ordering for subgroups of a group G, we define a partial ordering for orbits of channels under the action of G, such that

$$\mathcal{M}(\mathcal{E}) \prec \mathcal{M}(\mathcal{F}) \iff \mathrm{Iso}(\mathcal{E}) \succ \mathrm{Iso}(\mathcal{F}) \tag{6.6}$$

up to isomorphism of the isotropy subgroups. Likewise for states, $\mathcal{M}(\rho) \prec \mathcal{M}(\sigma)$ iff $\mathrm{Iso}(\rho) \succ \mathrm{Iso}(\sigma)$ (up to isomorphism).

The ordering is reversed to match the convention for resource measures of asymmetry [87,89], such that a smaller isotropy subgroup maps to a 'bigger' shape. This is because states which maximally break the symmetry are the most useful resources for lifting symmetry constraints and simulating channels in this context.

The partial ordering on subgroups is deeply connected with the interconversion structure of states under symmetric dynamics, and with the resources required to lift these constraints. We can use the subgroup lattice, or alternatively the partial ordering on the group orbit shapes, to describe aspects of dynamics under symmetry constraints.

6.3 H-Twirling

When coarse-graining the state space according to isotropy subgroups, it is helpful to project onto sets of states with a given set of residual symmetries, achieved with the H-twirling map

$$\mathcal{P}_H(\rho) := \int_H dh\, \mathcal{U}_h(\rho) \ , \tag{6.7}$$

averaging over a subgroup H, weighted by the invariant Haar measure dh. If H is a finite subgroup of size $|H|$, the integral $\int_H dh$ is replaced by the sum $\frac{1}{|H|}\sum_{h\in H}$. H-twirling has important applications in various physically motivated scenarios.

Example 6.3: G-twirling

G-twirling $\mathcal{P}_G$ averages over the full group G. This arose in Example 6.1 from having no information about the relative orientation between reference frames.

Example 6.4: Dephasing as H-twirling

The qubit dephasing channel [3] can be expressed as

$$\mathcal{P}_{U(1)}(\rho) = \frac{1}{2\pi}\int_0^{2\pi} d\theta\ e^{i\theta Z}\rho e^{-i\theta Z}, \tag{6.8}$$

sending ρ to the nearest incoherent state while preserving the diagonal terms of the density matrix. This is important in the resource theory of coherence [11].

H-twirling does not necessarily output a state with isotropy subgroup H; it can return states with larger isotropy subgroup. For example, under an SU(2) group action $\mathrm{Iso}\big(\mathcal{P}_{\mathrm{U}(1)}(|+\rangle\langle+|)\big) = \mathrm{Iso}\big(\mathbb{I}/2\big) = \mathrm{SU}(2)$. $\mathcal{P}_H$ projects onto a state with H *contained within* the isotropy subgroup. It is therefore useful to collect together the sets $C(H)$ in the following way.

Definition 6.3: $\hat{C}(H)$

States that are invariant under the subgroup H are collected together in the set

$$\hat{C}(H) := \bigcup_{W \succ H} C(W). \tag{6.9}$$

Note that $\hat{C}(H)$ is a convex set, a simple consequence of Lemma 4.1, because convex combinations have the same residual symmetries (if not more) as the states or channels being mixed. H-twirling therefore projects onto these convex sets.

Lemma 6.1: H-Twirling as a projector

$\mathcal{P}_H$ is an orthogonal projector onto $\hat{C}(H)$.

Proof. Firstly, $H \subseteq \mathrm{Iso}\big(\mathcal{P}_H(\rho)\big)$ because

$$\mathcal{U}_h \circ \mathcal{P}_H(\rho) = \mathcal{U}_h \left(\int_H dh' \, \mathcal{U}_{h'}(\rho) \right) = \int_H dh' \, \mathcal{U}_{hh'}(\rho) = \int_H dh'' \, \mathcal{U}_{h''}(\rho) = \mathcal{P}_H(\rho) \tag{6.10}$$

for all $h \in H$ because of the group rearrangement theorem. Furthermore, $\mathcal{P}_H$ is an orthogonal projector because

$$\mathcal{P}_H \circ \mathcal{P}_H = \int_H dh \, \int_H dh' \, \mathcal{U}_h \circ \mathcal{U}_{h'} = \int_H dh \, \int_H dh' \, \mathcal{U}_{hh'} = \int_H dh'' \, \mathcal{U}_{h''} = \mathcal{P}_H \tag{6.11}$$

and

$$\mathcal{P}_H^\dagger(\rho) = \int_H dh \, \mathcal{U}_{h^{-1}}(\rho) = \int_H dh \, \mathcal{U}_h(\rho) = \mathcal{P}_H(\rho) \ . \tag{6.12}$$

This proves that $\mathcal{P}_H$ is an orthogonal projector onto $\hat{C}(H)$. □

An example of this projection is indicated in Figure 6.1. Moreover, H-twirling outputs the closest state in $\hat{C}(H)$ (as measured by the relative entropy) to the input state [88]. This is useful when smoothing away small perturbations in our classification scheme, as in Section 4.5.

Lemma 6.2: H-twirling projects onto nearest state in $\hat{C}(H)$

$\mathcal{P}_H(\rho) = \arg\min_{\sigma \in \hat{C}(H)} S(\rho||\sigma)$, where $S(\rho||\sigma)$ is the relative entropy.

Proof. By the definition of relative entropy,

$$\min_{\sigma \in \hat{C}(H)} S(\rho||\sigma) = S(\rho) - \min_{\sigma \in \hat{C}(H)} \mathrm{tr}[\rho \log \sigma] \ . \tag{6.13}$$

Since log is an analytic function for density matrices, we can Taylor expand $\log \sigma$ in powers of σ, but because $\mathcal{P}_H(\sigma) = \sigma$, and therefore $\mathcal{P}_H(\sigma^k) = \sigma^k$, we have that $\log \sigma = \mathcal{P}_H(\log \sigma)$. Hence

$$\min_{\sigma \in \hat{C}(H)} S(\rho||\sigma) = S(\rho) - \min_{\sigma \in \hat{C}(H)} \mathrm{tr}[\rho \mathcal{P}_H(\log \sigma)] = S(\rho) - \min_{\sigma \in \hat{C}(H)} \mathrm{tr}[\mathcal{P}_H(\rho) \log \sigma] \tag{6.14}$$

because $\mathcal{P}_H = \mathcal{P}_H^\dagger$, but $\min_{\sigma \in \hat{C}(H)} S(\mathcal{P}_H(\rho)||\sigma) = 0$, obtained when $\sigma = \mathcal{P}_H(\rho)$, hence $\mathcal{P}_H(\rho) = \arg\min_{\sigma \in \hat{C}(H)} S(\rho||\sigma)$. □

H-twirling is used extensively in Appendix B for the classification of group orbits for two qubits under a tensor product SU(2) action.

6.4 A Resource Theory for Shapes?

In a sense, the isotropy subgroup, and hence the associated shape, measure how much a given state breaks the symmetry in question. Unlike other examples of resource measures in the literature, which map the state space onto the totally ordered set $\mathbb{R}$, we instead map onto a partial order (in this case a lattice), thus generalising the concept of a resource measure. Intermediate between the full interconversion structure of the constrained dynamics and a traditional real-valued measure, the shape of the group orbit captures more of the structure than an individual measure.

Can we take this a step further, viewing the subgroup lattice as a coarse-grained resource theory in its own right, thus elevating shape to a resource-theoretic property? The shape certainly encodes information about the full interconversion structure, as the following result makes clear.

Lemma 6.3: Isotropy subgroups are non-decreasing

Given a group action on a system, $\mathcal{M}\big(\mathcal{E}(\rho)\big) \prec \mathcal{M}(\rho)$ for any symmetric operation $\mathcal{E}$.

Proof. Suppose $g \in \mathrm{Iso}(\rho)$, such that $\mathcal{U}_g(\rho) = \rho$. Acting on $\mathcal{E}(\rho)$, the same group transformation gives $\mathcal{U}_g \circ \mathcal{E}(\rho) = \mathcal{E} \circ \mathcal{U}_g(\rho) = \mathcal{E}(\rho)$ because $\mathcal{E}$ is symmetric. Therefore $g \in \mathrm{Iso}(\mathcal{E}(\rho))$, hence $\mathrm{Iso}(\mathcal{E}(\rho)) \succ \mathrm{Iso}(\rho)$. This ordering carries over to shapes, such that $\mathcal{M}(\mathcal{E}(\rho)) \prec \mathcal{M}(\rho)$ for any symmetric operation $\mathcal{E}$. □

Can transformations on the state space be viewed as mappings on the subgroup lattice, specifically as functional mappings between the sets $C(H)$? Unfortunately, there exist states ρ_1 and ρ_2 in the same set $C(H)$ that each get sent to different sets $C(H_1')$ and $C(H_2')$ by the same symmetric channel.

Example 6.5: Failure of functional mapping between $C(H)$

Consider the qubit channel $\mathcal{E}$ such that $\mathcal{E}(\rho) = \mathbb{I}/2$ for all input states ρ. This is an SU(2) symmetric channel, as is $\mathrm{id} \otimes \mathcal{E}$. Although the states $\mathbb{I}/2 \otimes \rho$ and $\rho \otimes \mathbb{I}/2$ both belong to $C\big(\mathrm{U}(1)\big)$ for $\rho \neq \mathbb{I}/2$, they are mapped to states in $C\big(\mathrm{SU}(2)\big)$ and $C\big(\mathrm{U}(1)\big)$ respectively by $\mathcal{E}$.

This highlights that care must be taken when defining a coarse-grained version of the parent resource theory. Instead, we can define functional mappings $\hat{C}(H) \to \hat{C}(f(H))$ because of Lemma 6.3, which ensures that $\hat{C}(H)$ is closed under symmetric operations.

Symmetric operations preserve or add to the residual symmetries, therefore they move down chains of the subgroup lattice. However, there is not necessarily a symmetric operation that reaches all states in $\hat{C}(H)$, in other words $H_1 \prec H_2$ does not imply the existence of a symmetric channel that maps from $\hat{C}(H_1)$ to $\hat{C}(H_2)$. H-twirling does provide one way to move along these chains of subgroups, however it is only a symmetric operation when twirling over a normal subgroup. The following result characterises the residual symmetries of the H-twirling map, providing information on the resources required to implement the channel.

Lemma 6.4: Twirling Resources

Under the action of a group G, $\mathrm{Iso}(\mathcal{P}_{\mathrm{H}}) = \mathrm{N_G(H)}$, where $\mathrm{N}_G(H) := \{g \in G \ : \ gHg^{-1} = H\}$ is the normalizer of H in G. If $H \triangleleft G$, then $\mathcal{P}_H$ is a symmetric operation.

Proof. The group action on $\mathcal{P}_H$ is

$$\mathfrak{U}_g(\mathcal{P}_H) = \int_H dh \ \mathfrak{U}_g(\mathcal{U}_h) = \int_H dh \ \mathcal{U}_g \circ \mathcal{U}_h \circ \mathcal{U}_{g^{-1}} = \int_H dh \ \mathcal{U}_{ghg^{-1}} \ , \tag{6.15}$$

therefore $g \in \mathrm{Iso}(\mathcal{P}_H)$ iff $gHg^{-1} = H$. This implies that $\mathrm{Iso}(\mathcal{P}_{\mathrm{H}}) = \mathrm{N_G(H)}$, and this equals the full group G if $H \triangleleft G$, since the normalizer of a normal subgroup is the full group. □

This tells us the cost of projecting out residual symmetries, providing an upper bound on the type of resource needed to move down the chain of subgroups, although more efficient

transformations are likely to exist. It also brings us naturally to the question of lifting symmetry-induced superselection rules.

6.5 Asymmetry as a Resource

The shape of the group orbit clearly offers insight into the complex structures imposed on the state space by symmetry constraints. Given access to states that break this symmetry, we can lift these restrictive superselection rules and open up new transformations beyond the symmetric operations. But why are asymmetric states useful or important? The following examples provide some intuition, illustrating how asymmetric states can extend our capabilities.

Example 6.6: Signposts

Wayfinding requires asymmetry. Signposts point in a particular direction by breaking rotational symmetry. Clearly a spherical signpost would not work.

Example 6.7: Cams as asymmetric resources

For centuries, clockmakers and engineers have used cams to induce non-trivial patterns of motion from rotation. A cam is an object that is asymmetric under rotation, therefore it pushes a rod or piston a variable amount when turned. These motions would not be possible with purely rotationally-symmetric components.

Example 6.8: Clocks and coherence

Clocks must be asymmetric. The hands on a clockface break rotational symmetry, indicating the time measured by the clock. Analogously, coherence between different energy subspaces evolves non-trivially under unitary time translations generated by the Hamiltonian. This indicates how asymmetry can be a resource for metrology, detecting the action of the relevant symmetry group.

These examples highlight the importance and varied nature of asymmetric resources, and

this carries over into quantum information theory. Such resources act as pointers or reference frames, or as auxiliary systems for implementing useful operations that would otherwise be impossible, called simulating a channel in the resource theory framework. However, these states are often difficult to prepare and limited in number, therefore we should like to know the most efficient way to implement a given channel. Moreover, symmetries can be broken in different ways, as captured by the shape framework, hence there are often different types of resource that permit different sets of additional operations. The shape of the group orbit identifies the type of resource needed in a particular setting.

6.6 Efficient Simulation of Channels

A quantum channel can be implemented by appending an auxiliary resource state and applying a joint evolution on the combined state. We therefore define this to be the simulation of a channel under symmetry constraints.

Definition 6.4: Channel simulation under symmetry constraints

Given the action of a group G on a given quantum system, a channel $\mathcal{E}$ is simulated under symmetry constraints with an auxiliary resource state σ_B if there exists an isometry V such that

$$\mathcal{E}(\rho) = \mathrm{tr}_B\big(V(\rho \otimes \sigma_B)V^\dagger\big) \ , \tag{6.16}$$

where $\mathcal{U}_g(V) = V$ for all $g \in G$.

Our aim is to relate the group orbit shapes of a channel and the resource state needed to simulate it under symmetry constraints. We must therefore understand how the shape behaves under the composition of systems, as described by the following lemma.

Lemma 6.5: Combining channels

Given any quantum channels $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_{A'})$ and $\mathcal{F} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{B'})$, $\mathrm{Iso}(\mathcal{E} \otimes \mathcal{F}) \succ \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$ under the tensor product group action $\mathfrak{U}_g(\mathcal{E} \otimes \mathcal{F}) = \mathfrak{U}_g(\mathcal{E}) \otimes \mathfrak{U}_g(\mathcal{F})$. Similarly, for states $\rho \in \mathcal{B}(\mathcal{H}_A)$ and $\sigma \in \mathcal{B}(\mathcal{H}_B)$, $\mathrm{Iso}(\rho \otimes \sigma) \succ \mathrm{Iso}(\rho) \wedge \mathrm{Iso}(\sigma)$ under a tensor product group action $\mathcal{U}_g(\rho \otimes \sigma) = \mathcal{U}_g(\rho) \otimes \mathcal{U}_g(\sigma)$.

Proof. The tensor product group action gives $\mathfrak{U}_g(\mathcal{E} \otimes \mathcal{F}) = \mathfrak{U}_g(\mathcal{E}) \otimes \mathfrak{U}_g(\mathcal{F}) = \mathcal{E} \otimes \mathcal{F}$ when $g \in \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$, therefore $\mathrm{Iso}(\mathcal{E} \otimes \mathcal{F}) \succ \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$. An analogous proof holds for states. □

The residual symmetries of the combined state or channel must include the meet (intersection) of the respective isotropy subgroups, reflecting the subgroup lattice structure. Since appending an auxiliary state is an example of a quantum channel, this is a special case of the following result.

Lemma 6.6: Concatenating channels

Given two quantum channels $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ and $\mathcal{F} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_C)$, $\mathrm{Iso}(\mathcal{E} \circ \mathcal{F}) \succ \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$ under the action of a group G.

Proof. The group transformation acts on $\mathcal{E} \circ \mathcal{F}$ to give

$$\mathfrak{U}_g(\mathcal{E} \circ \mathcal{F}) = \mathcal{U}_g \circ \mathcal{E} \circ \mathcal{F} \circ \mathcal{U}_{g^{-1}} = \mathcal{U}_g \circ \mathcal{E} \circ \mathcal{U}_{g^{-1}} \circ \mathcal{U}_g \circ \mathcal{F} \circ \mathcal{U}_{g^{-1}} = \mathfrak{U}_g(\mathcal{E}) \circ \mathfrak{U}_g(\mathcal{F}) \ . \tag{6.17}$$

If $g \in \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$, then $g \in \mathrm{Iso}(\mathcal{E} \circ \mathcal{F})$, which implies $\mathrm{Iso}(\mathcal{E} \circ \mathcal{F}) \succ \mathrm{Iso}(\mathcal{E}) \wedge \mathrm{Iso}(\mathcal{F})$. □

Again, concatenation of channels respects the structure of the subgroup lattice. These results connect the shape of the group orbit for a channel to the resource state needed to simulate it, as this next key theorem demonstrates.

Theorem 6.2: Simulating Channels

If a state σ_B can be used to simulate a quantum channel $\mathcal{E}$ under symmetric dynamics $\mathcal{F}$, then $\mathcal{M}(\sigma_B) \succ \mathcal{M}(\mathcal{E})$.

Therefore the resource state σ_B must break the symmetry to at least the same degree as the channel being simulated. This can be seen as an immediate consequence of the previous lemma, since we can view appending the state σ_B as a CPTP map, therefore $\mathcal{E} = \mathcal{F} \circ \sigma_B$, and $\mathrm{Iso}(\mathcal{E}) = \mathrm{Iso}(\mathcal{F} \circ \sigma_\mathrm{B}) \succ \mathrm{Iso}(\mathcal{F}) \wedge \mathrm{Iso}(\sigma_\mathrm{B}) = \mathrm{Iso}(\sigma_\mathrm{B})$. However, we can also prove this with the Stinespring form.

Proof. From the Covariant Stinespring theorem [80, 93],

$$\mathcal{E}(\rho) = \mathcal{F}(\rho \otimes \sigma_B) = \mathrm{tr}_{BC}[V(\rho \otimes \sigma_B \otimes \gamma_C)V^\dagger], \tag{6.18}$$

where $\gamma_C = |\eta\rangle\langle\eta|$ is some pure symmetric state, and V is a globally symmetric unitary such that $[V, U_A(g) \otimes U_B(g) \otimes U_C(g)] = 0$. If $g \in \mathrm{Iso}(\sigma_\mathrm{B})$, then

$$\mathcal{E}(\rho) = \mathrm{tr}_{BC}[V(\rho \otimes \mathcal{U}_g(\sigma_B) \otimes \mathcal{U}_g(\gamma_C))V^\dagger] = \mathrm{tr}_{BC}[V(\mathbb{I} \otimes \mathcal{U}_g \otimes \mathcal{U}_g)(\rho \otimes \sigma_B \otimes \gamma_C)V^\dagger]. \tag{6.19}$$

Because V is symmetric,

$$V(\mathbb{I} \otimes \mathcal{U}_g \otimes \mathcal{U}_g) = V(\mathcal{U}_g \otimes \mathcal{U}_g \otimes \mathcal{U}_g)(\mathcal{U}_g^\dagger \otimes \mathbb{I} \otimes \mathbb{I}) = (\mathcal{U}_g \otimes \mathcal{U}_g \otimes \mathcal{U}_g)V(\mathcal{U}_g^\dagger \otimes \mathbb{I} \otimes \mathbb{I}) \ , \tag{6.20}$$

therefore

$$\mathcal{E}(\rho) = \mathrm{tr}_{BC}[\mathcal{U}_g(V(\mathcal{U}_{g^{-1}}(\rho) \otimes \sigma_B \otimes \gamma_C)V^\dagger)] \tag{6.21}$$

$$= \mathcal{U}_g\big(\mathrm{tr}_{BC}[V(\mathcal{U}_{g^{-1}}(\rho) \otimes \sigma_B \otimes \gamma_C)V^\dagger]\big) = \mathcal{U}_g \circ \mathcal{E} \circ \mathcal{U}_{g^{-1}}(\rho). \tag{6.22}$$

Hence $g \in \mathrm{Iso}(\sigma_\mathrm{B})$ implies $g \in \mathrm{Iso}(\mathcal{E})$, proving that $\mathrm{Iso}(\sigma_B) \prec \mathrm{Iso}(\mathcal{E})$, and $\mathcal{M}(\sigma_B) \succ \mathcal{M}(\mathcal{E})$. □

This places a lower bound on the resources required to simulate a channel under symmetry constraints, however this bound is also tight, in the sense that there exists an efficient simulation from the perspective of shapes.

Theorem 6.3: Efficient simulation of channels

If $\mathcal{E}$ has isotropy group $\mathrm{Iso}(\mathcal{E})$, then there exists a quantum system B and quantum state σ_B with $\mathcal{M}(\sigma_B) \cong \mathcal{M}(\mathcal{E})$ that allows the simulation of $\mathcal{E}$.

Proof. The channel $\mathcal{E}$ decomposes into irreducible process modes [79],

$$\mathcal{E} = \sum_{\lambda,k} \alpha_{\lambda^*,k}(\boldsymbol{x})\Phi_{\lambda,k} \tag{6.23}$$

where λ labels irreducible representations, and k labels the basis vectors of those representations. The coefficients $\alpha_{\lambda,k}(\boldsymbol{x})$ are harmonic wavefunctions on the group orbit $\mathcal{M}(\mathcal{E}) \cong G/\mathrm{Iso}(\mathcal{E})$. The point $\boldsymbol{x}$ encodes the reference data needed to simulate $\mathcal{E}$ under the symmetry constraint [79].

To simulate $\mathcal{E}$, we choose a reference system whose Hilbert space is $\mathcal{L}^2(\mathcal{M}(\mathcal{E}), \mathbb{C})$, explicitly using the group orbit . The required reference data is then encoded in the state $\sigma_B = |\psi\rangle\langle\psi|$, where $\psi(\boldsymbol{z}) = \delta(\boldsymbol{z} - \boldsymbol{x})$. This ensures that the action of the group on the reference system is the same as the action on $G/\mathrm{Iso}(\mathcal{E})$, hence $\mathrm{Iso}(\sigma_B) = \mathrm{Iso}(\mathcal{E})$ and $\mathcal{M}(\sigma_B) \cong \mathcal{M}(\mathcal{E})$. By measuring this reference system and implementing the corresponding channel on the primary system, this simulates $\mathcal{E}$ because

$$\mathcal{E} = \sum_{\lambda,k} \alpha_{\lambda^*,k}(\boldsymbol{x})\ \Phi_{\lambda,k} = \mathrm{tr}_B \sum_{\lambda,k} |\boldsymbol{x}\rangle_B\langle\boldsymbol{x}| \otimes \alpha_{\lambda^*,k}(\boldsymbol{x})\ \Phi_{\lambda,k} \tag{6.24}$$

$$= \sum_{\lambda,k} \int dz\ \mathrm{tr}[M(\boldsymbol{z})\ \sigma_B]\ \alpha_{\lambda^*,k}(\boldsymbol{z})\ \Phi_{\lambda,k} \tag{6.25}$$

where $\{M(\boldsymbol{z})\ dz\}$ is a symmetric position measurement on the group orbit. □

This shows that there exist auxiliary systems in states with the *same* residual symmetries as the given channel that allow the channel's simulation under symmetry constraints. In other words, we can use resources efficiently, not needing to prepare states that break the symmetry to a greater extent than the channel.

Previously, we discussed how the shape framework could be made robust to small perturbations. This robustness carries over to analysing resources for channel simulation. In reality,

our simulation will involve some symmetric noise. Given a channel $\mathcal{E}$, there is a noisy version $\mathcal{E}_\epsilon = (1-\epsilon)\ \mathcal{E} + \epsilon\ \mathcal{D}$, where $\mathcal{D}$ is the complete depolarization map $\mathcal{D}(\rho) = \frac{1}{d}\mathbb{I}$ for all ρ. We can therefore extend our notion of simulation to noisy simulation, where σ_B enables us to simulate $\mathcal{E}_\epsilon$ for some $0 < \epsilon < 1$. The continuity of the Stinespring dilation [94] means that the noisy simulation should be possible using a resource with the same shape (with some error rate ϵ).

6.7 Conclusions & Outlook

In this chapter, we have considered how the shape of group orbits provides a natural coarse-grained description of symmetry-constrained dynamics. Furthermore, we can see the lattice of isotropy subgroups – and the partial ordering induced for the shapes – as a coarse-grained version of the full resource theory, generalising the concept of a resource measure. For resource theories that involve highly complex structures, simpler coarse-grained descriptions provide useful information about resources and how they can be manipulated, without delving into a full analysis. It also points towards looking for coarse-grained versions of a given resource theory, retaining some aspect of the full structure but in a more amenable setting.

We have also demonstrated the importance of group orbits in understanding and analysing resources in this setting, providing a bound for efficient simulation of a channel in terms of group orbit shapes. In practical settings, where quantum resources can be difficult to maintain, such information is very valuable. Moreover, this information is much simpler to compute than the complete set of resource measures found for the resource theory of asymmetry in [73] and, although the group orbit shape gives only coarse-grained information, it could potentially simplify the necessary computations when using those more detailed measures.

Group orbits are obviously a useful tool for understanding the symmetry properties of states and channels. While the shape of the group orbit has important consequences for the efficient implementation of quantum channels, a deeper understanding of the geometric aspects of group orbits is required, perhaps linking to metrology [69] and differential geometry [95]. Future work

might also connect with the group-theoretic approach to coarse-graining dynamics presented in [96].

There is still much work to be done in applying these concepts in more concrete scenarios, whether in the study of multiple non-commuting charges in thermodynamics [86], or in understanding the computational power of quantum computing, for instance in the resource theory of magic [13]. Our framework sits alongside classifications of Clifford gate-sets in a lattice hierarchy [56], and work looking at the symmetry constraints on quantum speed-ups [61]. Hopefully these few chapters have demonstrated the advantages of using the group orbit to classify and manipulate resources under symmetry constraints and that these techniques will find further applications in the ongoing study of resources.

Part III

Resampling Algorithms for Quantum Computers

Chapter 7

Parameter Estimation in Classical & Quantum Bernoulli Factories

This part of the thesis focuses on the advantages that quantum computational resources offer over their classical counterparts. We approach this task by considering a simple, yet practically important, example: the Bernoulli factory. This is a useful tool in statistical modelling, however our interest is motivated by the proven quantum advantage for this algorithm.

- After motivating our interest in Bernoulli factories, we define Bernoulli factories and introduce the quantum advantage for this setting.
- We describe a general construction for Bernoulli factories that puts parameter estimation at its heart, providing intuition through digit functions.
- We demonstrate how the constraints on classical Bernoulli factories arise from limitations in parameter estimation, placing upper bounds on the resources required to implement a given Bernoulli factory.
- Finally, we show how quantum parameter estimation provides a reduction in the required resources, and how the quantum advantage for Bernoulli factories arises from destructive interference.

The main novel contributions of this chapter are:

- The interpretation of the general Bernoulli factory construction in terms of a generalised binary expansion, as described in Sections 7.6 and 7.7.
- Using this construction to place upper bounds on the resources required to simulate a given function in Section 7.8, demonstrating how the constraints on classical Bernoulli factories arise.
- In Section 7.9, applying quantum amplitude estimation to the general Bernoulli factory algorithm to demonstrate not only how the quantum advantage arises for Bernoulli factories, but also the quadratic speed-up allowed in the oracular setting.

The work is from a manuscript in preparation, in collaboration with David Jennings.

7.1 Background and Motivation

Resampling algorithms are amongst the most important and widely used. For example, the Metropolis-Hastings algorithm [97,98] was selected as one of the ten most important algorithms of the twentieth century in a recent review [99]. Resampling algorithms process randomness, using samples from one distribution to simulate those from another that is perhaps difficult to directly simulate. Whilst of interest from a fundamental computer science standpoint, their importance stems from their widespread applications throughout statistical modelling, where one attempts to determine the underlying parameters of a distribution given data [100].

Bernoulli factories [101] are a family of resampling algorithms that have found wide application in statistical modelling, in particular as a key tool for Approximate Bayesian Computation [102]. For many important situations, the likelihood function is either unknown or computationally expensive to evaluate. Various techniques have been proposed to bypass the likelihood function, and Bernoulli factories form part of that armoury, finding application in settings as diverse as disease modelling [103], phylogeography [104] and ocean circulation [105].

In recent years, algorithms for quantum computers have promised computational advantages over their classical counterparts. Two archetypal examples of quantum advantage are Grover's search algorithm [106] and Shor's prime factoring algorithm [107]. Although there is often strong evidence of computational speed-ups in a given setting, the quantum advantage is usually difficult to prove, the need for which is perhaps more pressing in light of recent work reducing the expected quantum speed-ups for various algorithms [108–110].

Bernoulli factories provide a *proven* quantum advantage in the context of randomness processing [111]. This is of theoretical interest, shedding light on how a quantum advantage arises in this simple setting, whilst also providing a good candidate for applying a quantum advantage in real-world computational scenarios, such as statistical modelling. Furthermore, its implementation is amenable to near-term quantum computing platforms, requiring only single-qubit unitary gates and classical post-processing.

Although mathematically rigorous, the proof presented in [111] does not provide an intuitive understanding of the source of this quantum advantage. Such an understanding might point to how other quantum algorithms outcompete their classical counterparts, and whether this advantage can be exploited for applications in statistical modelling.

This chapter examines the role of parameter estimation in determining the resources required by classical and quantum Bernoulli factories. Using a general Bernoulli factory construction first developed by Keane & O'Brien [101], and later Wästlund [112], we demonstrate how parameter estimation constrains the functions that can be simulated in the Bernoulli factory setting. Furthermore, we provide a novel perspective on this construction, decomposing the desired function in terms of 'fuzzy' digit functions. Finally, we show that quantum amplitude estimation provides a speed-up in the number of inputs required, and how destructive interference extends the class of simulable functions beyond the classical restrictions.

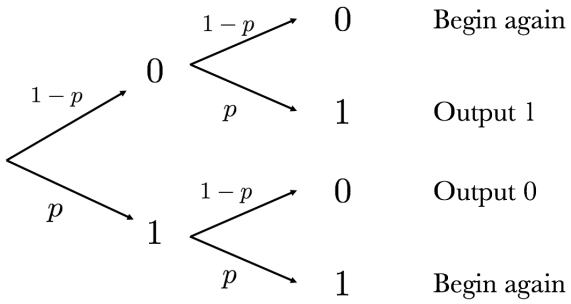


Figure 7.1: The von Neumann coin trick. *The probabilities for each outcome are indicated on the branches of the decision tree.*

7.2 The von Neumann Coin Trick

For a concrete example of a Bernoulli factory, consider an algorithm designed by von Neumann in the context of random number generation [113]. The algorithm simulates unbiased coin flips from independent outcomes of a coin with some unknown fixed bias. This is an example of randomness extraction [114, 115], processing outcomes from a fixed entropy source to distil unbiased random bits.

The input for the algorithm is a (potentially infinite) string of independent outcomes from a Bernoulli random variable $X \sim \text{Bern}(p)$, which produces outcomes 0 and 1 with probability $1-p$ and p respectively. For obvious reasons, this random variable is called a *p-coin*. Crucially, the parameter $p \in (0, 1)$ is not known to us, nonetheless the algorithm uses the string of binary inputs to simulate outcomes from an unbiased coin.

Example 7.1: von Neumann coin trick

Consider the first two binary outcomes (X_1, X_2) in the input string.

1. If $X_1 = X_2$, discard these entries from the input string and begin again.
2. If $X_1 \neq X_2$, then output X_2 and terminate the algorithm.

The decision tree is shown in Figure 7.1.

The algorithm simulates outcomes from an unbiased coin because of the symmetry between outcomes 01 and 10, which both have probability $p(1-p)$. Post-selecting these equally-likely outcomes generates the balanced probabilities of an unbiased coin, without knowledge of the value of p. The algorithm fails when $p = 0$ or 1 because all input bits would be identical and the algorithm would never terminate. Although it is possible that a coin with bias $p \in (0,1)$ produces an infinite string of identical values, such as $(1,1,1,\dots)$, the probability of the algorithm not terminating is of measure zero, ensuring that the expected number of input bits for the algorithm is finite for all $p \in (0,1)$.

Intuitively, if p is close to 0 or 1, pairs of outcomes are more likely to be equal, hence more outcomes are discarded by the algorithm before the output is generated. This is clearly inefficient, however there are more efficient adaptations of this algorithm that use all input bits, thereby attaining the optimal binary entropy bound for the expected number of inputs [116].

7.3 Classical Bernoulli Factories

The von Neumann coin trick illustrates the essential features of Bernoulli factories. They are algorithms that take in a (potentially infinite) string of binary inputs and generate a single binary output. Crucially, the algorithm is expected to terminate after a finite number of input bits, corresponding to a finite-time algorithm, although the actual number of inputs can vary – this is known as a 'Las Vegas' algorithm [117].

A Bernoulli factory generalises the von Neumann algorithm, simulating outcomes from an $f(p)$-coin, where $f : (0,1) \to [0,1]$, such that the probability of the algorithm generating the output 1 is $f(p)$.

Example 7.2: $f(p) = 2p(1-p)$

To simulate outcomes from a $2p(1-p)$-coin, consider two binary input bits (X_1, X_2) from a p-coin. If $X_1 \neq X_2$, output 1, otherwise output 0.

In general, the number of inputs is allowed to vary, and probabilities must be assigned to sets of infinite binary strings. This is achieved with cylinder sets [118].

Definition 7.1: Cylinder sets and probability measures

Given a sample space $\Omega := \{(X_1, X_2, \dots) : X_i \in \{0,1\}\ \forall i \in \mathbb{N}\}$ of infinite binary strings, a probability measure P_p is defined on cylinder sets

$$U(x_1, \dots, x_n) := \{X \in \Omega : X_i = x_i \text{ for } 1 \leq i \leq n\}\ , \tag{7.1}$$

such that

$$P_p(U(\mathbf{x})) = \prod_{i=1}^{n}\ p^{x_i}\ (1-p)^{1-x_i} \tag{7.2}$$

for $p \in (0,1)$.

This specifies a probability measure on all open subsets of Ω because the cylinder sets form a basis for the open sets. The probability measure extends to any subset $S \subseteq \Omega$ by defining an outer measure $P_p(S) := \inf\{P_p(E) : E \text{ is open and } S \subseteq E\}$ [112]. Cylinder sets are sets of infinite strings that begin with the same finite sequence of outcomes, and the measure on them implicitly assumes independent coin flips, whose probabilities multiply. A Bernoulli factory is completely specified by two sets A and B of finite binary strings, corresponding to the outputs 0 and 1 respectively.

Example 7.3: von Neumann coin trick in terms of input sets

The von Neumann algorithm is specified by the sets $A = \{10, 0010, 1110, \dots\}$ and $B = \{01, 0001, 1101, \dots\}$. Output 0 if the input belongs to A, and output 1 if in B.

The algorithm terminates in finite time if it corresponds to an *observable event* [112], illustrated in Figure 7.2.

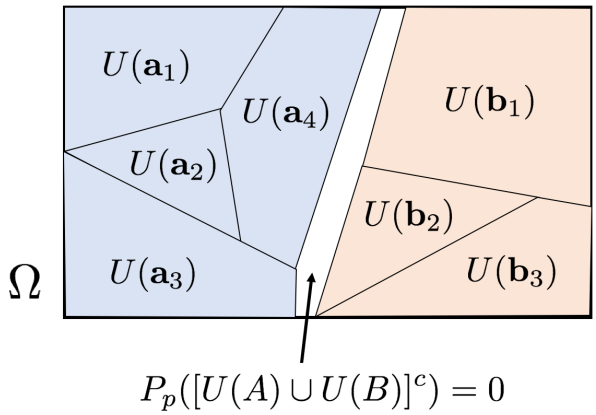


Figure 7.2: An observable event. *The state space Ω has a probability measure imposed on it by cylinder sets. The event occurs if a string in B occurs (the infinite string must lie in $U(B)$, shaded in beige), and does not occur if the string is in $U(A)$ (blue). The probability that the string does not lie in $U(A)$ or $U(B)$ (white) is 0, guaranteeing that the event occurs after a finite number of input bits.*

Definition 7.2: Observable event

An event $E \subseteq \Omega$ is observable if there are two sets $A = \{\mathbf{a}_1, \mathbf{a}_2, \dots\}$ and $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots\}$ of finite binary strings such that

1. $U(\mathbf{a}_i) \cap U(\mathbf{a}_j) = U(\mathbf{b}_k) \cap U(\mathbf{b}_l) = U(\mathbf{a}_m) \cap U(\mathbf{b}_n) = \emptyset$ for all distinct finite binary strings, and
2. $P_p(U(A)) + P_p(U(B)) = 1$, where $U(A) = \bigcup_i U(\mathbf{a}_i)$ and $U(B) = \bigcup_i U(\mathbf{b}_i)$.

The event E occurs if the outcome sequence $X \in \Omega$ belongs to $U(B)$ and does not occur if it belongs to $U(A)$.

A Bernoulli factory corresponds to an observable event E, such that $P_p(E) = f(p)$. Given the simplicity of the Bernoulli factory algorithm for $f(p) = 2p(1-p)$, one might expect a similar scheme for $f(p) = 4p(1-p)$, however the following theorem [101,112] shows this is not the case.

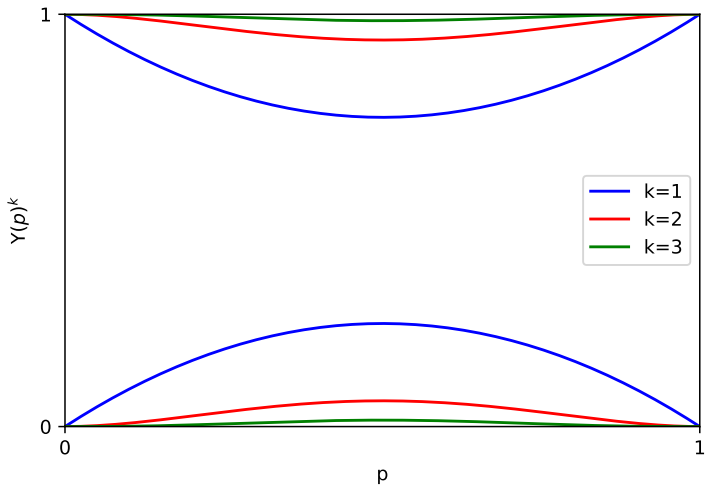


Figure 7.3: Bounds on simulable functions for classical Bernoulli factories. *The upper and lower bounds are given by* $\Upsilon(p)^k$ *and* $1 - \Upsilon(p)^k$ *respectively, with* $k = 1, 2, 3$ *depicted.*

Theorem 7.1: Classical Bernoulli factory constraints

There exists a classical Bernoulli factory for the function $f : (0, 1) \to [0, 1]$ iff

1. f is continuous, and
2. either f is a constant function, or for all $p \in (0, 1)$, $\Upsilon(p)^k < f(p) < 1 - \Upsilon(p)^k$ for some $k \in \mathbb{N}$,

where $\Upsilon(p) := p(1 - p)$.

We say $f(p)$ is *simulable* if there exists a corresponding Bernoulli factory. Intuitively, these conditions specify that a classically simulable function cannot evaluate to 0 or 1 within the domain $(0, 1)$, which implies there is no Bernoulli factory for $f(p) = 4p(1-p)$. Simulable functions are bounded away from 0 and 1 by the fluctuation level $\Upsilon(p)$, as illustrated in Figure 7.3, and more coins will be required the closer $f(p)$ gets to 0 or 1.

7.4 Quantum Bernoulli Factories

Besides their practical importance in statistical modelling, our interest in Bernoulli factories arises from their quantum analogue [111], which makes two key changes from the classical setting.

(Coins to Quoins) The input for a quantum Bernoulli factory is a string of two-dimensional quantum systems (qubits) prepared in the pure state $|p\rangle = \sqrt{1-p}\ |0\rangle + \sqrt{p}\ |1\rangle$, where $p \in (0,1)$ is again unknown. These are called *p-quoins*, and measuring them in the computational basis $\{|0\rangle, |1\rangle\}$ simulates outcomes from a classical p-coin.

(Unitary Operations) Arbitrary single-qubit unitaries can be performed on the input qubits prior to measurement in the computational basis. The measurement outcomes can then undergo classical processing. One could allow multi-qubit gates in this scheme to reduce the number of quoins required, however these are more difficult to implement with comparable fidelities and they do not extend the class of simulable functions beyond those available with single-qubit operations [111].

Using p-quoins allows the simulation of a wider class of functions.

Example 7.4: Quantum advantage from a single quoin

Applying a Hadamard gate to a single p-quoin, followed by measurement in the computational basis, generates outcomes from an $f(p)$-coin, where

$$f(p) = \frac{1}{2} - \sqrt{p(1-p)}. \tag{7.3}$$

This function is not classically simulable because $f(1/2) = 0$.

In [111], the conditions for simulability with a quantum Bernoulli factory were derived.

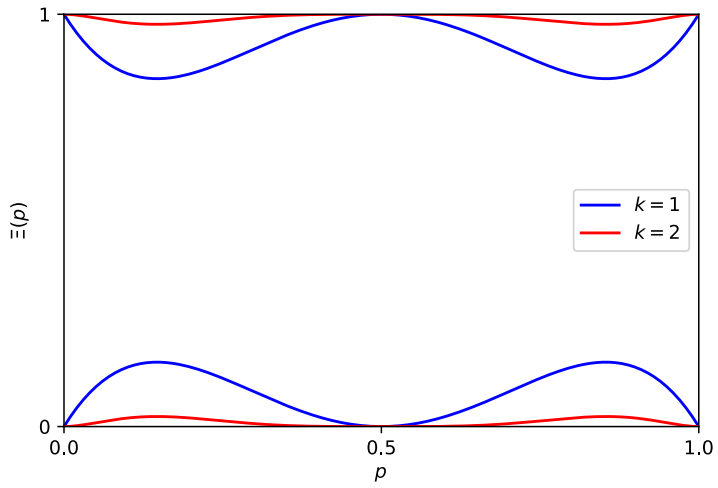


Figure 7.4: Bounds on simulable functions for quantum Bernoulli factories. *The corresponding bounds for functions that are simulable with a quantum Bernoulli factory are shown, namely* $\Xi_Z(p)^k$ *and* $1-\Xi_Z(p)^k$ *for* $k=1,2$*. Here* $Z=\{1/2\}$*.*

Theorem 7.2: Quantum Bernoulli factory constraints

There exists a quantum Bernoulli factory for the function $f:(0,1)\to[0,1]$ that uses single-qubit unitaries and classical post-processing if

1. f is piecewise continuous, and
2. either f is a constant function, or for all $p\in(0,1)$, $\Xi_Z(p)^k \le f(p) \le 1-\Xi_Z(p)^k$ for some $k\in\mathbb{N}$,

where, for a finite set $Z\subset(0,1)$,

$$\Xi_Z(p) := \Upsilon(p)\prod_{z\in Z} h_z(p)\big(1-h_z(p)\big) \tag{7.4}$$

with $h_z(p) := \left(\sqrt{p(1-z)}-\sqrt{z(1-p)}\right)^2$.

These conditions are reminiscent of the classical constraints, but with modified bounding functions $\Xi(p)$ in place of $\Upsilon(p)$, as shown in Figure 7.4. This allows a simulable function to evaluate to 0 or 1 a finite number of times in the domain $(0,1)$, proving the quantum advantage

for Bernoulli factories.

Quantum Bernoulli factories are also amenable to implementation on present and near-term quantum computers. The use of single-qubit unitaries eliminates problems arising from lower-fidelity multi-qubit gates and architectural constraints arising from the arrangement of qubits. A proof-of-principle demonstration has already been performed with a photonic platform [119]. Quantum Bernoulli factories are therefore a prime target for exploiting a quantum advantage on NISQ-era platforms.

Although the mathematical constructions used to prove the simulability conditions for classical and quantum Bernoulli factories are well-founded, they do not provide an intuitive understanding of how these constraints arise. The aim of this chapter is to demonstrate how parameter estimation determines which functions are simulable, and importantly to pinpoint the source of the quantum advantage in this setting.

7.5 Bernoulli Factories for Constant Functions

To build a general Bernoulli factory, we adapt a simple construction for constant functions, based around binary expansions. From the von Neumann scheme, which simulates the constant function $f(p) = \frac{1}{2}$, a $\frac{1}{4}$-coin can be simulated by generating two unbiased coins and outputting 1 only when both give outcome 1. This clearly generalises, using n outcomes from the von Neumann algorithm to simulate a $\frac{1}{2^n}$-coin.

Bernoulli factories for any constant function $f(p) = c$ can then be constructed, where $c \in [0, 1]$. The algorithm utilises the binary expansion of the constant, $c = 0.b_1b_2b_3\ldots$, which encodes the infinite series

$$c = \sum_{n=1}^{\infty} b_n \frac{1}{2^n} \quad \text{where } b_n \in \{0, 1\}. \tag{7.5}$$

An iterative procedure builds up probability bit-by-bit until the desired value is reached. Beginning with the function $f_1(p) := f(p) = c$, the algorithm checks whether $\frac{1}{2}$ is part of the

summation, as encoded in the first binary digit b_1. If $b_1 = 1$, the algorithm contributes $\frac{1}{2}$ to Eq. (7.5) by simulating an unbiased coin, otherwise it outputs 0. The next iteration defines a new function

$$f_2(p) := f_1(p) - \frac{1}{2}b_1 = 0.0b_2b_3\ldots \tag{7.6}$$

that represents the remainder of the summation. The algorithm then decides whether to contribute a probability of $\frac{1}{4}$, as determined by the next binary digit b_2. This iterative procedure continues for all $n \in \mathbb{N}$, iteratively defining functions

$$f_{n+1}(p) := f_n(p) - \frac{1}{2^n}b_n = 0.0\ldots 0b_{n+1}b_{n+2}\ldots \tag{7.7}$$

to account for the remainder of the binary expansion, and letting the binary digits decide whether to contribute $\frac{1}{2^n}$ to the sum.

In this iterative form, it might appear as though all of the potentially infinitely many contributions to the summation must be physically implemented. However, sampling over the iterations with the appropriate probabilities effectively sums all of the contributions. The algorithm to simulate a c-coin is as follows:

1. Simulate n unbiased coins until outcome 0 is obtained.

2. If $b_n = 1$, then output 1, otherwise output 0.

The first step selects the n^{th} binary digit with probability $\frac{1}{2^n}$, while the second checks the binary expansion to decide whether to contribute $\frac{1}{2^n}$ to the infinite sum in Eq. (7.5). This simulates any constant function $f(p) = c \in [0, 1]$ with a finite expected number of coins, even if c is irrational, as the following example demonstrates.

Example 7.5: $f(p) = 1/\pi$

The irrational number $\frac{1}{\pi}$ has the binary expansion $0.01010001\ldots$. The associated Bernoulli factory obtains a series of n unbiased coin flips until outcome 0 is obtained. The output for the algorithm is then b_n, as summarised in Table 7.1.

n	1	2	3	4	5	6	7	8	...
b_n	0	1	0	1	0	0	0	1	...

Table 7.1: Bernoulli factory for $f(p) = \frac{1}{\pi}$. *Here n is the number of unbiased coin flips before 0 is obtained, and b_n is the corresponding binary digit.*

The simulation splits into two components: (a) sampling $n \in \mathbb{N}$, using only unbiased coins, and (b) filtering the contributions to the total probability, controlled by the binary digits b_n. For constant functions, the value of p is not explicitly used, however it comes into play when simulating non-constant functions, as the next section demonstrates.

7.6 Digit Functions

If $f(p)$ is not a constant, the construction of the previous section must be generalised. Once again, the binary expansion takes centre stage, with the function at p evaluating to

$$f(p) = \sum_{n=1}^{\infty} b_n(p) \, \frac{1}{2^n} \quad \text{where } b_n(p) \in \{0, 1\}. \tag{7.8}$$

Crucially, the binary digits $b_n(p)$ are now themselves functions of p, and we refer to them as *digit functions*. As before, an iterative scheme builds up the contributions bit-by-bit, adding $\frac{1}{2^n}$ if required.

Example 7.6: $f(p) = p$

For a given value of p, the probability $f(p) = p$ can be built up from contributions of size $\frac{1}{2^n}$, as described by the binary expansion. This is visualised in Figure 7.5.

This construction can be applied to any function $f : (0, 1) \to [0, 1]$. Beginning with the function $f_1(p) := f(p)$, the first digit function is

$$b_1(p) := \begin{cases} 1 & f_1(p) \geq \frac{1}{2} \\ \\ 0 & f_1(p) < \frac{1}{2} \end{cases}, \tag{7.9}$$

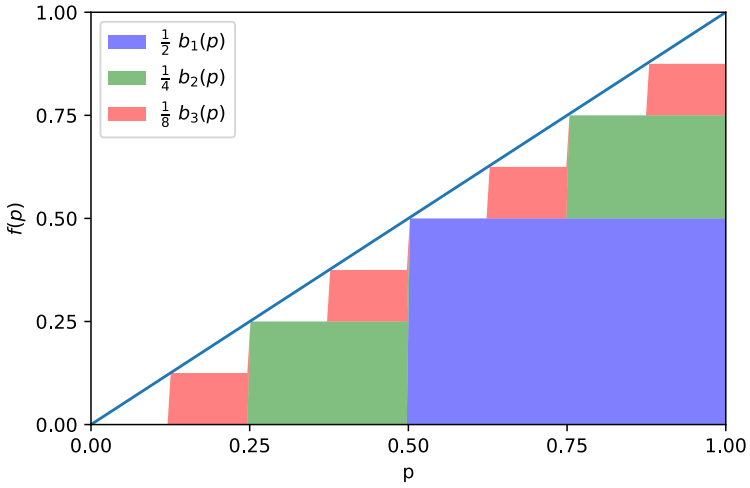


Figure 7.5: Building up $f(p)$ with digit functions. *The function $f(p) = p$ is built up bit by bit from digit functions $b_n(p)$. The first three contributions are shown.*

explicitly the first digit in the binary expansion of $f(p)$ evaluated at p. The next stage is defined by subtracting the contribution of the first iteration, namely $\frac{1}{2}b_1(p)$, leaving

$$f_2(p) := f(p) - \frac{1}{2}b_1(p) = 0.0b_2b_3b_4\ldots(p) \tag{7.10}$$

This can be repeated indefinitely, defined as follows.

Definition 7.3: Digit functions

Given a function $f(p)$, iteratively define

$$f_1(p) := f(p) \qquad f_{n+1}(p) := f_n(p) - \frac{1}{2^n}\, b_n(p)\ , \tag{7.11}$$

where the n^{th} digit function is

$$b_n(p) = \begin{cases} 1 & f_n(p) \geq \frac{1}{2^n} \\ \\ 0 & f_n(p) < \frac{1}{2^n} \end{cases}\ . \tag{7.12}$$

Examples of the iteratively defined functions $f_n(p) = 0.0\ldots0b_nb_{n+1}\ldots(p)$ are shown in Figure 7.6, encoding the contribution from the n^{th} digit onwards in the binary expansion of $f(p)$.

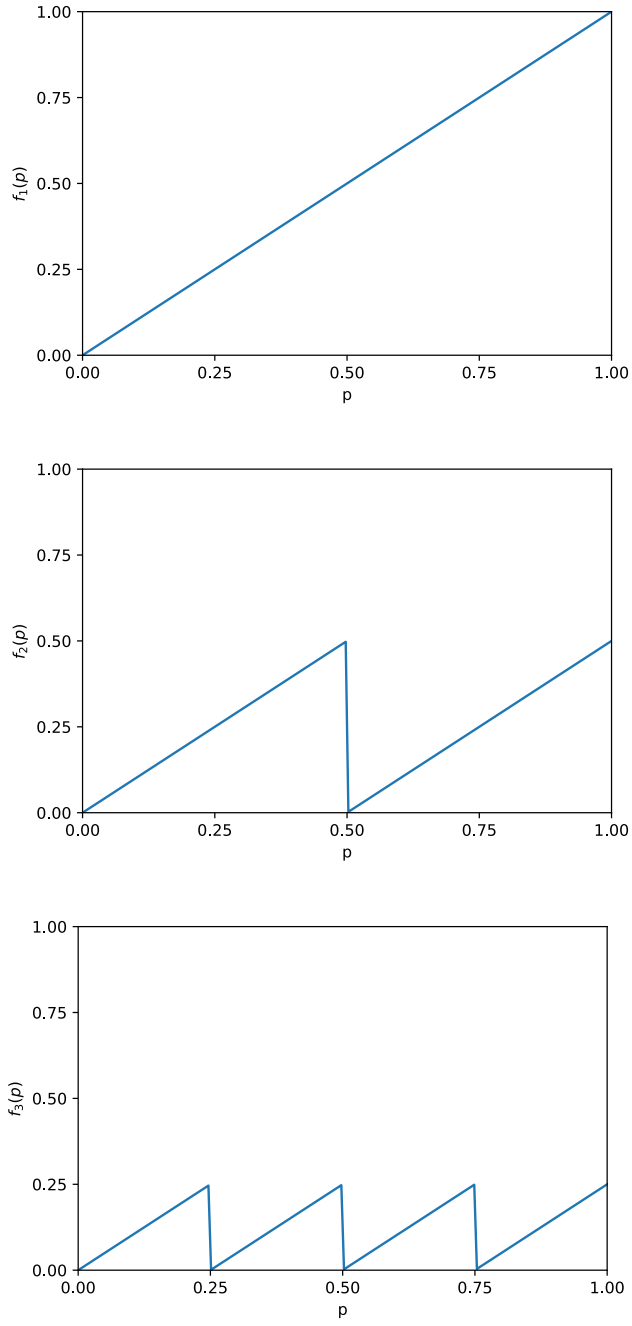


Figure 7.6: $f_n(p)$ encodes the remainder of $f(p)$ to be contributed by digit functions. *For $f(p) = p$, the functions $f_1(p)$, $f_2(p)$ and $f_3(p)$ are pictured from top to bottom, illustrating how the iterative scheme builds up $f(p)$ at each stage. Note that $f_n(p) < \frac{1}{2^{n-1}}$.*

Supposing that the digit functions can be simulated, a Bernoulli factory for $f(p)$ is as follows:

1. Simulate an unbiased coin n times until outcome 0 is obtained.

2. Simulate a $b_n(p)$-coin, taking its outcome as the output for the algorithm.

Once again, the construction of $f(p)$ separates into two components: (a) sampling $n \in \mathbb{N}$ with an unbiased coin, and (b) filtering the contributions with digit functions. However, in this more general case, the value of p now determines the output of the digit functions.

7.7 Fuzzy Digit Functions

Given that digit functions can be defined for any function $f : (0,1) \to [0,1]$, where do the constraints of Theorem 7.1 come from? The problem comes from the digit functions, $b_n(p)$, which are not simulable with a classical Bernoulli factory (unless $b_n(p)$ is a constant) because they evaluate to 0 or 1 within the domain $(0,1)$. To see why this is the case, consider their construction. Since the output depends on p, an obvious algorithm would be:

1. Use parameter estimation to obtain the value of p.

2. If $f_n(p) \geq \frac{1}{2^n}$, then output 1, otherwise output 0.

However, Bernoulli factories use a *finite* number of p-coins, so there will necessarily be some uncertainty in the estimate of $p \in (0,1)$. Simulation of $b_n(p)$ requires infinitely many p-coins.

Nonetheless, $f(p)$ can be expanded in terms of simulable functions that play the role of digit functions. We call these *fuzzy digit functions*, denoted $\hat{b}_n(p)$. The iterative scheme then plays out as before, but with modified functions $\hat{f}_n(p)$. For the first iteration, $\hat{f}_1(p) := f(p)$ and the contribution from the corresponding fuzzy digit function is $\frac{1}{2}\hat{b}_1(p)$. However, $\hat{b}_1(p)$ is constrained by the next iteration,

$$\hat{f}_2(p) := \hat{f}_1(p) - \frac{1}{2}\hat{b}_1(p) \ . \tag{7.13}$$

(Positivity) $\hat{f}_2(p)$ must also be simulable. If $\hat{f}_2(p) < 0$ for any $p \in (0,1)$, negative probabilities would be required. If $\hat{f}_2(p) = 0$ for any $p \in (0,1)$, it must be a constant function. Therefore, $\hat{b}_1(p)$ should be chosen such that $\hat{f}_1(p) > \frac{1}{2}\ \hat{b}_1(p)$ for all $p \in (0,1)$, allowing equality if $\hat{f}_2(p)$ is a constant function.

(Bounded difference) To ensure that $\hat{f}_2(p)$ can be simulated by later fuzzy digit functions,

$$\hat{f}_1(p) - \frac{1}{2}\hat{b}_1(p) < \frac{1}{2}\ , \tag{7.14}$$

such that $\hat{f}_2(p) = 0.0\hat{b}_2\hat{b}_3\ldots(p)$ for all $p \in (0,1)$. Again, this bound can be made an equality if $\hat{f}_2(p)$ is then a constant function.

Definition 7.4: Fuzzy digit functions

Given the function $f(p)$, iteratively define

$$\hat{f}_1(p) := f(p) \qquad \hat{f}_{n+1}(p) := \hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p)\ , \tag{7.15}$$

where the n^{th} fuzzy digit function $\hat{b}_n(p)$ is a simulable function satisfying

$$0 < \hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) < \frac{1}{2^n} \tag{7.16}$$

for all $p \in (0,1)$. These inequalities can be made equalities if $\hat{f}_{n+1}(p)$ is a constant function.

Fuzzy digit functions adapt to the uncertainty in the parameter estimation, shifting the discrepancy from the original digit functions $b_n(p)$ into the following iteration. The algorithm is then implemented much as before:

1. Simulate an unbiased coin n times until outcome 0 is obtained.

2. Simulate an $\hat{b}_n(p)$-coin, and use the outcome as the output for the algorithm.

The fuzzy digit functions play the same role as the binary digits, but have the advantage of being simulable in the context of Bernoulli factories.

7.8 Fuzzy Digit Functions from Parameter Estimation

Given that parameter estimation with an infinite number of p-coins can simulate the digit function $b_n(p)$, it is natural to simulate fuzzy digit functions $\hat{b}_n(p)$ in much the same way, only with a finite number of p-coins. In essence, the aim is to decide whether p belongs to the region

$$I_n := \left\{ p \in (0,1) : \ \hat{f}_n(p) \geq \frac{1}{2^n} \right\} \tag{7.17}$$

with sufficient confidence. The algorithm for simulating $\hat{b}_n(p)$ is as follows:

1. Use parameter estimation to obtain an estimate $\hat{p}$ for the value of p.

2. If $\hat{f}_n(\hat{p}) \geq \frac{1}{2^n}$ (i.e. $\hat{p} \in I_n$), then output 1, otherwise output 0.

If N samples from the p-coin are used, this simulates the function

$$\hat{b}_n(p) = \sum_{j=0}^{N} H_n\big(\hat{f}_n(j/N)\big) \binom{N}{j} p^j (1-p)^{N-j} \ , \tag{7.18}$$

where

$$H_n(x) = \begin{cases} 1 & x \geq \frac{1}{2^n} \\ \\ 0 & x < \frac{1}{2^n} \ . \end{cases} \tag{7.19}$$

This is an example of a Bernstein polynomial, a class of functions that play an important role in constructive proofs of the Weierstrass approximation theorem, which states that any real-valued continuous function on the interval $[a, b]$ can be uniformly approximated by real polynomials on $\mathbb{R}$ [120]. Furthermore, as $N \to \infty$, the fuzzy digit function uniformly converges to $H_n\big(\hat{f}_n(p)\big)$, recovering $b_n(p)$ in the limit, illustrated in Figure 7.7.

The defining constraints on the fuzzy digit functions place requirements on the number of p-coins that must be sampled in the parameter estimation step. These conditions provide the source of the constraints in Theorem 7.1, as we now demonstrate.

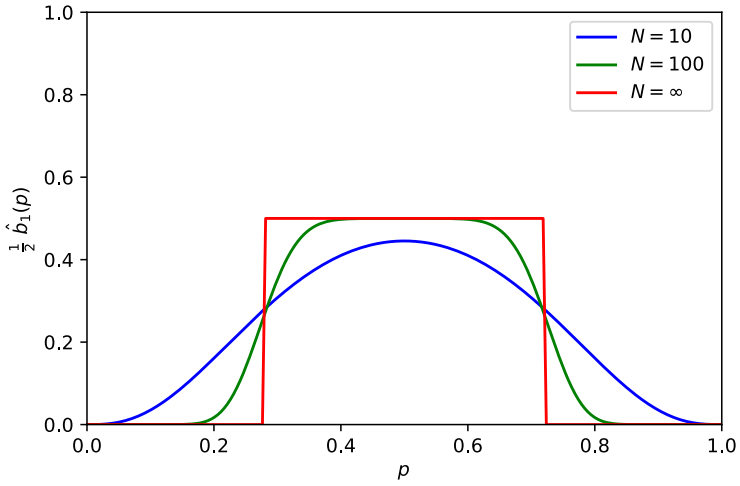


Figure 7.7: Convergence of $\frac{1}{2}\hat{b}_1(p)$ to $\frac{1}{2}b_1(p)$. *As the number N of p-coins used in the parameter estimation step increases, the closer the fuzzy digit function gets to its exact counterpart. Here $f(p) = \frac{5}{2}p(1-p)$.*

7.8.1 Bounding $f(p)$ away from 0

Parameter estimation forces $f(p)$ to be bounded away from 0 within its domain. To see why, suppose that $f(p) = 0$ for some $p \in (0, 1)$. It is not possible to lower bound this function with a non-constant fuzzy digit function simulated with a finite number of p-coins. $\hat{b}_n(p)$ is non-zero for all $p \in (0, 1)$, because there is always some finite probability of the estimate $\hat{p}$ belonging to I_n, implying that $\hat{b}_n(p)$ is non-zero outside of I_n.

However, using enough p-coins in for the parameter estimation step ensures that $\hat{b}_n(p)$ satisfies the requirements in Definition 7.4, such that

$$0 < \hat{f}_n(p) - \frac{1}{2^n}\,\hat{b}_n(p) < \frac{1}{2^n} \tag{7.20}$$

for all $p \in (0, 1)$. As $f(p)$ gets closer to zero, more coins are required to ensure $\frac{1}{2}\hat{b}_1(p)$ lower bounds $f(p)$. The following lemma quantifies the number of p-coins needed to guarantee the fuzzy digit function satisfies the defining conditions outside of I_n.

Lemma 7.1: Parameter estimation bounds outside I_n

Consider a function $\hat{f}_n(p) > \Upsilon(p)^k$. For $p \in (0,1)$ such that the distance to the set I_n,

$$\Delta := \min_{q \in I_n} |p - q| \ , \tag{7.21}$$

is greater than some value $\delta > 0$, then a non-constant fuzzy digit function $\hat{b}_n(p)$ satisfies

$$0 < \hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) < \frac{1}{2^n} \tag{7.22}$$

if the number of coins used in the parameter estimation step is $k\left\lceil 2^{-n/k}\delta^{-2}\right\rceil$.

Proof. For the p being considered, $\hat{f}_n(p) < \frac{1}{2^n}$, therefore $\hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) < \frac{1}{2^n}$. To show that $\hat{f}_n(p) > \frac{1}{2^n}\ \hat{b}_n(p)$, we use Chebyshev's inequality [121], which states that

$$P\big(|\hat{p} - p| > \delta\big) < \frac{\Upsilon(p)}{N\delta^2} \ . \tag{7.23}$$

Using k rounds of this parameter estimation allows us to upper bound the fuzzy digit function, as seen in Figure 7.8, such that

$$\hat{b}_n(p) < P\big(|\hat{p} - p| > \delta\big)^k < \frac{\Upsilon(p)^k}{(N\delta^2)^k} \ , \tag{7.24}$$

only outputting 1 if $\hat{p} \in I_n$ for all k rounds of parameter estimation. Therefore $\hat{f}_n(p) > \frac{1}{2^n}\ \hat{b}_n(p)$ if

$$\frac{1}{2^n(N\delta^2)^k} \leq 1 \ , \quad \text{or equivalently} \quad N \geq \frac{1}{2^{n/k}\delta^2} \ . \tag{7.25}$$

Over the k rounds of parameter estimation, this uses $k\left\lceil 2^{-n/k}\delta^{-2}\right\rceil$ samples, with the ceiling function ensuring an integer number of samples. □

This construction is unlikely to be optimal, however it is guaranteed to satisfy the required bounds. The $\Upsilon(p)^k$ bounds in Theorem 7.1 can be naturally interpreted as fluctuations in the parameter estimation step. Moreover, as the function $f(p)$ gets closer to zero, the value of k gets larger and thus more coins are required.

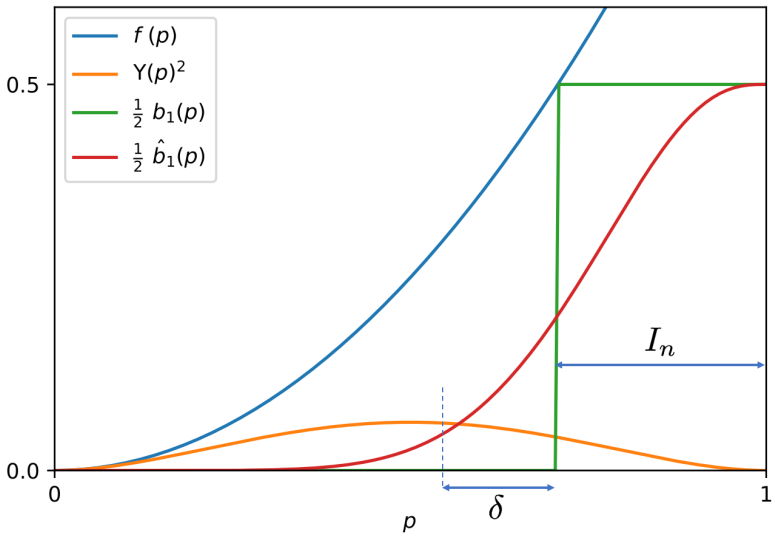


Figure 7.8: Lower bounding $\hat{f}_1(p)$ with $\frac{1}{2}\,\hat{b}_1(p)$. *Here $f(p) = p^2$, which is lower bounded by $\Upsilon(p)^2$. To ensure that $\hat{f}_1(p) > \frac{1}{2}\,\hat{b}_1(p)$, enough coins are used in the parameter estimation to ensure that, for some distance δ from I_n, $P(|\hat{p} - p| > \delta)$ is upper bounded by $\Upsilon(p)$. This is done twice to upper bound $\frac{1}{2}\,\hat{b}_1(p)$ by $\Upsilon(p)^2$.*

7.8.2 Bounding $f(p)$ away from 1

Similarly, the fuzzy digit function construction demonstrates why classically simulable functions cannot evaluate to 1 within the domain $(0, 1)$. Suppose that $f(p_1) = 1$ for some $p_1 \in (0, 1)$, then the constraint that

$$\hat{f}_1(p_1) - \frac{1}{2}\,\hat{b}_1(p_1) < \frac{1}{2} \tag{7.26}$$

cannot be satisfied (although $\hat{f}_1(p) - \frac{1}{2}\,\hat{b}_1(p) = \frac{1}{2}$ if $\hat{b}_1(p)$ is a constant function). With a finite number of coins, there is some finite probability that $\hat{p} \notin I_1$ when $p \in I_1$. Once again, parameter estimation constrains the set of simulable functions. The number of coins needed to simulate the required fuzzy digit functions can be derived in a similar manner to Lemma 7.1.

Lemma 7.2: Parameter estimation bounds inside I_n

Consider a function $\hat{f}_n(p) < \frac{1}{2^{n-1}} - \Upsilon(p)^k$, where $k \in \mathbb{N}$. For $p \in (0,1)$ such that the distance to the set $I_n^{\mathrm{C}} = (0,1) - I_n$,

$$\Delta := \min_{q \in I_n^{\mathrm{C}}} |p - q| \ , \tag{7.27}$$

is greater than some value $\delta > 0$, a non-constant fuzzy digit function $\hat{b}_n(p)$ satisfies

$$0 < \hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) < \frac{1}{2^n} \tag{7.28}$$

if the number of coins used in the parameter estimation step is $k \left\lceil 2^{-n/k} \delta^{-2} \right\rceil$.

Proof. For the p being considered, $\hat{f}_n(p) > \frac{1}{2^n}$, therefore $\hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) > 0$. To ensure that $\hat{f}_n(p) - \frac{1}{2^n}\ \hat{b}_n(p) < \frac{1}{2^n}$, we use Chebyshev's inequality to lower bound $\hat{b}_n(p)$ for the values of p being considered, such that

$$P(\hat{p} \notin I_n) < \frac{\Upsilon(p)}{N\delta^2}. \tag{7.29}$$

Repeating this k times, and only outputting 1 if all k runs give $\hat{p} \in I_n$, ensures that $\hat{b}_n(p) > 1 - 2^n \Upsilon(p)^k$, and hence

$$\hat{f}_n(p) - \frac{1}{2^n}\hat{b}_n(p) < \frac{1}{2^{n-1}} - \Upsilon(p)^k - \frac{1}{2^n} + \Upsilon(p)^k = \frac{1}{2^n} \tag{7.30}$$

if each round of parameter estimation uses at least $2^{-n/k}\delta^{-2}$ samples from the p-coin. This is pictured in Figure 7.9. In total, this scheme uses $k \left\lceil 2^{-n/k} \delta^{-2} \right\rceil$ samples. □

7.8.3 Continuity of $f(p)$

The final condition on simulable functions in Theorem 7.1 is continuity, which also stems from using parameter estimation with only a finite number of samples. This condition arises because $\hat{b}_n(p)$ drops off at the boundary of I_n over some finite region, rather than the immediate cut-off provided by $b_n(p)$. For functions $\hat{f}_n(p)$ that vary rapidly, the parameter estimation step must be more accurate to ensure that the required bounds on $\hat{b}_n(p)$ are satisfied.

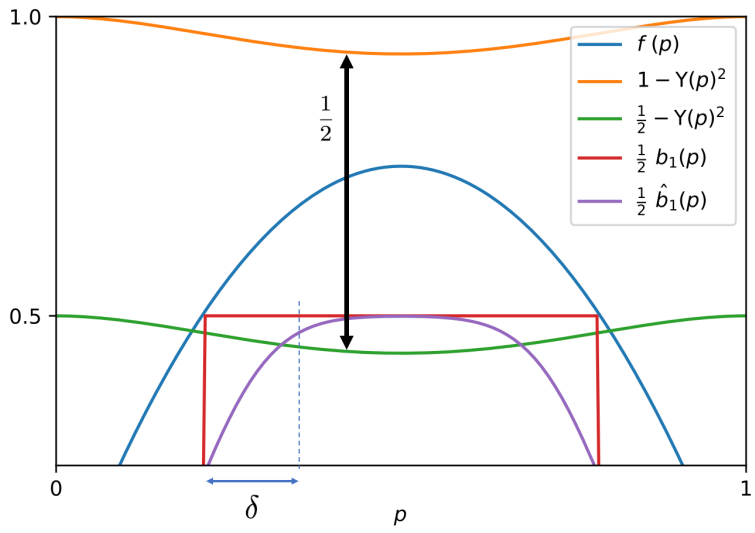


Figure 7.9: Bounding $\frac{1}{2}\hat{b}_1(p)$ such that $\hat{f}_1(p)-\frac{1}{2}\hat{b}_1(p)<\frac{1}{2}$. *Here $f(p)=3p(1-p)$, which is upper bounded by $1-\Upsilon(p)^2$. To guarantee that $\hat{f}_1(p)-\hat{b}_1(p)/2<1/2$, enough coins are used in the parameter estimation to ensure that $\hat{b}_1(p)>1-2\Upsilon(p)^k$ for the values of p in question. This is easily seen because the gap between $f(p)$ and $\frac{1}{2}\hat{b}_1(p)$ is clearly smaller than the gap between $1-\Upsilon(p)^2$ and $1/2-\Upsilon(p)^2$.*

Consider values of $p\notin I_n$ within δ of the boundary of I_n. Clearly $\frac{1}{2^n}\hat{b}_n(p)<\frac{1}{2^n}$, and we have already specified that $\frac{1}{2^n}\hat{b}_n(p)<\Upsilon(p)^k$ at a distance δ from the boundary. As illustrated in Figure 7.10, choosing δ such that the gradient of the line joining these two bounds is steeper than the maximum gradient of $\hat{f}_n(p)$ ensures that the straight line lower bounds $\hat{f}_n(p)$ in the region around the boundary of I_n. Furthermore, $\hat{b}_n(p)$ dies away outside I_n, therefore $\frac{1}{2^n}\hat{b}_n(p)$ must remain below this same straight line if it is also below it at the endpoints of this δ-sized region. Therefore the required bounds on $\hat{b}_n(p)$ are satisfied if

$$\left|\frac{\frac{1}{2^n}-\Upsilon(p)^k}{\delta}\right|>\max_{p\in(0,1)}\left|\frac{d\hat{f}_n}{dp}\right|. \tag{7.31}$$

For simplicity, we choose

$$\delta=\frac{1}{2^n}\left(\max_{p\in(0,1)}\left|\frac{d\hat{f}_n}{dp}\right|\right)^{-1}. \tag{7.32}$$

A similar reasoning holds for δ-sized region around the boundary inside I_n. This places an upper bound on the number of coins required to simulate the required fuzzy digit function.

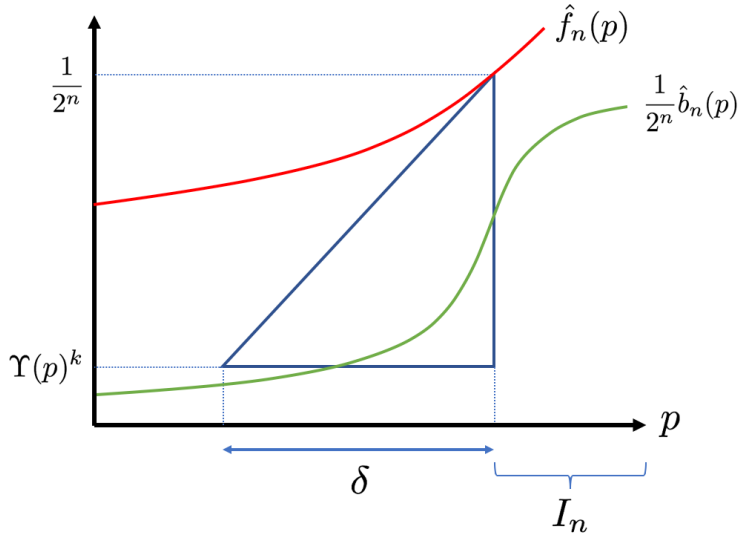


Figure 7.10: Bounding $\frac{1}{2}\hat{b}_1(p)$ near the boundary of I_n. *To ensure that $\hat{f}_n(p) > \frac{1}{2}\hat{b}_1(p)$ in the region around the boundary of I_n, choose δ such that the absolute value of the gradient of the triangle in this figure is greater than $\max_{p\in(0,1)} \left|\frac{d\hat{f}_n}{dp}\right|$.*

Theorem 7.3: Classical fuzzy digit function resources

Given a differentiable function $\hat{f}_n(p)$, bounded such that $\Upsilon(p)^k < \hat{f}_n(p) < \frac{1}{2^{n-1}} - \Upsilon(p)^k$, then a non-constant fuzzy digit function $\hat{b}_n(p)$, satisfying $0 < \hat{f}_n(p) - \frac{1}{2^n}\,\hat{b}_n(p) < \frac{1}{2^n}$ for all $p \in (0,1)$, can be simulated with

$$N = k\left\lceil \frac{1}{2^{n/k}\delta^2} \right\rceil = k\left\lceil 2^{\left(2n-\frac{n}{k}\right)} \max_{p\in(0,1)} \left|\frac{d\hat{f}_n}{dp}\right|^2 \right\rceil \tag{7.33}$$

samples from the p-coin.

Functions with larger derivatives require more coins in the parameter estimation step. If $\hat{f}_n(p)$ has a discontinuity, (i.e. $\hat{f}_n' \to \infty$), the number of coins required $N \to \infty$. It is likely that the conditions for the fuzzy digit function are satisfied with fewer samples from the p-coin, and further work is required to find better bounds on the number of coins required. The coming sections describe how the δ^{-2} dependence in N can be improved to δ^{-1} in the quantum setting, and how the proven quantum advantage for Bernoulli factories arises.

7.9 Quantum Speed-up for Bernoulli Factories

The fuzzy digit function construction highlights the central role of parameter estimation in restricting the randomness processing power of the classical Bernoulli factory. Is it advantageous to replace classical parameter estimation by its quantum analogue in this construction? We now show that it offers a square-root speed-up in the number of p-coins required to simulate a given function, much like Grover's algorithm [106].

Amplitude estimation is a quantum mechanical parameter estimation scheme [122], estimating the value of p encoded in the p-quoins. However, it requires a slightly stronger form of the quantum Bernoulli factory scenario, namely that rather than pre-prepared p-quoins, we instead access a unitary oracle O_p that prepares p-quoins. For simplicity, we consider a p-quoin oracle of the form $O_p = e^{-i\theta Y}$, where $\sin\theta = \sqrt{p}$, thus ensuring that $O_p |0\rangle = |p\rangle$. This scenario was first considered by Yoder [123].

7.9.1 Quantum Amplitude Estimation

The amplitude estimation algorithm obtains an estimate for p by performing quantum phase estimation [3, 107] on the p-quoin oracle. The algorithm requires two registers:

- a register A of $s = \log N$ qubits in which to output the estimate,
- a register B composed of a single qubit on which the quoin oracle O_p can act.

The registers are initialised in the state $|0\rangle_A |0\rangle_B$, then a Fourier Transform unitary, which acts on the computational basis as

$$Q |j\rangle_A = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{2\pi i j k}{N}} |k\rangle_A , \tag{7.34}$$

is applied to the first register to prepare the state

$$Q_A |0\rangle_A |0\rangle_B = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle_A |0\rangle_B . \tag{7.35}$$

The key to the phase estimation algorithm, and the reason for switching to the oracular setting, is a coherent querying of the oracle, $\Lambda(O_p)\,|j\rangle\,|\psi\rangle = |j\rangle\left(O_p^j\ |\psi\rangle\right)$, where $|\psi\rangle$ is an arbitrary state. Applied to our state,

$$\Lambda(O_p)\ Q_A\,|0\rangle_A\,|0\rangle_B = \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle\left(e^{-ij\theta Y}\,|0\rangle\right) = \frac{1}{\sqrt{2N}}\sum_{j=0}^{N-1}e^{-ij\theta}\,|j\rangle\,|+i\rangle + e^{ij\theta}\,|j\rangle\,|-i\rangle\,, \tag{7.36}$$

where $|\pm i\rangle$ are the ± 1 eigenvectors of Y. The inverse fourier transform is then applied to the first register to leave the state

$$Q_A^\dagger\ \Lambda(O_p)\ Q_A\,|0\rangle_A\,|0\rangle_B = \frac{1}{\sqrt{2}N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}e^{-ij\left(\theta+\frac{2\pi k}{N}\right)}\,|k\rangle_A\,|+i\rangle_B + e^{ij\left(\theta-\frac{2\pi k}{N}\right)}\,|k\rangle_A\,|-i\rangle_B\,. \tag{7.37}$$

The aim is to estimate $p=\sin^2\theta$, so there is no difference between the choices $\theta=\pm\frac{2\pi k}{N}$. Considering just one of these possibilities, we consider the above state to be

$$|\phi\rangle_{AB} = \frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}e^{ij\left(\theta-\frac{2\pi k}{N}\right)}\,|k\rangle_A\,|+i\rangle_B \tag{7.38}$$

for simplicity. The probability of measuring the value k in the first register is

$$P(k) = \frac{1}{N^2}\left|\sum_{j=0}^{N-1}e^{ij\left(\theta-\frac{2\pi k}{N}\right)}\right|^2 = \frac{\sin^2\left(\frac{N}{2}\left[\theta-\frac{2\pi k}{N}\right]\right)}{N^2\sin^2\left(\frac{1}{2}\left[\theta-\frac{2\pi k}{N}\right]\right)}\,, \tag{7.39}$$

which provides an estimate $\hat{p}$ for p through the relation $\hat{p}=\sin^2\left(\frac{2\pi k}{N}\right)$.

Coherent querying of the oracle permits a reduction in the number of 'coins' required for the parameter estimation step in the simulation of fuzzy digit functions. Moreover, just like the classical parameter estimation scheme used in earlier sections, when $p=0$ or 1, then $P(\hat{p}=p)=1$. The certainty in the estimate at these special values of p is the source of the quantum advantage, since these special points can be changed by application of a unitary before the amplitude estimation step.

7.9.2 Amplitude Estimation and Bernoulli Factories

In this oracular setting, the resources required to simulate a function are quantified by the number of calls to the p-quoin oracle, analogous to the number of p-coins required in the

classical setting. The fuzzy digit construction can be used in the quantum case, utilising the amplitude estimation algorithm to estimate p. As before, enough oracle calls are required to bound the fuzzy digit functions, and this is provided by a quantum analogue of the Chebyshev inequality.

Lemma 7.3: Quantum Chebyshev inequality from amplitude estimation

Given access to a p-quoin oracle O_p, quantum amplitude estimation provides an estimate $\hat{p}$ for the parameter p such that for $\delta > 0$,

$$P\big(|\hat{p} - p| > \delta\big) < \frac{\pi\sqrt{\Upsilon(p)}}{N\delta} + \mathcal{O}\left(\frac{\pi^2\Upsilon(p)}{N^2\delta^2}\right), \tag{7.40}$$

where N is the number of calls to the oracle, and $\Upsilon(p) = p(1-p)$. For $p = 0$ or 1, $P(\hat{p} = p) = 1$.

Proof. It was shown in [122] that for any integer $j > 0$, the amplitude estimation algorithm generates an estimate $\hat{p}$ for the parameter p such that

$$P\left(|\hat{p} - p| > \frac{2\pi j\sqrt{\Upsilon(p)}}{N} + \frac{j^2\pi^2}{N^2}\right) \leq \begin{cases} 1 - \frac{8}{\pi^2} & j = 1 \\ \frac{1}{2(j-1)} & j \geq 2, \end{cases} \tag{7.41}$$

with $P(\hat{p} = p) = 1$ for $p = 0$ or 1. To obtain a similar form to the Chebyshev bound, namely $P\left(|\hat{p} - p| > \delta\right) < \epsilon$, we set

$$\delta = \frac{2\pi j\sqrt{\Upsilon(p)}}{N} + \frac{j^2\pi^2}{N^2}, \tag{7.42}$$

which can be rearranged into standard quadratic form

$$\left(\frac{j\pi}{N}\right)^2 + 2\sqrt{\Upsilon(p)}\left(\frac{j\pi}{N}\right) - \delta = 0 \tag{7.43}$$

with a positive solution

$$j = \frac{N}{\pi}\left(\sqrt{\Upsilon(p) + \delta} - \sqrt{\Upsilon(p)}\right) = \frac{N\delta}{2\pi\sqrt{\Upsilon(p)}} + \mathcal{O}\left(\frac{\delta^2}{\Upsilon(p)^{\frac{3}{2}}}\right). \tag{7.44}$$

Assuming that $j \geq 2$ and $\epsilon = (2j-1)^{-1}$,

$$P\left(|\hat{p} - p| > \delta\right) < \frac{1}{2\left(j-1\right)} = \frac{1}{2j} + \frac{1}{2j(j-1)} \tag{7.45}$$

$$= \frac{\pi\sqrt{\Upsilon(p)}}{N\delta} + \mathcal{O}\left(\frac{\pi^2\Upsilon(p)}{N^2\delta^2}\right) . \tag{7.46}$$

The higher order terms can be neglected with $\frac{\pi}{N\delta} < 1$. □

There are two important differences to the classical Chebyshev bound. Firstly $\sqrt{\Upsilon(p)}$ has replaced $\Upsilon(p)$. More importantly, the probability scales with δ^{-1} instead of δ^{-2}, and this is the source of a quantum speed-up over the classical fuzzy-digit construction in the following result.

Theorem 7.4: Quantum fuzzy digit function resources

Given a differentiable function $\hat{f}_n(p)$, bounded such that $\Upsilon(p)^k < \hat{f}_n(p) < \frac{1}{2^{n-1}} - \Upsilon(p)^k$, then a non-constant fuzzy digit function $\hat{b}_n(p)$, satisfying $0 < \hat{f}_n(p) - \frac{1}{2^n}\,\hat{b}_n(p) < \frac{1}{2^n}$ for all $p \in (0,1)$, can be simulated with

$$N = 2k\left\lceil 2^{\left(n-\frac{n}{2k}\right)}\pi \max_{p\in(0,1)}\left|\frac{d\hat{f}_n}{dp}\right|\right\rceil \tag{7.47}$$

calls to the p-quoin oracle.

Proof. We use the same construction as the classical case, simply replacing the classical estimate $\hat{p}$ with that obtained from quantum amplitude estimation, and using the Chebyshev-like bound derived in Lemma 7.3. To guarantee the bound of $\Upsilon(p)^k$, the amplitude estimation algorithm is run $2k$ times, only generating outcome 1 if all $2k$ parameter estimation runs give $\hat{p} \in I_n$. The fuzzy digit conditions are satisfied if

$$\frac{1}{2^n}\frac{\pi^{2k}}{(N\delta)^{2k}} \leq 1 \tag{7.48}$$

Given $2k$ rounds of estimation, and the same choice of δ as for the classical construction, this uses

$$N = 2k\left\lceil\frac{\pi}{2^{n/2k}\delta}\right\rceil = 2k\left\lceil 2^{\left(n-\frac{n}{2k}\right)}\pi \max_{p\in(0,1)}\left|\frac{d\hat{f}_n}{dp}\right|\right\rceil \tag{7.49}$$

samples from the p-quoin oracle, the ceiling function arising from the requirement that an integer number of samples is required on each round of parameter estimation. □

This explains the square-root speed-up for the quantum Bernoulli factory previously noted in [119]. Although the class of simulable functions is not extended here by the use of quantum parameter estimation, it is able to simulate the same functions using fewer coins, reducing the resources needed for the simulation. Even if a particular function is not simulable, the quantum Bernoulli factory simulates a better approximation with the same number of inputs.

7.10 Destructive Interference and the Quantum Advantage

Our interest in Bernoulli factories was motivated by the proven quantum advantage in this setting. Quantum parameter estimation provides a speed-up for the fuzzy digit function construction, however we have yet to show how the use of p-quoins extends the class of simulable functions. The source of this quantum advantage is destructive interference, specifically the ability to unitarily transform a particular p-quoin, whose measurement outcomes have non-zero entropy in the computational basis, to $|0\rangle$ or $|1\rangle$, where the outcomes are totally predictable.

Lemma 7.4: 'Derandomizing' a p-quoin

The unitary $e^{-i\theta Y}$ acts on a p-quoin to produce a new quoin with parameter $q = \sin^2\left(\arcsin\left(\sqrt{p}\right) + \theta\right)$. When $\theta = -\arcsin\left(\sqrt{p}\right)$, this outputs the state $|0\rangle$.

Proof. Setting $\sqrt{p} = \sin\alpha$, the unitary acts on a p-quoin to give

$$e^{-i\theta Y}|p\rangle = (\cos\theta\ \mathbb{I} - i\sin\theta\ Y)(\cos\alpha|0\rangle + \sin\alpha|1\rangle) \tag{7.50}$$

$$= (\cos\alpha\cos\theta - \sin\alpha\sin\theta)|0\rangle + (\sin\alpha\cos\theta + \cos\alpha\sin\theta)|1\rangle \tag{7.51}$$

$$= \cos(\alpha+\theta)|0\rangle + \sin(\alpha+\theta)|1\rangle \tag{7.52}$$

This is a quoin with parameter $q = \sin^2(\alpha+\theta) = \sin^2\left(\arcsin\left(\sqrt{p}\right) + \theta\right)$. Setting $\theta = -\alpha$ gives $e^{-i\theta Y}|p\rangle = |0\rangle$. □

The cases where $p = 0$ or 1 are special from the point of view of parameter estimation. For these values, the estimate $\hat{p}$ always equals p, and there is no chance of error when deciding whether $p \in I_n$. This explains why the bounding functions $\Upsilon(p)$ tend to zero at these points. Applying a unitary prior to parameter estimation moves these special points within the domain.

Lemma 7.5: 'Shifted' parameter estimation

After transforming p-quoin samples with the unitary $e^{-i\theta Y}$, classical parameter estimation can decide whether $p \in I_n$, with error probability upper bounded by $\Upsilon(q)$, where $q = \sin^2\left(\arcsin\left(\sqrt{p}\right) + \theta\right)$.

Proof. Apply $e^{-i\theta Y}$ to the p-quoins to obtain N samples from q-quoins, and obtain an estimate $\hat{q}$ for the transformed parameter $q = \sin^2\left(\arcsin\left(\sqrt{p}\right) + \theta\right)$. The Chebyshev bound implies that

$$P\big(|\hat{q} - q| > \delta\big) < \frac{\Upsilon(q)}{N\delta^2} \ , \tag{7.53}$$

which can be used just as before to decide whether $q \in \{\sin^2\left(\arcsin\left(\sqrt{p}\right) + \theta\right) : p \in I_n\}$ – and therefore whether $p \in I_n$ – with error probability upper bounded by $\Upsilon(q)$. □

The same procedure holds when using amplitude estimation, where the error probability can instead by upper bounded by $\sqrt{\Upsilon(q)}$. Crucially, this bound goes to zero at the value q, so this technique can lower bound a function that evaluates to zero at that point. This is the source of the proven quantum advantage. To make this clear, we reframe the conditions on simulability for quantum Bernoulli factories.

Theorem 7.5: Quantum constraints made simple

There exists a quantum Bernoulli factory for the function $f : (0,1) \to [0,1]$ that uses single-qubit unitaries and classical post-processing if

1. f is continuous, and
2. either $f(p)$ is a constant, or for all $p \in (0,1)$, $\Xi_Z(p)^k \leq f(p) \leq 1 - \Xi_Z(p)^k$ for some $k \in \mathbb{N}$,

where, for a finite set $Z \subset (0,1)$,

$$\Xi_Z(p) := \prod_{z \in Z} \Upsilon(q_z) \tag{7.54}$$

with $q_z = \sin^2\left(\arcsin\left(\sqrt{p}\right) - \arcsin\left(\sqrt{z}\right)\right)$.

Proof. Again the fuzzy digit functions can be constructed, using either classical or quantum parameter estimation to decide whether $p \in I_n$. To bound the error probability by $\Xi_Z(p)$, apply $e^{-i\theta Y}$ with $\theta = -\arcsin\sqrt{z}$, perform the required parameter estimation and repeat for each $z \in Z$. If all of these rounds of parameter estimation agree that $p \in I_n$, then output 1. Moreover, this can be repeated k times to obtain the $\Xi_Z(p)^k$ bounds. □

These conditions are the same as those in Theorem 7.2 because

$$q_z = \sin^2\left(\arcsin\left(\sqrt{p}\right) - \arcsin\left(\sqrt{z}\right)\right) = \left(\sqrt{p(1-z)} - \sqrt{z(1-p)}\ \right)^2 , \tag{7.55}$$

however the source of the bounds is made obvious. In particular, it highlights the importance of destructive interference for extending the class of simulable functions, while parameter estimation constrains the resources required to perform the algorithm.

7.11 Outlook

In summary, we identified parameter estimation as the source of constraints on simulable functions for a classical Bernoulli factory. We presented a construction based on digit functions, providing an intuitive interpretation for the general constructions of Keane & O'Brien [101] and Wästlund [112]. Furthermore, we demonstrated how the quantum advantage consists of two components:

- a square-root reduction in the number of coins required, arising from using quantum amplitude estimation in the construction of the fuzzy digit functions, and
- the extension of the class of simulable functions by means of destructive interference.

In doing so, we derived upper bounds on the resources required to simulate the required fuzzy digit functions.

Future work could improve these bounds, perhaps by being less conservative in bounding the fuzzy digit functions. More efficient constructions exist for most functions, however the fuzzy digit functions may pave the way for more practically useful Bernoulli factory algorithms. Perhaps the most pressing direction for future research is to explore whether the quantum advantage can be exploited for statistical modelling. One might naively attempt to upload a classical p-coin to the quantum setting by classically estimating p and uploading that estimate to give a 'noisy' p-quoin, upon which the quantum algorithm can be implemented. The resulting quantum state that this scheme uses as a p-quoin would be

$$\rho = \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \ |\hat{p}_n(j)\rangle\langle\hat{p}_n(j)| \ , \tag{7.56}$$

where $|\hat{p}_n(j)\rangle = \sqrt{1-\frac{j}{n}}\,|0\rangle + \sqrt{\frac{j}{n}}\,|1\rangle$. This can be used for approximate simulation, however it is unclear if these noisy p-quoins offer any advantage over their classical counterparts, since the same number of p-coins can match the accuracy of the approximation offered by the noisy p-quoins. Further work is required to exploit the quantum advantage, and we conjecture that a

fully coherent version of the algorithm must be used, such as the version described in [124], where both the inputs and outputs are quantum states. Future research could apply this fully coherent Bernoulli factory algorithm in conjunction with other quantum resampling algorithms, such as the quantum Metropolis algorithm [125] and quantum rejection sampling algorithm [126], analogous to the use of classical Bernoulli factories in statistical modelling applications.

Chapter 8

Kitaev-Webb Resampling

Quantum computing is still in its infancy, and scientists are grappling with how best to build algorithms for these platforms. At this stage, one programme of research is to collect together useful computational primitives, routines that form part of several larger algorithms. From this toolkit, quantum information scientists can then construct new algorithms in a more systematic way. We should view these computational building blocks as resources that can be exploited and manipulated in the design of quantum algorithms.

One class of such building blocks consists of resampling algorithms. Bernoulli factories, discussed in the previous chapter, constitute just one type of resampling algorithm. We often require more flexible resampling algorithms, capable of manipulating wider classes of probability distribution. Moreover, we should like to achieve this on quantum hardware, with the resampling achieved coherently for use in further algorithmic steps. This chapter demonstrates how the Kitaev-Webb (KW) algorithm fulfils these goals, and looks at the computational resources required for its successful implementation. The outline is as follows:

- We define the resampling problem for continuous probability distributions and motivate why such a building block is useful for quantum simulation.

- An idealised version of the KW resampling algorithm is described, with analytical results

for the fidelity of this algorithm.

- We present preliminary numerical results for the fidelity of the KW algorithm, highlighting the effects of discreteness.

- The KW algorithm is extended to non-linear rescalings, and we give analytical results for the fidelity in the idealised setting.

A key aim of this chapter is to develop an intuitive presentation of the KW algorithm and to suggest it as a useful component in our quantum algorithm toolkit. The novel contributions of this chapter are:

- A simpler, idealised version of the KW algorithm for infinite continuous-variable quantum registers is given in Section 8.2, from which analytical results for the fidelity of the algorithm are derived (Theorem 8.1).

- Example 8.5 presents numerical calculations of the fidelity for the KW algorithm on qubit-based quantum registers that are within reach with near-term platforms.

- A new generalisation of the KW algorithm is described in Section 8.4, together with analytical results for the fidelity of an idealised version of this algorithm.

This chapter describes unpublished work in collaboration with David Jennings.

8.1 Resampling Continuous Probability Distributions

Bernoulli factories transform one Bernoulli distribution into another. It is natural to ask whether there exist resampling algorithms that can be implemented on a quantum computer for other types of distribution. This chapter is concerned with the following problem:

> *Use a continuous random variable $X \sim p(x)$ to simulate another continuous random variable $Y = f(X) \sim q(y)$ for some function $f : \mathbb{R} \to \mathbb{R}$.*

Throughout, we assume that f is invertible and analytic (and hence differentiable). This implies that the output distribution takes a particular form.

Lemma 8.1: Transforming probability density functions

Given a continuous random variable $X \sim p(x)$ and an invertible differentiable function $f : \mathbb{R} \to \mathbb{R}$, the probability density function for $Y = f(X)$ is

$$q(y) = \frac{1}{f'\big(f^{-1}(y)\big)}\, p\big(f^{-1}(y)\big)\ , \tag{8.1}$$

assuming $f'\big(f^{-1}(y)\big) \neq 0$.

Proof. The probability that X produces an outcome in the region $R \subseteq \mathbb{R}$ is given by

$$P(X \in R) = \int_R dx\ p(x). \tag{8.2}$$

Similarly, for the random variable Y with probability density function $q(y)$,

$$P(Y \in R) = \int_R dy\ q(y). \tag{8.3}$$

Given that $Y = f(X)$ for some invertible function f,

$$P(Y \in R) = P(f(X) \in R) = P\big(X \in f^{-1}(R)\big) = \int_{f^{-1}(R)} dx\ p(x). \tag{8.4}$$

Under a change of variables $y = f(x)$, this becomes

$$P(Y \in R) = \int_R dy\ \frac{1}{f'\big(f^{-1}(y)\big)}\, p\big(f^{-1}(y)\big), \tag{8.5}$$

assuming that $f'\big(f^{-1}(y)\big) \neq 0$ for all $y \in \mathbb{R}$. We therefore identify $q(y) = \frac{1}{f'\left(f^{-1}(y)\right)}\, p\big(f^{-1}(y)\big)$ as the transformed probability density function. □

Example 8.1: Linear rescaling

An important instance of this problem that we shall discuss extensively in this chapter is linear rescaling: $f(x) = ax$ for some constant $a \in \mathbb{R}$ (where $a \neq 0$). In this case, $q(y) = \frac{1}{a} p\left(\frac{y}{a}\right)$.

A quantum-mechanical analogue of this linear rescaling was described by Kitaev & Webb [127]. The input and output distributions are encoded as quantum states. Measuring the system in these states corresponds to drawing samples from the corresponding random variables. To take into account the Born rule, the probability density function $p(x)$ is encoded in the quantum state

$$|\psi\rangle = \int_{\mathbb{R}} dx\ \psi(x)\,|x\rangle \tag{8.6}$$

where $\psi(x) = \sqrt{p(x)}$ (we assume the wavefunction is real for simplicity). The aim of the protocol is to transform this state into

$$|\xi\rangle = \int_{\mathbb{R}} dx\ \frac{1}{\sqrt{f'(f^{-1}(x))}}\ \psi(f^{-1}(x))\,|x\rangle\,. \tag{8.7}$$

Example 8.2: Kitaev-Webb rescaling

The original Kitaev-Webb (KW) scheme focuses on the case of Example 8.1, namely rescaling by a linear function $f(x) = ax$. The target state is

$$|\xi\rangle = \int_{\mathbb{R}} dx\ \frac{1}{\sqrt{a}}\ \psi\left(\frac{x}{a}\right)|x\rangle\,. \tag{8.8}$$

Note that, like the Bernoulli factory, the transformation should be independent of the input distribution. This coherent resampling scheme would be useful for quantum simulation, as the following example demonstrates.

Example 8.3: Rescaling Gibbs states

Consider a continuous coherent Gibbs state,

$$|\gamma_\beta\rangle = \int_0^\infty dE\ \sqrt{\beta e^{-\beta E}}\ |E\rangle\,, \tag{8.9}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature. The KW algorithm implements a linear rescaling on this distribution, which is transformed to

$$\left|\gamma_{\beta/a}\right\rangle = \int_0^\infty dE\ \sqrt{\frac{\beta}{a}\ e^{-\frac{\beta}{a}E}}\ |E\rangle\,, \tag{8.10}$$

the corresponding coherent Gibbs state at inverse temperature $\frac{\beta}{a}$.

The ability to manipulate the temperature of the Gibbs state is clearly desirable in simulations of quantum thermodynamics, but also more broadly in quantum many-body physics and quantum gravity. For example, the thermo-field double state

$$|\text{TFD}\rangle = \frac{1}{\sqrt{Z}} \sum_n e^{-\frac{\beta E_n}{2}} |n\rangle \otimes |n\rangle \tag{8.11}$$

is dual to an eternal black hole in the AdS-CFT correspondence [128], and it has also been used to implement local scale transformations in simulations of the thermal properties of the quantum XY/Ising chain [129].

8.2 The Core KW Algorithm

Given the aim of providing an intuitive understanding of the KW algorithm, let us consider an idealised version of the algorithm, without the discreteness arising from finite numbers of qubits and free from the effects of noise.

The KW algorithm uses two quantum registers, which we take to be infinite and continuous. The first register X is initialised in the state

$$|\psi\rangle_X = \int_{\mathbb{R}} dx \ \psi(x) \, |x\rangle_X \, , \tag{8.12}$$

where $|\psi(x)|^2 = p(x)$ encodes the input distribution. The second register Y is prepared in the state

$$|\phi_n\rangle_Y = \frac{1}{\sqrt{2n}} \int_{-n}^{n} dy \ \ |y\rangle_Y \, , \tag{8.13}$$

where n is assumed to be small. Intuitively, these two registers generate x and y-coordinate axes, with a straightforward mapping between a point (x, y) in $\mathbb{R}^2$ and the bipartite basis state $|x\rangle_X |y\rangle_Y$. The wavefunction of the two-register state can be plotted on a third axis, as in Figure 8.1.

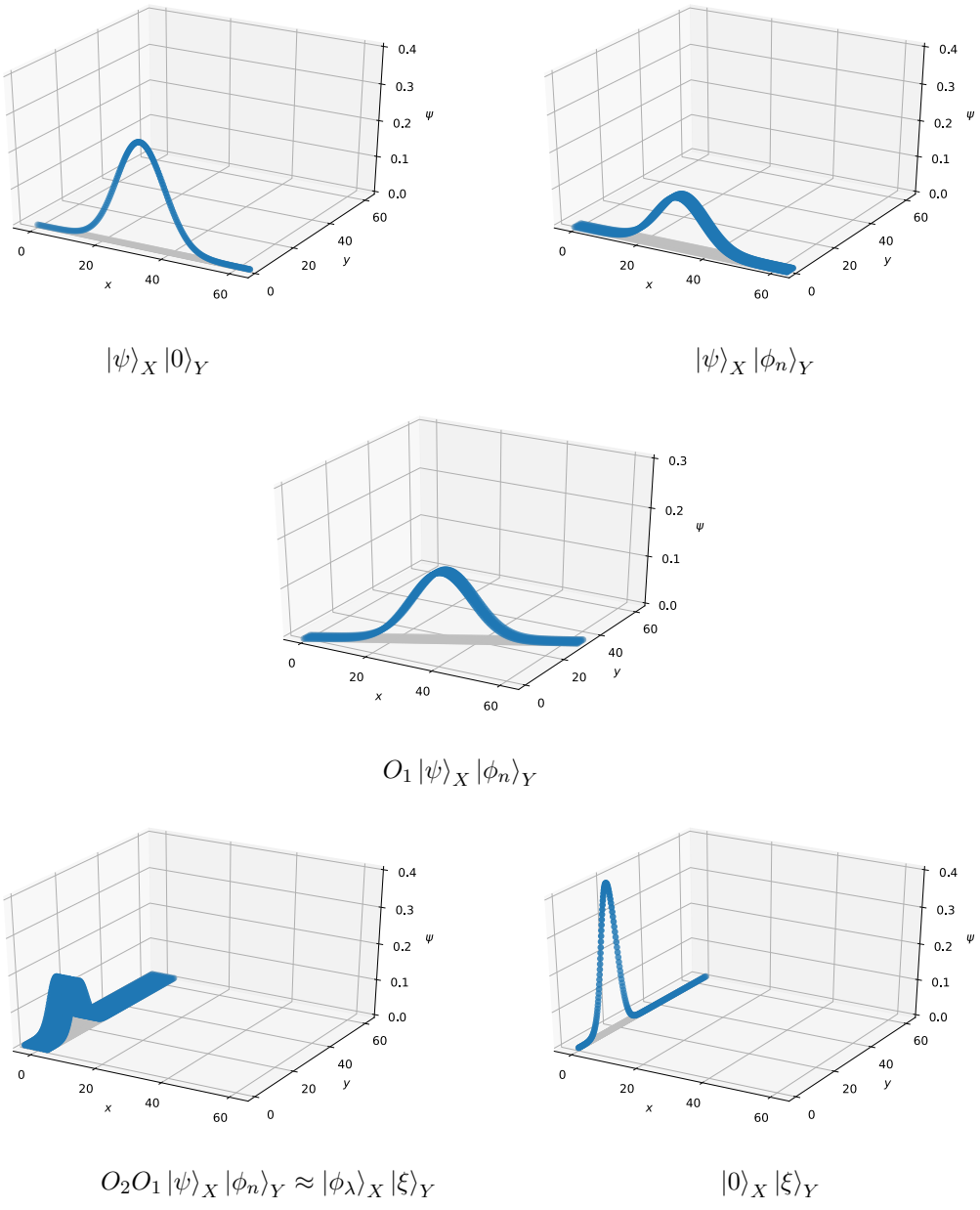


Figure 8.1: Graphical intuition for KW resampling. *The initial distribution is converted into a thin band of wavefunction over the two registers. The first oracle O_1 shifts this band from the x-axis up to the line $y = ax$, then the second oracle O_2 shifts over to the y-axis. Finally the first register is reset to $|0\rangle$, leaving the second register with the transformed distribution.*

First Oracle Two oracle calls perform the rescaling, enacted as unitary operations on the two registers. The first unitary O_1 acts upon the computational basis as

$$O_1 \left|x\right\rangle_X \left|y\right\rangle_Y = \left|x\right\rangle_X \left|ax+y\right\rangle_Y , \tag{8.14}$$

being an oracle that evaluates the function $f(x) = ax$. The registers are now in the state

$$O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y = \frac{1}{\sqrt{2n}} \int_{-n}^{n} dy \int_{\mathbb{R}} dx \ \psi(x) \left|x\right\rangle_X \left|ax+y\right\rangle_Y . \tag{8.15}$$

As pictured in Figure 8.1, the band of non-zero wavefunction has shifted from the x-axis up to the line $y = ax$.

Second Oracle A second unitary evaluates the inverse function $f^{-1}(x) = \frac{x}{a}$, such that

$$O_2 \left|x\right\rangle_X \left|y\right\rangle_Y = \left|x - \frac{y}{a}\right\rangle_X \left|y\right\rangle_Y \tag{8.16}$$

on the computational basis. This is similar to the first oracle, but with the role of the registers swapped. This leaves the state

$$O_2 \ O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y = \frac{1}{\sqrt{2n}} \int_{-n}^{n} dy \int_{\mathbb{R}} dx \ \psi(x) \left|-\frac{y}{a}\right\rangle_X \left|ax+y\right\rangle_Y . \tag{8.17}$$

The band of non-zero wavefunction moves over to the y-axis in Figure 8.1. After a change of variables to $u = -\frac{y}{a}$ and $v = ax + y$, this becomes

$$O_2 \ O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y = \frac{1}{\sqrt{2n}} \int_{-\lambda}^{\lambda} du \int_{\mathbb{R}} dv \ \psi\left(\frac{v}{a} + u\right) \left|u\right\rangle_X \left|v\right\rangle_Y \tag{8.18}$$

where $\lambda := \frac{n}{a}$. We then make two assumptions:

1. n is small, such that $|u| \leq n/a$ is also small, and

2. $\psi(x)$ is smooth and continuous,

which ensure that $\psi\left(\frac{v}{a} + u\right) \approx \psi\left(\frac{v}{a}\right)$, hence

$$O_2 \ O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y \approx \frac{1}{\sqrt{2\lambda}} \int_{-\lambda}^{\lambda} du \left|u\right\rangle_X \int_{\mathbb{R}} dv \ \frac{1}{\sqrt{a}} \psi\left(\frac{v}{a}\right) \left|v\right\rangle_Y = \left|\phi_\lambda\right\rangle_X \left|\xi\right\rangle_Y , \tag{8.19}$$

where

$$|\xi\rangle_B = \int_{\mathbb{R}} dx \; \frac{1}{\sqrt{a}} \; \psi\left(\frac{x}{a}\right) |x\rangle_Y \tag{8.20}$$

encodes the target distribution $q(x) = \frac{1}{a} p\left(\frac{x}{a}\right)$. The next section quantifies this approximation. The support of the uniform superposition changes during the protocol from $[-n, n]$ to $[-\lambda, \lambda]$, providing the $1/a$ factor in $q(x)$. If the second register were instead initialised in the state $|0\rangle_Y$, this prefactor would not have appeared. Finally, the first register should be uncomputed if the output is required within the rest of a larger quantum algorithm.

8.2.1 Fidelity of the KW Algorithm

The KW algorithm only approximately resamples the probability distribution. We require that $\psi\left(\frac{v}{a} + u\right) \approx \psi\left(\frac{v}{a}\right)$, therefore the fidelity of the protocol depends on two factors:

- *(protocol)* the ratio $\lambda = n/a$,
- *(input distribution)* the gradient $\frac{d\psi}{dx}$ of the input wavefunction.

Theorem 8.1: KW fidelity

If $\psi(x)$ is analytic and $\lambda \ll 1$, the final-state fidelity of the KW algorithm (with $a \neq 0$) on infinite, continuous registers is $F = 1 - \frac{\lambda^2}{3} \mathrm{Var}_\psi(\hat{p}) + \mathcal{O}(\lambda^4)$, where $\hat{p} = -i\frac{\partial}{\partial x}$.

Proof. The density matrix of the final state in the KW scheme is

$$\rho_{XY} = O_2 \; O_1 \; |\psi\rangle\langle\psi|_X \otimes |\phi_n\rangle\langle\phi_n|_Y \; O_1^\dagger \; O_2^\dagger \tag{8.21}$$

$$= \frac{1}{2n} \int_{\mathbb{R}} dx \int_{\mathbb{R}} dx' \int_{-\lambda}^{\lambda} dy \int_{-\lambda}^{\lambda} dy' \; \psi\left(\frac{x}{a} + y\right) \; \psi^*\left(\frac{x'}{a} + y'\right) |y\rangle\langle y'|_X \otimes |x\rangle\langle x'|_Y \,, \tag{8.22}$$

therefore the reduced density matrix for the second register is

$$\rho_Y = \frac{1}{2n} \int_{\mathbb{R}} dx \int_{\mathbb{R}} dx' \int_{-\lambda}^{\lambda} dy \; \psi\left(\frac{x}{a} + y\right) \; \psi^*\left(\frac{x'}{a} + y\right) |x\rangle\langle x'|_Y \,. \tag{8.23}$$

The success of the protocol is measured by the fidelity between ρ_Y and the intended state,

$$|\xi\rangle = \int_{\mathbb{R}} dx \; \frac{1}{\sqrt{a}} \; \psi\left(\frac{x}{a}\right) |x\rangle_Y \,. \tag{8.24}$$

This fidelity is given by

$$F = \langle\xi|\,\rho_Y\,|\xi\rangle \tag{8.25}$$

$$= \frac{1}{2na}\int_{\mathbb{R}^4} dz\ dz'\ dx\ dx' \int_{-\lambda}^{\lambda} dy\ \psi^*\left(\frac{z}{a}\right)\psi\left(\frac{x}{a}+y\right)\psi^*\left(\frac{x'}{a}+y\right)\psi\left(\frac{z'}{a}\right)\langle z|x\rangle\left\langle x'\middle|z'\right\rangle \tag{8.26}$$

$$= \frac{1}{2n}\int_{-n}^{n} dj \int_{\mathbb{R}} dx\ \psi^*(x)\ \psi\left(x+\frac{j}{a}\right)\int_{\mathbb{R}} dx'\ \psi(x')\ \psi^*\left(x'+\frac{j}{a}\right). \tag{8.27}$$

Given that $\psi(x)$ is analytic, Taylor expand in terms of j/a,

$$F = \frac{1}{2n}\int_{-n}^{n} dj \left[1+\frac{j}{a}\int_{\mathbb{R}} dx\ \psi^*(x)\frac{\partial}{\partial x}\psi(x)+\frac{j^2}{2a^2}\int_{\mathbb{R}} dx\ \psi^*(x)\frac{\partial^2}{\partial x^2}\psi(x)+\mathcal{O}\left(\frac{j^3}{a^3}\right)\right]$$
$$\left[1+\frac{j}{a}\int_{\mathbb{R}} dx\ \psi(x)\frac{\partial}{\partial x}\psi^*(x)+\frac{j^2}{2a^2}\int_{\mathbb{R}} dx\ \psi(x)\frac{\partial^2}{\partial x^2}\psi^*(x)+\mathcal{O}\left(\frac{j^3}{a^3}\right)\right] \tag{8.28}$$

$$= \frac{1}{2n}\int_{-n}^{n} dj\ \left[1+i\frac{j}{a}\langle\hat{p}\rangle-\frac{j^2}{2a^2}\langle\hat{p}^2\rangle+\mathcal{O}\left(\frac{j^3}{a^3}\right)\right]\left[1-i\frac{j}{a}\langle\hat{p}\rangle-\frac{j^2}{2a^2}\langle\hat{p}^2\rangle+\mathcal{O}\left(\frac{j^3}{a^3}\right)\right] \tag{8.29}$$

$$= 1-\frac{\lambda^2}{3}\left[\langle p^2\rangle-\langle p\rangle^2\right]+\mathcal{O}(\lambda^4) = 1-\frac{\lambda^2}{3}\mathrm{Var}_\psi(\hat{p})+\mathcal{O}(\lambda^4) \tag{8.30}$$

where $\hat{p} = -i\frac{\partial}{\partial x}$. We neglect higher powers of λ since it is assumed that $\lambda \ll 1$. □

Higher fidelities are achieved when the the derivative of the distribution (encapsulated in the variance of the momentum) is lower. Intuitively, for rapidly varying wavefunctions, the necessary approximations will not hold so well and the algorithm will be less effective. Likewise, the size n of the support for the uniform superposition in $|\phi_n\rangle$ must be balanced against the rescaling factor a, such that $\lambda = \frac{n}{a}$ is minimised. For some simple examples, the fidelity can be expressed analytically.

Example 8.4: Normal Distribution in the KW Algorithm

A Normal distribution with standard deviation σ is encoded in the state

$$|\psi\rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} \int_{\mathbb{R}} dx \ \exp\left\{-\frac{x^2}{4\sigma^2}\right\} |x\rangle \,. \tag{8.31}$$

After the KW algorithm, the fidelity of the final state is

$$F = \frac{1}{2n} \int_{-n}^{n} dj \int_{\mathbb{R}} dx \ \psi^*(x) \ \psi\left(x + \frac{j}{a}\right) \int_{\mathbb{R}} dx' \ \psi(x') \ \psi^*\left(x' + \frac{j}{a}\right) \tag{8.32}$$

$$= \frac{1}{4\pi n\sigma^2} \int_{-n}^{n} dj \ \left| \int_{\mathbb{R}} dx \ e^{-\frac{x^2}{4\sigma^2}} e^{-\frac{(x+j/a)^2}{4\sigma^2}} \right|^2 \tag{8.33}$$

$$= \frac{\sqrt{\pi}\sigma}{\lambda} \mathrm{Erf}\left(\frac{\lambda}{2\sigma}\right), \tag{8.34}$$

where Erf denotes the error function. The fidelity depends only upon the ratio between λ, which encapsulates the scale over which we hope the distribution is approximately constant, and σ, which determines how quickly the distribution varies with x.

8.3 Implementing the KW Protocol

The original KW scheme was tailored to qubit-based quantum computers [127]. Changes must be made to the idealised algorithm presented in the previous section for implementation in a more realistic setting.

- The quantum registers consist of N qubits, with computational basis $\{|0\rangle, \ldots, |M-1\rangle\}$, where $M = 2^N$.

- The first register X is initialised in the state

$$|\psi\rangle_X = \sum_{j=0}^{M-1} \psi(j) \, |j\rangle_X \tag{8.35}$$

  where $P(j) = |\psi(j)|^2$ is a discretized (possibly binned) representation of the probability density $p(x)$ (see [5, 127, 130] for possible preparation algorithms of efficiently integrable

distributions). The second register Y begins in the state

$$|\phi_n\rangle_Y = \frac{1}{\sqrt{2n}} \sum_{k=-n}^{n-1} |k\rangle_Y \tag{8.36}$$

where $|-k\rangle = |M-k\rangle$.

- The unitary oracles are discretized, and act on the computational basis as

$$O_1 |j\rangle_X |k\rangle_Y = |j\rangle_X |\lfloor aj \rfloor + k\rangle_Y \tag{8.37a}$$

$$O_2 |j\rangle_X |k\rangle_Y = \left| j - \left\lfloor \frac{k}{a} \right\rfloor \right\rangle_X |k\rangle_Y \tag{8.37b}$$

where the arithmetic is calculated modulo M.

The protocol proceeds as in the idealised case, but with the above modifications to handle the finiteness and discreteness of the registers.

Even on a fault-tolerant, noiseless quantum computer, the KW scheme relies on approximation, and there would be some loss of fidelity because of the discreteness introduced by the qubit registers. For small enough systems, such as those that we realistically expect to access in the near term, we can numerically calculate the fidelity of this algorithm. The density matrix for the state on the second register at the end of the protocol is

$$\rho_Y = \frac{1}{2n} \sum_{x,y=0}^{M-1} \sum_{j,k=-n}^{n-1} \psi(x)\psi^*(y) \left\langle y - \left\lfloor \frac{\lfloor ay \rfloor + k}{a} \right\rfloor \middle| x - \left\lfloor \frac{\lfloor ax \rfloor + j}{a} \right\rfloor \right\rangle |\lfloor ax \rfloor + j\rangle \langle \lfloor ay \rfloor + k| \,. \tag{8.38}$$

The target state is

$$|\xi\rangle = \sum_{x=0}^{M-1} \xi(x) |x\rangle = \sum_{x=0}^{M-1} \frac{1}{\sqrt{a}} \psi\left(\frac{x}{a}\right) |x\rangle \,. \tag{8.39}$$

The final-state fidelity is therefore

$$F = \langle\xi|\,\rho_Y\,|\xi\rangle \tag{8.40}$$

$$= \frac{1}{2n}\sum_{x,y,u,v}\sum_{j,k=-n}^{n-1}\psi(x)\psi^*(y)\xi^*(u)\xi(v)\left\langle y - \left\lfloor\frac{\lfloor ay\rfloor + k}{a}\right\rfloor\middle| x - \left\lfloor\frac{\lfloor ax\rfloor + j}{a}\right\rfloor\right\rangle\langle u|\lfloor ax\rfloor + j\rangle\,\langle\lfloor ay\rfloor + k|v\rangle \tag{8.41}$$

$$= \frac{1}{2n}\sum_{x,y}\sum_{j,k=-n}^{n-1}\psi(x)\psi^*(y)\xi^*(\lfloor ax\rfloor + j)\xi(\lfloor ay\rfloor + k)\left\langle y - \left\lfloor\frac{\lfloor ay\rfloor + k}{a}\right\rfloor\middle| x - \left\lfloor\frac{\lfloor ax\rfloor + j}{a}\right\rfloor\right\rangle, \tag{8.42}$$

where x, y, u and v are all summed from 0 to $M-1$. This again neglects the effects of noise, which would further reduce the fidelity.

Example 8.5: Normal Distribution Fidelity

Consider the practically important example of a normal distribution. The input state on the first N-qubit register is

$$|\psi\rangle_X = \sum_{j=0}^{M-1}\psi(j)\,|j\rangle_X \tag{8.43}$$

where

$$\psi(j) = \left[\int_{j-\frac{1}{2}}^{j+\frac{1}{2}} dx\ \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right]^{\frac{1}{2}} \tag{8.44}$$

for $1 \leq j < M-1$, while for $j = 0$ and $j = M-1$ one of the integration limits extends out to $-\infty$ and ∞ respectively. This encodes a binned version of the probability density function. We assume that μ and σ are such that

$$\int_0^M dx\ \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \approx 1, \tag{8.45}$$

so that the endpoints of the wavefunction $\psi(j)$ are approximately zero, thus boundary effects can be neglected.

It is reasonable to expect access to 4 quantum registers (2 primary registers, 2 auxiliary workspace registers for evaluating the function), each with 5 qubits, in the near term. The final-state fidelity of the KW algorithm is plotted in Figure 8.2 as a function of λ. The data collapse onto the idealised case at higher values of λ, although the fidelities are

low in that regime. At lower λ, the fidelity is higher, however there is more variation because of discreteness.

Given the higher fidelities for lower values of λ, one might choose n to be as small as possible, say $n = 1$, however Figure 8.2 shows that the algorithm becomes more susceptible to the discreteness. For $n = 1$, the fidelity varies considerably for different values of a, corresponding to regularities in the rounding of the floor function (note the peaks at $a = 1/2$ and $2/3$). Although the fidelity is lower for $n = 4$, the effects of discreteness are much reduced.

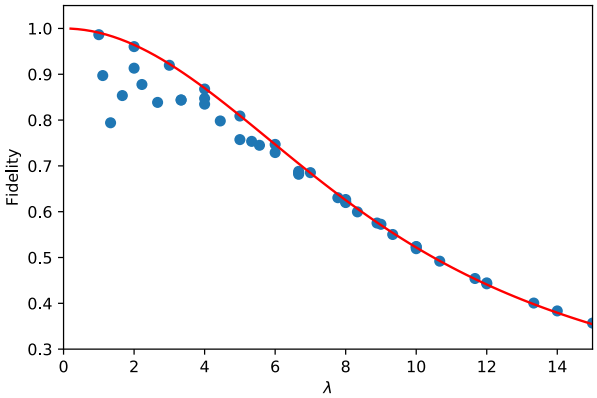
Reasonable fidelities can be achieved with as few as 5 qubits in each register, although more are obviously desirable. One should strike a balance between lowering n to increase the fidelity, and mitigating against the effects of discreteness when n is too small. The uniform superposition $|\phi_n\rangle$ not only helps to renormalise the distribution, but also to smooth out the discreteness.

The KW algorithm is clearly a practically useful algorithm that could be implemented on quantum computers in the relatively near future. Further optimisation given constraints on allowed gates and architectures is obviously needed, but there is clear promise for implementation on NISQ platforms.

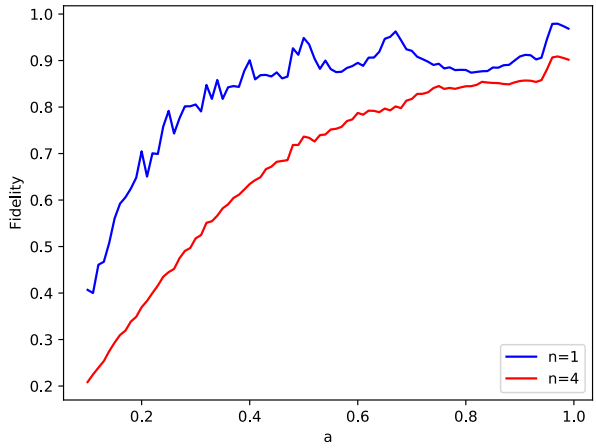
8.4 Extending the KW Protocol to Non-linear Rescaling

The original KW algorithm implements a linear rescaling of the input distribution using oracles that evaluate $f(x) = ax$. This immediately extends to a general linear function $f(x) = ax + b$. Such rescaling is a useful tool for quantum simulation, however non-linear rescaling of distributions is also of potential interest, for example when simulating phenomena on a curved spacetime.

Consider again the idealised setting of infinite and continuous quantum registers, but with



Final-state fidelity for the discrete KW protocol. *Fidelities for a range of a values are included (0.5, 0.6, 0.75, 0.9 and 1). The red curve shows the fidelity for the idealised KW algorithm.*



Effects of discretization. *The KW algorithm with a lower value of n has a higher fidelity. However, very low n are susceptible to the effects of discreteness in the algorithm, as seen by the variability in the fidelity for $n = 1$.*

Figure 8.2: Fidelity for Discretized KW Algorithm. *These fidelities are calculated on registers of $N = 5$ qubits for the Normal distribution example with $\mu = 16$ and $\sigma = 3$.*

modified unitary oracles

$$O_1 \left|x\right\rangle_X \left|y\right\rangle_Y = \left|x\right\rangle_X \left|f(x)+y\right\rangle_Y \tag{8.46a}$$

$$O_2 \left|x\right\rangle_X \left|y\right\rangle_Y = \left|x - f^{-1}(y)\right\rangle_X \left|y\right\rangle_Y , \tag{8.46b}$$

explicitly requiring that f be invertible. As in the linear case, the states $\left|\psi\right\rangle_X$ and $\left|\phi_n\right\rangle_Y$ are initialised in the two quantum registers. The above oracles are then called, which gives

$$O_2\ O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y = \frac{1}{\sqrt{2n}} \int_{-n}^{n} dy \int_{\mathbb{R}} dx\ \psi(x) \left|x - f^{-1}\big(f(x)+y\big)\right\rangle_X \left|f(x)+y\right\rangle_Y . \tag{8.47}$$

If f is differentiable, the change of coordinates to $u = f(x)+y$ and $v = x - f^{-1}\big(f(x)+y\big)$ gives

$$O_2\ O_1 \left|\psi\right\rangle_X \left|\phi_n\right\rangle_Y \approx \frac{1}{\sqrt{2n}} \int_{\mathbb{R}} du \int_{-\lambda(u)}^{\lambda(u)} dv\ \psi\big(f^{-1}(u)+v\big) \left|v\right\rangle_X \left|u\right\rangle_Y \tag{8.48}$$

$$\approx \frac{1}{\sqrt{2n}} \int_{\mathbb{R}} du \int_{-\lambda(u)}^{\lambda(u)} dv\ \sqrt{f'\big(f^{-1}(u)\big)}\ \xi(u) \left|v\right\rangle_X \left|u\right\rangle_Y \tag{8.49}$$

$$= \int_{\mathbb{R}} du\ \xi(u) \left|\phi_{\lambda(u)}\right\rangle_X \left|u\right\rangle_Y \tag{8.50}$$

where $\lambda(u) := \frac{n}{f'(f^{-1}(u))}$. To avoid problematic divergences, we require that $f'(x) \neq 0$ for any $x \in \mathbb{R}$. The two registers remain entangled, however the state of the first register $\left|\phi_{\lambda(u)}\right\rangle$ can be uncomputed conditioned on the second register.

Theorem 8.2: Non-linear KW Fidelity

If $\lambda(x) \ll 1$ for all $x \in \mathbb{R}$ and $\psi(x)$ is analytic on $\mathbb{R}$, the KW scheme for non-linear rescaling by the invertible and differentiable function $f(x)$, with $f'(x) > 0$ for all $x \in \mathbb{R}$, has final-state fidelity

$$F = \left\langle f'^{\frac{1}{2}} \right\rangle \left\langle \frac{1}{f'^{\frac{1}{2}}} \right\rangle + \frac{n^2}{6} \left[2 \left\langle f'^{\frac{1}{2}} \hat{p} \right\rangle \left\langle \frac{\hat{p}}{f'^{\frac{5}{2}}} \right\rangle - \left\langle f'^{\frac{1}{2}} \hat{p}^2 \right\rangle \left\langle \frac{1}{f'^{\frac{5}{2}}} \right\rangle - \left\langle f'^{\frac{1}{2}} \right\rangle \left\langle \frac{\hat{p}^2}{f'^{\frac{5}{2}}} \right\rangle \right] + \mathcal{O}(n^4) \tag{8.51}$$

where $\hat{p} = -i\frac{\partial}{\partial x}$ and $\left\langle A(x) \right\rangle := \int_{\mathbb{R}} dx\ \psi^*(x)\ A(x)\ \psi(x)$.

Proof. We follow a similar strategy to the proof of Theorem 8.1. The final state in the second

register is

$$\rho_B = \frac{1}{2n}\int_{\mathbb{R}} dx \int_{\mathbb{R}} dx' \int_{-\lambda(x)}^{\lambda(x)} dy\ \psi(f^{-1}(x)+y)\ \psi^*(f^{-1}(x')+y)\ \left|x\right\rangle\left\langle x'\right|. \tag{8.52}$$

The final-state fidelity is

$$F = \langle\xi|\,\rho_B\,|\xi\rangle \tag{8.53}$$

$$= \int_{\mathbb{R}} dz\ \frac{\psi^*(f^{-1}(z))}{\sqrt{f'(f^{-1}(z)}}\,\langle z| \int_{\mathbb{R}^2} dx\ dx' \int_{-\lambda(x)}^{\lambda(x)} dy\ \frac{\psi(f^{-1}(x)+y)\ \psi^*(f^{-1}(x')+y)}{2n}\,|x\rangle\langle x'| \times \int_{\mathbb{R}} dz'\ \frac{\psi(f^{-1}(z'))}{\sqrt{f'(f^{-1}(z')}}\,|z'\rangle \tag{8.54}$$

$$= \int_{\mathbb{R}^2} dx\ dx' \int_{-\lambda(x)}^{\lambda(x)} dy\ \frac{\psi^*(f^{-1}(x))\ \psi(f^{-1}(x)+y)\ \psi^*(f^{-1}(x')+y)\ \psi(f^{-1}(x'))}{2n\sqrt{f'(f^{-1}(x)f'(f^{-1}(x')}} \tag{8.55}$$

$$= \int_{-n}^{n} dy \int_{\mathbb{R}} dx\ \frac{1}{2n\sqrt{f'(x)}}\ \psi^*(x)\psi\left(x+\frac{y}{f'(x)}\right)\int_{\mathbb{R}} dx'\ \sqrt{f'(x')}\ \psi^*\left(x'+\frac{y}{f'(x)}\right)\psi(x') \tag{8.56}$$

Expanding to second order,

$$\int_{\mathbb{R}} dx'\ \sqrt{f'(x')}\ \psi^*\left(x'+\frac{y}{f'(x)}\right)\psi(x')$$

$$= \int_{\mathbb{R}} dx'\ \sqrt{f'(x')}\ \psi(x')\left[\psi^*(x') + \frac{y}{f'(x)}\frac{\partial}{\partial x'}\psi^*(x') + \frac{y^2}{2f'(x)^2}\frac{\partial^2}{\partial x'^2}\psi^*(x') + \mathcal{O}(y^3)\right] \tag{8.57}$$

$$= \langle\sqrt{f'}\rangle - \frac{iy}{f'(x)}\langle\sqrt{f'}\hat{p}\rangle - \frac{y^2}{2f'(x)^2}\langle\sqrt{f'}\hat{p}^2\rangle + \mathcal{O}(y^3). \tag{8.58}$$

Substituting this into the fidelity and again expanding to second order,

$$F = \int_{-n}^{n} dy \int_{\mathbb{R}} dx\ \frac{\psi^*(x)\ \psi\left(x+\frac{y}{f'(x)}\right)}{2n\sqrt{f'(x)}}\ \left[\langle\sqrt{f'}\rangle - \frac{iy\langle\sqrt{f'}\hat{p}\rangle}{f'(x)} - \frac{y^2\langle\sqrt{f'}\hat{p}^2\rangle}{2f'(x)^2} + \mathcal{O}(y^3)\right] \tag{8.59}$$

$$= \left\langle\sqrt{f'}\right\rangle\left\langle\frac{1}{\sqrt{f'}}\right\rangle + \frac{n^2}{6}\left[2\left\langle\sqrt{f'}\hat{p}\right\rangle\left\langle\frac{\hat{p}}{(f')^{\frac{5}{2}}}\right\rangle - \left\langle\sqrt{f'}\hat{p}^2\right\rangle\left\langle\frac{1}{(f')^{\frac{5}{2}}}\right\rangle - \left\langle\sqrt{f'}\right\rangle\left\langle\frac{\hat{p}^2}{(f')^{\frac{5}{2}}}\right\rangle\right] + \mathcal{O}(n^4) \tag{8.60}$$

where $\hat{p} = -i\frac{\partial}{\partial x}$, neglecting higher order terms since $\lambda(x) \ll 1$. □

This result agrees with the fidelity in the linear case where $f'(x) = a$, since

$$F = \langle\sqrt{a}\rangle\left\langle\frac{1}{\sqrt{a}}\right\rangle + \frac{n^2}{6}\left[2\langle\sqrt{a}\hat{p}\rangle\left\langle\frac{\hat{p}}{a^{\frac{5}{2}}}\right\rangle - \langle\sqrt{a}\hat{p}^2\rangle\left\langle\frac{1}{a^{\frac{5}{2}}}\right\rangle - \langle\sqrt{a}\rangle\left\langle\frac{\hat{p}^2}{a^{\frac{5}{2}}}\right\rangle\right] + \mathcal{O}(n^4) \tag{8.61}$$

$$= 1 - \frac{\lambda^2}{3}\mathrm{Var}_\psi(\hat{p}) + \mathcal{O}(n^4). \tag{8.62}$$

This scheme can also be implemented on near-term devices with finite and discrete quantum registers, with similar modifications to those used in the linear case.

8.5 Conclusions & Outlook

This chapter has aimed to place the KW resampling algorithm firmly in the quantum programmer's toolkit as a useful building block for more complex algorithms. We have described an idealised version of the algorithm to highlight the core conceptual features. We presented a preliminary analysis of the fidelity of the algorithm, paying heed to the effects of discretization. We then considered how the scheme could be generalised to non-linear rescaling.

Further work is required to successfully implement this algorithm on a real quantum computer. Realistic implementations should attempt to optimise the circuits to reduce noise or errors, and take into account architectural constraints, such as connectivity between qubits. However, these problems do not seem insurmountable and pose no fundamental barrier to the use of this algorithm.

Another possible direction is to introduce elements from quantum-to-quantum Bernoulli factories into the KW scheme. Individual qubits of the quantum registers in the KW algorithm could be treated as quoins to be manipulated in a Bernoulli factory. Does this extend the coherent resampling options available to the quantum programmer?

The KW algorithm is clearly a useful computational resource, and further work should look to find new applications for this scheme beyond those discussed here. Thermodynamic and many-body physics applications of the algorithm are tantalising beginnings, and further directions in statistical modelling are to be expected.

Chapter 9

Concluding Remarks

Although this thesis divides into three main parts, a number of common themes run between them. As noted in the introduction, resource-led approaches are leading to many advances in quantum information theory. This thesis has discussed just a few aspects of this movement. Resource theories provide a useful umbrella for this approach, but hopefully this thesis has highlighted a broader meaning for resources in quantum information theory that includes computational resources and techniques.

The thesis also splits into two strands – structure and manipulation. The first, more abstract chapters on thermodynamics and symmetry focused more on the mathematical structures associated with resources, while the second strand on resampling algorithms was based on manipulating more concrete computational resources. These two strands are intertwined, with the structure of resources revealing how resources can be utilised, while new applications reveal new relationships and structures.

In this concluding chapter, we highlight a number of themes that reach across the various parts of this thesis.

Coarse-graining In both the thermodynamics and symmetry parts of this thesis, coarse-grained descriptions of physical systems played a central role. The structure of relationships

between resources is often highly complex, so different levels of description can be very useful in deriving important results. In the thermodynamics chapters, there was first coarse-graining by unitary orbits, before moving on to consider sets of asymptotically equivalent states, where the pair (E, S) is used to describe the many-copy limit. This highlights several different levels of coarse-graining, each of which may capture different elements of the physics; unitary orbits are relevant to passivity, while the asymptotically equivalent sets are key for describing transformations in the many-copy regime.

In the symmetry part of the thesis, the 'shape' of the group orbit provided the most coarse-grained non-trivial description of quantum states and channels under symmetry constraints. Here the coarse-graining eliminates much of the complexity inherent to the full description, yet useful information and constraints on dynamics can still be extracted. Once again, it is key to be able to move between different levels of description. For example, the group orbit description of symmetry properties given in this thesis dovetails with the more complex description of modes of asymmetry [79, 80], where representation theory is harnessed in a much more fine-grained description. The group orbit classification is a foundation for other descriptions that can supplement this coarse-grained information to varying extent.

When analysing constraints in classical and quantum Bernoulli factories, the coarser 'fuzzy' digit functions were at the heart of the general construction used. Inspired by the sharp digit functions in a traditional binary expansion of a function, the coarser description allowed parameter estimation from finite numbers of samples to be used. Considering the idealised KW algorithm in infinite and continuous quantum registers also allowed analytical results on the fidelity to be derived, smoothing out the effects of discretization and finite-size effects in real qubit-based implementations. In both these cases, coarse-grained perspectives provide useful insights into the algorithms inner workings and how they perform.

Group orbits Featuring in both Parts I and II of this thesis, group orbits are clearly a versatile tool when studying and manipulating resources. Passivity was defined over unitary orbits, with passive states being energetic minima of these orbits, which partitioned the state space. In Part II, other group actions were used to partition this space, focusing on physically important symmetry groups.

Group orbits fit within the coarse-graining discussed above, but they also provide a simple mathematical structure that can be built upon with other techniques. This might be the majorization-based analysis used in our novel proof of the passivity conditions for a quantum state, or the harmonic functions that arise in the modes of asymmetry (see [79] for an extended treatment of this). Group orbits have also been used in the study of entanglement [10], further supporting their inclusion in the quantum information theorist's toolkit when analysing resources.

Geometric perspectives Another theme that has cropped up throughout this thesis is the power of geometric perspectives. Many derivations of the Gibbs states as the completely passive states existed in the literature, but none as simple and transparent as the geometric derivation presented here. This geometric perspective on passivity reveals new results along the way, as well as connecting with previous work using the geometry of E-S diagrams to study thermodynamics in the many-copy limit [47].

Visualising group orbits or the structures connected with them is also helpful in the sections on symmetry properties. For example, plotting the T-states with each group orbit shape revealed a new level of structure in this important and well-studied set of quantum states. It is also likely that techniques from information geometry could be used in the future to build on the foundations provided by the group orbits.

We made extensive use of visual representations of algorithms in the resampling chapters, highlighting the power of geometric perspectives when manipulating resources in these more

applied settings, whether this was in seeing how the wavefunction was manoeuvred by the KW algorithm, or in understanding how the constraints on functions arose for classical and quantum Bernoulli factories. This power to aid intuition is by no means restricted to the analysis of resources, but it is nonetheless important to this approach.

Partial orderings Another recurring topic in the first half of the thesis was that of ordering. In the thermodynamics chapters, a novel proof of the passivity conditions was presented, based on the relative orderings of the energies and populations. Passivity was then recast as a total ordering of vectors in the two-dimensional plane, and this geometric ordering was then at the heart of deriving the Gibbs states. In the following chapter, the geometric athermality provided a new measure for deviation from equilibrium, another type of ordering when considering transformations that do not smuggle work or order into the system in the many-copy limit.

Partial orderings arise in any resource theory, caused by the chosen constraints. When these are symmetry constraints, the shape of the group orbit encodes coarse-grained information about this partial ordering. In fact, classifying group orbits in this way can be considered as a new type of resource measure that does not map states or channels onto a totally ordered set such as the real numbers. Instead it preserves some aspects of the full structure but in a far simpler setting. This opens up a new avenue for studying resources by mapping complex scenarios to simpler resource theories that encode some key features.

It is important to compare different resources in quantum information. Some may be easier to create or maintain than others, so these comparisons can have ramifications for information processing protocols. For example, the shape of the group orbit tells us what the minimal resources are to simulate a given channel under symmetry constraints. If procuring these resources is costly, it is important to compare different resources and decide on the best implementation.

Constraints Finally, resources are defined by constraints. These may be fundamental constraints imposed by the physics of the situation, or more practical constraints relating to exper-

imental difficulty. This thesis has been about finding new ways to analyse these constraints and to understand their consequences and the structures they create, presenting a number of new tools and features that will continue to be built upon and used in this direction of research.

Constraints are inevitable and often one of our main considerations in quantum information theory. As an approach to the subject, the study of constraints and the resources needed to overcome them has been a fruitful and effective philosophy. Moreover, constraints are often vital to our information processing tasks, preserving and protecting information, and the structures they build are some of the most fascinating to study. Quantum information theory will continue to be the study of the structure and manipulation of resources.

Part IV

Appendices

Appendix A

Proof of geometric athermality monotonicity

In the main text, we provided a proof for the monotonicity of $\mathcal{A}$ *for qutrits*. However, the same monotonicity holds for systems of higher dimension that are transformed under activation trajectories, however extending the proof of the qutrit case to this more general setting introduces several technicalities. The proof uses transformations on virtual qutrit systems within the full $d > 3$-dimensional system, deforming the asymptotic ensemble in a way that ensures monotonicity of $\mathcal{A}$.

A.1 Virtual Qutrit Deformations

The key tools in this proof are virtual qutrit systems within the full d-dimensional system.

Definition A.1: Virtual qutrit

Consider a d-dimensional ($d \geq 3$) quantum system with non-degenerate Hamiltonian H. A virtual qutrit is isolated from a passive non-Gibbsian state ρ by selecting three non-colinear elements $\mathcal{V}_{\text{qut}}(\rho) \coloneqq \{\boldsymbol{v}_{k_1}, \boldsymbol{v}_{k_2}, \boldsymbol{v}_{k_3}\} \subseteq \mathcal{V}(\rho)$.

This selects a triangle of points from within $\mathcal{V}(\rho)$. Considering transformations of a virtual

qutrit system allows us to isolate the changes in area and see how (E, S) changes. Such qutrit deformations can be composed to give any allowed transformation.

Definition A.2: Qutrit deformation

A qutrit deformation is an infinitesimal transformation of ρ generated by varying the three occupation probabilities $p_{k_i} \to p_{k_i} + \mathrm{d}p_{k_i}$ of the virtual qutrit $\mathcal{V}_{\mathrm{qut}}(\rho)$ in isolation, such that $\sum_{k_i} \mathrm{d}p_{k_i} = 0$, inducing the change $A(\rho) \to A(\rho) + \mathrm{d}A = A(\rho')$ in the area of $\mathcal{V}_\infty(\rho)$.

We want to use qutrit deformations to manipulate (E, S) in the E-S diagram, shown to always be possible in the following lemma.

Lemma A.1: Non-trivial qutrit deformations

Given ρ and H as in Definition A.1, there exist non-trivial qutrit deformations to $\rho \in \mathcal{D}_{(E,S)}$ that move (E, S) either isentropically ($\mathrm{d}S = 0$) or isoenergetically ($\mathrm{d}E = 0$).

Proof. Consider a virtual qutrit with vertices $\mathcal{V}_{\mathrm{qut}}(\rho)$ which we label $\boldsymbol{v}_1 \leq \boldsymbol{v}_2 \leq \boldsymbol{v}_3$ without loss of generality, further assuming that the points in $\mathcal{V}_{\mathrm{qut}}(\rho)$ are not colinear (i.e. our virtual qutrit is not Gibbsian). Such a virtual qubit can always be found if ρ is not completely passive. A qutrit deformation is constrained to preserve normalisation of the virtual qutrit, such that $\sum_{i=1}^{3} \mathrm{d}p_i = 0$.

- For an isentropic transformation ($\mathrm{d}S = 0$),

$$\mathrm{d}E = \begin{cases} \Delta_{3,1}\left[\frac{\beta_{1,3}}{\beta_{2,3}} - 1\right] \mathrm{d}p_1, & p_2 \neq p_3, \\ -\Delta_{3,2}\mathrm{d}p_2, & p_2 = p_3. \end{cases} \tag{A.1}$$

- For an isoenergetic transformation ($\mathrm{d}E = 0$),

$$\mathrm{d}S = \log\frac{p_3}{p_1}\left[1 - \frac{\beta_{2,3}}{\beta_{1,3}}\right] \mathrm{d}p_1, \tag{A.2}$$

where $p_1 \neq p_3$ because, by assumption, $\mathcal{V}_{\mathrm{qut}}(\rho)$ is not colinear and $\boldsymbol{v}_1 \leq \boldsymbol{v}_2 \leq \boldsymbol{v}_3$.

Therefore finite changes to the qutrit probabilities generate finite changes in E and S. □

At the same time, these non-trivial qutrit deformations will also potentially change $A(\rho)$.

Lemma A.2: Infinitesimal change in Area

Given ρ and H as in Definition A.2, the differential of area $A(\rho)$ under a qutrit deformation is $\mathrm{d}A = \partial_E A \, \mathrm{d}E + \partial_S A \, \mathrm{d}S$, where

$$\partial_x A = \frac{1}{4A_{\mathrm{qut}}} \sum_{i=1}^{3} \frac{\Delta_{k_i+1,k_i-1}}{p_{k_i}} \left[(s_{k_{(i+1)}} - s_{k_{(i-1)}})\delta_{x,E} - \Delta_{k_{(i+1)},k_{(i-1)}}\delta_{x,S} \right] \chi[\boldsymbol{v}_{k_i}], \tag{A.3}$$

for $x \in \{E, S\}$.

- The labels i run modulo 3 over points in $\mathcal{V}_{\mathrm{qut}}(\rho)$. The labels k_i run modulo n clockwise around the vertices $\mathcal{V}_{\mathrm{vert}}(\rho)$.
- A_{qut} is the area formed by the points in $\mathcal{V}_{\mathrm{qut}}(\rho)$, guaranteed to be non-zero by assuming $\mathcal{V}_{\mathrm{qut}}(\rho)$ is not colinear.
- $\chi[\boldsymbol{v}_k]$ is the characteristic function such that

$$\chi[\boldsymbol{v}_k] \coloneqq \begin{cases} 1, & \text{if } \boldsymbol{v}_k \in \mathcal{V}_{\mathrm{vert}}(\rho), \\ 0, & \text{otherwise.} \end{cases} \tag{A.4}$$

Proof. For a qutrit deformation $\mathrm{d}p_{k_i} = 0$ for $i \notin \{1, 2, 3\}$ (only the virtual qutrit probabilities vary), therefore Eq. (3.5) gives

$$\mathrm{d}A = \frac{1}{2} \sum_{i=1}^{3} \frac{\Delta_{k_i+1,k_i-1}}{p_{k_i}} \chi[\boldsymbol{v}_{k_i}] \mathrm{d}p_{k_i}. \tag{A.5}$$

We now find expressions for the $\mathrm{d}p_{k_i}$ in terms of $\mathrm{d}E$ and $\mathrm{d}S$ by defining vectors $\mathbf{dp} \coloneqq (\mathrm{d}p_{k_1}, \mathrm{d}p_{k_2}, \mathrm{d}p_{k_3})^T$ and $\mathbf{dV} \coloneqq (\mathrm{d}E, \mathrm{d}S, 0)^T$, allowing us to express the differentials of energy,

entropy and normalisation as

$$\mathbf{d}\boldsymbol{V} = M\mathbf{d}\boldsymbol{p}, \quad \text{where} \quad M \coloneqq \begin{pmatrix} \epsilon_{k_1} & \epsilon_{k_2} & \epsilon_{k_3} \\ s_{k_1} & s_{k_2} & s_{k_3} \\ 1 & 1 & 1 \end{pmatrix}. \tag{A.6}$$

The matrix M is invertible because we assume the virtual qutrit is not completely passive, therefore $\det M = 2A_{\text{qut}} > 0$. Therefore $\mathbf{d}\boldsymbol{p} = M^{-1}\mathbf{d}\boldsymbol{V}$, for which we obtain

$$\mathrm{d}p_{k_i} = \boldsymbol{e}_{k_i} \cdot \mathbf{d}\boldsymbol{p} = \frac{1}{2A_{\text{qut}}} \left[(s_{k_{(i+1)}} - s_{k_{(i-1)}})\mathrm{d}E - \Delta_{k_{(i+1)},k_{(i-1)}}\mathrm{d}S \right], \quad i \in \{1,2,3\}, \tag{A.7}$$

where $\boldsymbol{e}_{k_1} \coloneqq (1,0,0)$ etc, and substituting Eq. (A.7) into Eq. (A.5) gives the result. □

A.2 Monotonically Non-Increasing Deformations

Now that qutrit deformations have been introduced, we want to show that there exist transformations that do not increase $A(\rho)$ along activation trajectories. Specifically, we are interested in the following transformations.

Definition A.3: Monotonically non-increasing deformation

Given ρ and H as in Definition A.2, a monotonically non-increasing deformation is some infinitesimal change $A(\rho) \to A(\rho) + \mathrm{d}A = A(\rho')$ to the area of $\mathcal{V}_\infty(\rho)$, generated by varying any number of occupation probabilities $p_{k_i} \to p_{k_i} + \mathrm{d}p_{k_i}$ and $\sum_{k_i} \mathrm{d}p_{k_i} = 0$, for $i \in \{1,\ldots,m\}$, where $m \leq d$, in such a way that $\mathrm{d}A \leq 0$ for $\mathrm{d}E < 0$ or $\mathrm{d}S > 0$. Given that $\mathrm{d}A = \partial_E A\ \mathrm{d}E + \partial_S A\ \mathrm{d}S$, this requires $\partial_E A \geq 0$ and $\partial_S A \leq 0$.

To prove the existence of such transformations, we need to define various features of $\mathcal{V}(\rho)$.

Definition A.4: Upper and lower branch

Given an ensemble $\mathcal{V}(\rho)$, the upper branch $\mathcal{V}_{\text{upper}}(\rho) \subseteq \mathcal{V}_{\text{vert}}(\rho)$ is

$$\mathcal{V}_{\text{upper}}(\rho) \coloneqq \{\boldsymbol{v}_k \in \mathcal{V}_{\text{vert}}(\rho)\ :\ \boldsymbol{v}_{k-1} \leq \boldsymbol{v}_k \leq \boldsymbol{v}_{k+1}\}, \tag{A.8}$$

where the k labels are ordered clockwise mod n around $\mathcal{V}_\infty(\rho)$. The lower branch

$\mathcal{V}_{\text{lower}}(\rho) \subset \mathcal{V}_{\text{vert}}(\rho)$ is then $\mathcal{V}_{\text{lower}}(\rho) \coloneqq \mathcal{V}_{\text{vert}}(\rho) \setminus \mathcal{V}_{\text{upper}}(\rho)$, where $\setminus$ denotes the set difference.

These should be thought of as the sets of vertices along the top and bottom of $\mathcal{V}(\rho)$. Finally, proving the existence of such transformations raises technicalities involving points along the edges of $\mathcal{V}(\rho)$.

Definition A.5: Face points

The face points $\mathcal{V}^k_{\text{face}}(\rho) \subset \mathcal{V}(\rho)$ associated with a vertex point $\boldsymbol{v}_k \in \mathcal{V}_{\text{vert}}(\rho)$ are

$$\mathcal{V}^k_{\text{face}}(\rho) \coloneqq \left\{\boldsymbol{v}_j \in \mathcal{V}(\rho) \ : \ \boldsymbol{v}_j \coloneqq q_j \boldsymbol{v}_k + (1-q_j)\boldsymbol{v}_{k+1} \text{ or } \boldsymbol{v}_j \coloneqq q_j \boldsymbol{v}_k + (1-q_j)\boldsymbol{v}_{k-1}\right\} \quad \text{(A.9)}$$

for any real-valued $\{q_j\}$ satisfying $0 < q_j < 1, \forall j$.

We are now in a position to prove that there is always a monotonically non-increasing transformation for any non-Gibbsian passive state.

Lemma A.3: Existence of monotonically non-increasing deformation

Let ρ and H be defined as in Definition A.2. For any such state $\rho \in \mathcal{D}_{(E,S)}, \forall E, S$ there exists a monotonically non-increasing deformation.

Proof. Our aim is to show that we can always find a qutrit deformation of ρ such that we obtain $\partial_E A \geq 0$ and $\partial_S A \leq 0$. Inspecting Eq. (A.3), this can be achieved by choosing a virtual qutrit such that

$$\Delta_{k_i+1,k_i-1}(\boldsymbol{v}_{k_{(i+1)}} - \boldsymbol{v}_{k_{(i-1)}}) \geq \mathbf{0} \quad \text{for } i \in \{1,2,3\}. \quad \text{(A.10)}$$

Assuming that ρ is passive means that the required entropy constraint in Eq. (A.10) is guaranteed by satisfying the energy constraint, therefore the virtual qutrit is chosen such that

$$\underbrace{\Delta_{k_i+1,k_i-1}}_{\text{mod } n}\underbrace{\Delta_{k_{(i+1)},k_{(i-1)}}}_{\text{mod } 3} \geq 0 \quad \text{for } i \in \{1,2,3\}. \quad \text{(A.11)}$$

All that remains is to show that such a choice of virtual qutrit exists. Suppose for the moment that there are no face points in $\mathcal{V}(\rho)$ – the only boundary points are the vertices. Given that the constraints in Eq. (A.11) depend only upon the relative signs of the energy differences, we project onto the ϵ-axis and look at whether the vertices belong to the upper or lower branches of $\mathcal{V}(\rho)$. The assumption that H is non-degenerate ensures the mapping $\boldsymbol{v}_k \to \epsilon_k, \forall \boldsymbol{v}_k \in \mathcal{V}_{\text{vert}}$ is a bijection. The choice of points now splits into two situations:

1. *($\mathcal{V}_{\text{upper}}(\rho)$ is non-empty)* There are four cases we need to consider, shown in Figure A.1(a)-(d). Choosing the highlighted qutrit in each case, the three constraints of Eq. (A.11) are satisfied.

2. *($\mathcal{V}_{\text{upper}}(\rho)$ is empty)* By inspection, the choice highlighted in Figure A.1(e) again satisfies the conditions of Eq. (A.11).

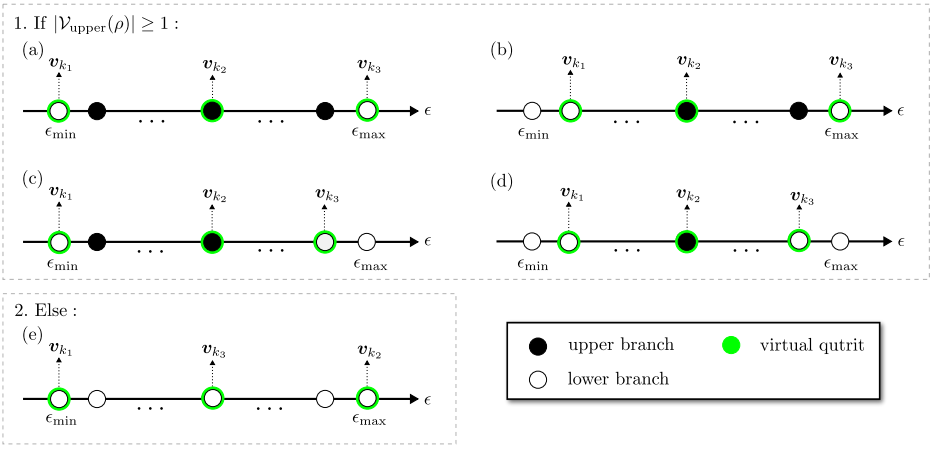


Figure A.1: Virtual qutrit selection. *Reproduced from [2].*

Face points complicate the picture, since Eq. (3.5) may not be continuous over throughout the deformation However, problematic vertices can be substituted with face or interior points to ensure the expression for the area remains continuous. If all points of the virtual qutrit must be swapped for non-vertex points, the transformation gives $\mathrm{d}A = 0$, which is still monotonically non-increasing.

Therefore, given any passive state that is not completely passive, there is some choice of virtual qutrit on which we can perform a monotonically non-increasing deformation. □

A.3 Proving monotonicity

We are now onto the final strait of the proof of Theorem 3.1, which we restate here.

Theorem A.1: Restatement of Theorem

For a d-dimensional system with non-degenerate Hamiltonian H, the geometric athermality $\mathcal{A}$ is monotonically non-increasing along all activation trajectories.

Proof. The theorem is trivially true for $d = 1$ and 2 because $A(\rho) = 0$ for all ρ in those settings. Now consider a $(d \geq 3)$-dimensional quantum system with non-degenerate Hamiltonian H, prepared in a passive, full rank, non-Gibbsian state $\rho \in \mathcal{D}_{(E,S)}$. For simplicity, assume that ρ is such that $A(\rho) = \mathcal{A}(E, S)$.

By Lemma A.3, there exists an isentropic monotonically non-increasing deformation that takes us to some new state $\rho' \in \mathcal{D}_{(E-\mathrm{d}E,S)}$, such that the area $\mathcal{A}(E, S) \to \mathcal{A}(E, S) - \partial_E A \, \mathrm{d}E \coloneqq A(\rho')$. By definition, the new area must be lower bounded by $\mathcal{A}(E - \mathrm{d}E, S)$, and because $\partial_E A \geq 0$, this implies that $\mathcal{A}(E, S) \geq A(\rho') \geq \mathcal{A}(E - \mathrm{d}E, S)$, from which we deduce

$$\partial_E \mathcal{A}(E, S) \, \mathrm{d}E = \mathcal{A}(E, S) - \mathcal{A}(E - \mathrm{d}E, S) \geq 0, \tag{A.12}$$

and thus $\partial_E \mathcal{A}(E, S) \geq 0$.

We can do the same with an isoenergetic monotonically non-increasing deformation to the state $\tilde{\rho} \in \mathcal{D}_{(E,S+\mathrm{d}S)}$, such that $\mathcal{A}(E, S) \to \mathcal{A}(E, S) + \partial_S A \, \mathrm{d}S \coloneqq A(\tilde{\rho})$. Since $\partial_S A < 0$ for such a transformation $\mathcal{A}(E, S) \geq A(\tilde{\rho}) \geq \mathcal{A}(E, S + \mathrm{d}S)$, therefore

$$\partial_S \mathcal{A}(E, S) \, \mathrm{d}S = \mathcal{A}(E, S + \mathrm{d}S) - \mathcal{A}(E, S) \leq 0, \tag{A.13}$$

which implies $\partial_S \mathcal{A}(E, S) \leq 0$.

This implies that $\partial_E \mathcal{A}(E,S) \geq 0$ and $\partial_S \mathcal{A}(E,S) \leq 0$, for all (E,S) strictly below the thermal curve, therefore $\nabla_{\boldsymbol{u}} \mathcal{A}(E,S) \leq 0$, whenever $\boldsymbol{u}$ has components which satisfy $u_E \leq 0$ and $u_S \geq 0$ (an activation trajectory). Finally, the statement trivially holds for Gibbs states, because they have $A(\gamma_\beta) = 0$ and can only be the endpoints of activation trajectories. $\square$

Appendix B

Group Orbit 'Shapes' for Two Qubits under SU(2)

In Chapter 5, the possible isotropy subgroups for two-qubit quantum states under a tensor product SU(2) group action were enumerated. For some simple cases, the isotropy subgroup can be calculated directly, as demonstrated for the Bell states. However, for most states, the H-twirling projection operators provide a simpler method for identifying states with certain residual symmetries. Nonetheless, we must further narrow down the possible subgroups of SU(2) to be checked. The following lemma restricts the possible isotropy subgroups for two qubits.

Lemma B.1: Restricting the Two-Qubit Isotropy Subgroups

For a two-qubit state ρ under an SU(2) tensor product group action, the cyclic subgroups of Iso(ρ) must be either $\mathbb{Z}_2$, $\mathbb{Z}_4$ or U(1).

Proof. The condition for membership of Iso(ρ) is $\mathcal{U}_g(\rho) = \rho$, or equivalently

$$U(g) \otimes U(g) \otimes U^*(g) \otimes U^*(g) \;\; |\mathrm{vec}(\rho)\rangle = |\mathrm{vec}(\rho)\rangle \tag{B.1}$$

by vectorisation [84], therefore $|\mathrm{vec}(\rho)\rangle$ must belong to $\ker[\mathrm{U(g)} \otimes \mathrm{U(g)} \otimes \mathrm{U^*(g)} \otimes \mathrm{U^*(g)} - \mathbb{I}]$.

All groups contain cyclic subgroups $\{g^n\}$ constructed from repeated application of a single generator $g \in G$, and there must be some basis in which the representation $U(g)$ is diagonal, i.e. $U(g) = e^{i\theta} |0\rangle \langle 0| + e^{-i\theta} |1\rangle \langle 1|$. For a cyclic subgroup of $\mathrm{Iso}(\rho)$,

$$\begin{aligned} &U(g) \otimes U(g) \otimes U^*(g) \otimes U^*(g) - \mathbb{I} \\ &\quad = (e^{2i\theta} - 1)\big(|1\rangle \langle 1| + |2\rangle \langle 2| + |7\rangle \langle 7| + |11\rangle \langle 11| \big) + (e^{4i\theta} - 1) |3\rangle \langle 3| + (e^{-4i\theta} - 1) |12\rangle \langle 12| \\ &\qquad + (e^{-2i\theta} - 1)\big(|4\rangle \langle 4| + |8\rangle \langle 8| + |13\rangle \langle 13| + |14\rangle \langle 14| \big) \,, \end{aligned} \tag{B.2}$$

assuming without loss of generality that $U(g)$ is diagonal in the computational basis (with for example $|2\rangle = |0010\rangle$).

- Vectors in $\mathrm{span}\big(|0\rangle, |5\rangle, |6\rangle, |9\rangle, |10\rangle, |15\rangle \big)$ are in the kernel for all θ, corresponding to a U(1) subgroup.
- Vectors in $\mathrm{span}\big(|1\rangle, |2\rangle, |4\rangle, |7\rangle, |8\rangle, |11\rangle, |13\rangle, |14\rangle \big)$ also belong to the kernel when $\theta = 0, \pi$, corresponding to a $\mathbb{Z}_2$ subgroup.
- Vectors in $\mathrm{span}\big(|3\rangle, |12\rangle \big)$ are in the kernel when $\theta = 0, \pi/2, \pi, 3\pi/2$, corresponding to a $\mathbb{Z}_4$ subgroup.

Therefore any cyclic subgroup of $\mathrm{Iso}(\rho)$ must be either $\mathbb{Z}_2$, $\mathbb{Z}_2$ or $U(1)$. □

This result restricts the possible isotropy subgroups for two qubits, much as the crystallographic restriction theorem constrains the allowed crystal lattices [55]. We can therefore use projectors $\mathcal{P}_H$ to isolate states in $\hat{C}(H)$, from which we can infer the sets $C(H)$. We now present the calculations for two-qubit states under SU(2).

B.1 Enumeration of the Isotropy Subgroups

B.1.1 $\mathbb{Z}_2 \prec \mathrm{Iso}(\rho)$

All two-qubit states are symmetric under $\mathbb{Z}_2 = \{\pm\mathbb{I}\}$, as confirmed by $\mathcal{P}_{\mathbb{Z}_2}(\rho) = \frac{1}{2}[\, \mathcal{U}_{\mathbb{I}}(\rho) + \mathcal{U}_{-\mathbb{I}}(\rho) \,] = \rho$.

B.1.2 $\mathbb{Z}_4 \prec \mathrm{Iso}(\rho)$

The next possible isotropy subgroup is $\mathbb{Z}_4 = \{\pm\mathbb{I}, \pm i\boldsymbol{r}\cdot\sigma\}$. We use the identities $(\boldsymbol{r}\cdot\sigma)(\boldsymbol{v}\cdot\sigma) = (\boldsymbol{r}\cdot\boldsymbol{v})\mathbb{I} + i(\boldsymbol{r}\times\boldsymbol{v})\cdot\sigma$ and $(\boldsymbol{r}\cdot\sigma)(\boldsymbol{v}\cdot\sigma)(\boldsymbol{r}\cdot\sigma) = 2(\boldsymbol{r}\cdot\boldsymbol{v})(\boldsymbol{r}\cdot\sigma) - \boldsymbol{v}\cdot\sigma$ to calculate $\mathcal{P}_{\mathbb{Z}_4}(\rho)$. The local Bloch vector terms give

$$\mathcal{P}_{\mathbb{Z}_4}(\boldsymbol{a}\cdot\sigma\otimes\mathbb{I}) = \frac{1}{2}[\boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + (\boldsymbol{r}\cdot\sigma)(\boldsymbol{a}\cdot\sigma)(\boldsymbol{r}\cdot\sigma)\otimes\mathbb{I}] = (\boldsymbol{r}\cdot\boldsymbol{a})(\boldsymbol{r}\cdot\sigma)\otimes\mathbb{I} \tag{B.3}$$

and similarly $\mathcal{P}_{\mathbb{Z}_4}(\mathbb{I}\otimes\boldsymbol{b}\cdot\sigma) = (\boldsymbol{r}\cdot\boldsymbol{b})(\mathbb{I}\otimes\boldsymbol{r}\cdot\sigma)$. We express the correlation terms as $\sum_{i,j=1}^{3} T_{ij}\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma$ where $\{\boldsymbol{c}_i\}$ is an orthonormal basis for $\mathbb{R}^3$, and choose $\boldsymbol{c}_1 = \boldsymbol{r}$. Clearly $\mathcal{P}_{\mathbb{Z}_4}(\boldsymbol{c}_1\cdot\sigma\otimes\boldsymbol{c}_1\cdot\sigma) = \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma$. The identity

$$\begin{aligned}(\boldsymbol{r}\cdot\sigma)(\boldsymbol{u}\cdot\sigma)(\boldsymbol{r}\cdot\sigma)\otimes(\boldsymbol{r}\cdot\sigma)(\boldsymbol{v}\cdot\sigma)(\boldsymbol{r}\cdot\sigma) &= 4(\boldsymbol{r}\cdot\boldsymbol{u})(\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma - 2(\boldsymbol{r}\cdot\boldsymbol{u})\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{v}\cdot\sigma \\ &\quad - 2(\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{u}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \boldsymbol{u}\cdot\sigma\otimes\boldsymbol{v}\cdot\sigma.\end{aligned} \tag{B.4}$$

has the consequence that $\mathcal{P}_{\mathbb{Z}_4}(\boldsymbol{c}_1\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma) = \mathcal{P}_{\mathbb{Z}_4}(\boldsymbol{c}_j\cdot\sigma\otimes\boldsymbol{c}_1\cdot\sigma) = 0$ for $j \neq 1$. Finally, when $i, j \neq 1$, $\mathcal{P}_{\mathbb{Z}_4}(\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma) = \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma$. Therefore

$$\mathcal{P}_{\mathbb{Z}_4}(\rho) = \frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I} + (\boldsymbol{r}\cdot\boldsymbol{a})(\boldsymbol{r}\cdot\sigma)\otimes\mathbb{I} + (\boldsymbol{r}\cdot\boldsymbol{b})(\mathbb{I}\otimes\boldsymbol{r}\cdot\sigma) + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right], \tag{B.5}$$

with $\boldsymbol{c}_1 = \boldsymbol{d}_1 = \boldsymbol{r}$.

B.1.3 $\mathrm{Iso}(\rho) = \mathbb{Z}_2$

The states with $\mathrm{Iso}(\rho) = \mathbb{Z}_2$ correspond to those eliminated by $\mathcal{P}_{\mathbb{Z}_4}$. These are:

- states of the form

$$\rho = \frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I} + \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes\boldsymbol{b}\cdot\sigma + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right], \tag{B.6}$$

  where $\boldsymbol{c}_i \neq \boldsymbol{d}_i$ for $i = 1, 2, 3$.

- states of the form

$$\rho = \frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I} + \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes\boldsymbol{b}\cdot\sigma + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right], \tag{B.7}$$

where $\boldsymbol{c}_1 = \boldsymbol{d}_1 = \boldsymbol{r}$ but $\boldsymbol{a}$ and/or $\boldsymbol{b}$ are not parallel to $\boldsymbol{r}$ (there are no $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\boldsymbol{a} = \gamma_1 \boldsymbol{r}$ and $\boldsymbol{b} = \gamma_2 \boldsymbol{r}$).

- states of the form

$$\rho = \frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I} + \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes\boldsymbol{b}\cdot\sigma + \sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\ \right], \tag{B.8}$$

where the local Bloch vectors are not parallel to each other ($\boldsymbol{a} \neq \gamma\boldsymbol{b}$ for any $\gamma \in \mathbb{R}$).

B.1.4 $U(1) \prec \mathrm{Iso}(\rho)$

Consider $U(1) = \{e^{i\boldsymbol{r}\cdot\sigma} :\ 0 \leq \theta < 2\pi\}$. This acts on a Bloch vector as

$$e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{u}\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma} = \cos 2\theta\ \boldsymbol{u}\cdot\sigma + \sin 2\theta\ (\boldsymbol{u}\times\boldsymbol{r})\cdot\sigma + 2\sin^2\theta\ (\boldsymbol{r}\cdot\boldsymbol{u})\ \boldsymbol{r}\cdot\sigma, \tag{B.9}$$

therefore, for the local Bloch vector terms,

$$\begin{aligned}\mathcal{P}_{U(1)}(\boldsymbol{a}\cdot\sigma\otimes\mathbb{I}) &= \frac{1}{2\pi}\int_0^{2\pi} d\theta\ e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{a}\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma}\otimes\mathbb{I} &\text{(B.10)}\\ &= \frac{1}{2\pi}\int_0^{2\pi} d\theta\ \cos 2\theta\ \boldsymbol{a}\cdot\sigma\otimes\mathbb{I} + \sin 2\theta\ (\boldsymbol{a}\times\boldsymbol{r})\cdot\sigma\otimes\mathbb{I} + 2\sin^2\theta\ (\boldsymbol{r}\cdot\boldsymbol{a})\ \boldsymbol{r}\cdot\sigma\otimes\mathbb{I}\\ &= (\boldsymbol{r}\cdot\boldsymbol{a})\ \boldsymbol{r}\cdot\sigma\otimes\mathbb{I} &\text{(B.11)}\end{aligned}$$

and similarly $P_{U(1)}(\mathbb{I}\otimes\boldsymbol{b}\cdot\sigma) = (\boldsymbol{r}\cdot\boldsymbol{b})\ \mathbb{I}\otimes\boldsymbol{r}\cdot\sigma$. On the correlation terms,

$$\begin{aligned}&e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{u}\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma}\otimes e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{v}\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma}\\ &= \cos^2 2\theta\ \boldsymbol{u}\cdot\sigma\otimes\boldsymbol{v}\cdot\sigma + 4\sin^4\theta(\boldsymbol{r}\cdot\boldsymbol{u})(\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + 2\cos 2\theta\sin 2\theta\ [\boldsymbol{u}\cdot\sigma\otimes(\boldsymbol{v}\times\boldsymbol{r})\cdot\sigma\\ &\quad + (\boldsymbol{u}\times\boldsymbol{r})\cdot\sigma\otimes\boldsymbol{v}\cdot\sigma] + 2\cos 2\theta\sin^2\theta[(\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{u}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + (\boldsymbol{r}\cdot\boldsymbol{u})\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{v}\cdot\sigma]\\ &\quad + 2\sin 2\theta\sin^2\theta\ [(\boldsymbol{r}\cdot\boldsymbol{v})(\boldsymbol{u}\times\boldsymbol{r})\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + (\boldsymbol{r}\cdot\boldsymbol{u})\boldsymbol{r}\cdot\sigma\otimes(\boldsymbol{v}\times\boldsymbol{r})\cdot\sigma]\\ &\quad + \sin^2 2\theta(\boldsymbol{u}\times\boldsymbol{r})\cdot\sigma\otimes(\boldsymbol{v}\times\boldsymbol{r})\cdot\sigma\ ,\end{aligned} \tag{B.12}$$

which gives

$$\mathcal{P}_{U(1)}(\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma)=\frac{1}{2\pi}\int_0^{2\pi}d\theta\; e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{c}_i\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma}\otimes e^{i\theta\boldsymbol{r}\cdot\sigma}(\boldsymbol{c}_j\cdot\sigma)e^{-i\theta\boldsymbol{r}\cdot\sigma} \tag{B.13}$$

$$=\frac{1}{2}[\;\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma-(\boldsymbol{r}\cdot\boldsymbol{c}_j)\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma-(\boldsymbol{r}\cdot\boldsymbol{c}_i)\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma+(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma\otimes(\boldsymbol{c}_j\times\boldsymbol{r})\cdot\sigma$$

$$+3(\boldsymbol{r}\cdot\boldsymbol{c}_i)(\boldsymbol{r}\cdot\boldsymbol{c}_j)\boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma\;] \tag{B.14}$$

$$=\frac{1}{2}[\;\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma-\delta_{1j}\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_1\cdot\sigma-\delta_{1i}\boldsymbol{c}_1\cdot\sigma\otimes\boldsymbol{c}_j\cdot\sigma+(\boldsymbol{c}_i\times\boldsymbol{c}_1)\cdot\sigma\otimes(\boldsymbol{c}_j\times\boldsymbol{c}_1)\cdot\sigma$$

$$+3\;\delta_{1i}\delta_{1j}\;\boldsymbol{c}_1\cdot\sigma\otimes\boldsymbol{c}_1\cdot\sigma]\;. \tag{B.15}$$

In summary,

$$\mathcal{P}_{U(1)}(\rho)=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+(\boldsymbol{r}\cdot\boldsymbol{a})\;\boldsymbol{r}\cdot\sigma\otimes\mathbb{I}+(\boldsymbol{r}\cdot\boldsymbol{b})\;\mathbb{I}\otimes\boldsymbol{r}\cdot\sigma+\tau\sum_{i=1}^{3}\;\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right] \tag{B.16}$$

where $\boldsymbol{c}_1=\boldsymbol{d}_1=\boldsymbol{r}$ and $\tau_2=\tau_3$.

B.1.5 $K_2\prec\mathrm{Iso}(\rho)$

Consider $\mathrm{K}_2=\{\pm\mathbb{I},\pm i\boldsymbol{r}_1\cdot\sigma,\pm i\boldsymbol{r}_2\cdot\sigma,\pm i\boldsymbol{r}_3\cdot\sigma\}$ where $\{\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3\}$ forms an orthonormal basis of $\mathbb{R}^3$, and we can use that same basis for the correlation terms, such that

$$\rho=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+\boldsymbol{a}\cdot\sigma\otimes\mathbb{I}+\mathbb{I}\otimes\boldsymbol{b}\cdot\sigma+\sum_{i,j=1}^{3}T_{ij}\;\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma\;\right]. \tag{B.17}$$

Acting on a local Bloch vector $\boldsymbol{a}\cdot\sigma=\sum_{i=1}^{3}(\boldsymbol{r}_i\cdot\boldsymbol{a})\;\boldsymbol{r}_i\cdot\sigma$,

$$\mathcal{P}_{K_2}(\boldsymbol{a}\cdot\sigma\otimes\mathbb{I})=\frac{1}{4}\left[\boldsymbol{a}\cdot\sigma\otimes\mathbb{I}+\sum_{i=1}^{3}(2(\boldsymbol{r}_i\cdot\boldsymbol{a})\boldsymbol{r}_i\cdot\sigma-\boldsymbol{a}\cdot\sigma)\otimes\mathbb{I}\right] \tag{B.18}$$

$$=\frac{1}{2}\left[\;\sum_{i=1}^{3}(\boldsymbol{r}_i\cdot\boldsymbol{a})\;\boldsymbol{r}_i\cdot\sigma\otimes\mathbb{I}-\boldsymbol{a}\cdot\sigma\otimes\mathbb{I}\;\right]=0, \tag{B.19}$$

and similarly $\mathcal{P}_{K_2}(\mathbb{I}\otimes \boldsymbol{b}\cdot\sigma)=0$. On the correlation terms,

$$\mathcal{P}_{K_2}(\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma)=\frac{1}{4}\left[\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma+\sum_{k=1}^{3}(\boldsymbol{r}_k\cdot\sigma)(\boldsymbol{r}_i\cdot\sigma)(\boldsymbol{r}_k\cdot\sigma)\otimes(\boldsymbol{r}_k\cdot\sigma)(\boldsymbol{r}_j\cdot\sigma)(\boldsymbol{r}_k\cdot\sigma)\right] \tag{B.20}$$

$$=\frac{1}{4}\left[\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma+\sum_{k=1}^{3}\Big(4(\boldsymbol{r}_k\cdot\boldsymbol{r}_i)(\boldsymbol{r}_k\cdot\boldsymbol{r}_j)\boldsymbol{r}_k\cdot\sigma\otimes\boldsymbol{r}_k\cdot\sigma+\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma \right.$$
$$\left.-2(\boldsymbol{r}_k\cdot\boldsymbol{r}_i)\boldsymbol{r}_k\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma-2(\boldsymbol{r}_k\cdot\boldsymbol{r}_j)\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_k\cdot\sigma\Big)\right] \tag{B.21}$$

$$=\frac{1}{4}\left[\sum_{k=1}^{3}(4\delta_{ik}\delta_{jk}\boldsymbol{r}_k\cdot\sigma\otimes\boldsymbol{r}_k\cdot\sigma-2\delta_{ik}\boldsymbol{r}_k\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma-2\delta_{jk}\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_k\cdot\sigma)+4\boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma\right]$$

$$=\delta_{ij}\ \boldsymbol{r}_j\cdot\sigma\otimes\boldsymbol{r}_j\cdot\sigma\ , \tag{B.22}$$

therefore

$$\mathcal{P}_{K_2}(\rho)=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+\sum_{i=1}^{3}\tau_i\ \boldsymbol{r}_i\cdot\sigma\otimes\boldsymbol{r}_i\cdot\sigma\right]\ . \tag{B.23}$$

B.1.6 $\mathrm{Iso}(\rho)=\mathbb{Z}_4$

The states with $\mathrm{Iso}(\rho)=\mathbb{Z}_4$ are those in $\hat{C}(\mathbb{Z}_4)$ but not in $\hat{C}(\mathrm{K}_2)$ or $\hat{C}(\mathrm{U}(1))$. These are:

- states of the form

$$\rho=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+\tau_1\ \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma+\sum_{i=2}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right] \tag{B.24}$$

  where $\tau_2\neq\tau_3$ and $\boldsymbol{c}_i\neq\boldsymbol{d}_i$ for $i=2,3$.

- states of the form

$$\rho=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+a\ \boldsymbol{r}\cdot\sigma\otimes\mathbb{I}+b\ \mathbb{I}\otimes\boldsymbol{r}\cdot\sigma+\sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{d}_i\cdot\sigma\right] \tag{B.25}$$

  where $\boldsymbol{c}_1=\boldsymbol{d}_1=\boldsymbol{r}$ but $\boldsymbol{c}_i\neq\boldsymbol{d}_i$ for $i=2,3$, and $\tau_2\neq\tau_3$.

- states of the form

$$\rho=\frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I}+a\boldsymbol{c}_1\cdot\sigma\otimes\mathbb{I}+b\mathbb{I}\otimes\boldsymbol{c}_1\cdot\sigma+\sum_{i=1}^{3}\tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma\right] \tag{B.26}$$

  where $\tau_1=\tau_2\neq\tau_3$.

B.1.7 $K_\infty \prec \mathrm{Iso}(\rho)$

Consider $K_\infty = \{(i\ \boldsymbol{r}_j \cdot \sigma)^\alpha e^{i\theta\ \boldsymbol{r}_k \cdot \sigma} :\ 0 \leq \theta < 2\pi,\ \alpha = 0,1,2,3\}$, where $j \neq k$. Acting with $\mathcal{P}_{K_\infty}$ on a two-qubit state gives

$$\mathcal{P}_{K_\infty}(\rho) = \mathcal{P}_{U(1)} \circ \mathcal{P}_{K_2}(\rho) = \mathcal{P}_{U(1)} \left[\frac{1}{4} \left(\mathbb{I} \otimes \mathbb{I} + \sum_{i=1}^{3} \tau_i\ \boldsymbol{r}_i \cdot \sigma \otimes \boldsymbol{r}_i \cdot \sigma \right) \right] \tag{B.27}$$

$$= \frac{1}{4} \left(\mathbb{I} \otimes \mathbb{I} + \frac{1}{2} \left[\sum_{i \neq k} \tau_i \right] \left[\sum_{i \neq k} \boldsymbol{r}_i \cdot \sigma \otimes \boldsymbol{r}_i \cdot \sigma \right] + \tau_k\ \boldsymbol{r}_k \cdot \sigma \otimes \boldsymbol{r}_k \cdot \sigma \right) \tag{B.28}$$

B.1.8 $\mathrm{Iso}(\rho) = \mathrm{K}_2$

These are the (rotated) T-states, namely

$$\rho = \frac{1}{4} \left[\mathbb{I} \otimes \mathbb{I} + \sum_i \tau_i\ \boldsymbol{c}_i \cdot \sigma \otimes \boldsymbol{c}_i \cdot \sigma \right], \tag{B.29}$$

with $\tau_1 \neq \tau_2 \neq \tau_3$.

B.1.9 $\mathrm{Iso}(\rho) = \mathrm{U}(1)$

These states take the form

$$\rho = \frac{1}{4} \left[\mathbb{I} \otimes \mathbb{I} + a\boldsymbol{r} \cdot \sigma \otimes \mathbb{I} + b\mathbb{I} \otimes \boldsymbol{r} \cdot \sigma + \tau_1\ \boldsymbol{r} \cdot \sigma \otimes \boldsymbol{r} \cdot \sigma + \tau \sum_{i=2}^{3} \boldsymbol{c}_i \cdot \sigma \otimes \boldsymbol{d}_i \cdot \sigma \right] \tag{B.30}$$

where $\boldsymbol{c}_1 = \boldsymbol{d}_1 = \boldsymbol{r}$ but $\boldsymbol{c}_i \neq \boldsymbol{d}_i$ for $i = 2,3$, and $\tau_2 = \tau_3$.

B.1.10 $\mathrm{Iso}(\rho) = \mathrm{SU}(2)$

Applying $\mathcal{P}_{SU(2)}$ to a general two-qubit state,

$$\mathcal{P}_{SU(2)}(\rho) = \mathcal{P}_{SU(2)} \circ \mathcal{P}_{K_\infty}(\rho) = \frac{1}{4} \left(\mathbb{I} \otimes \mathbb{I} + \sum_{i=1}^{3} \tau_i\ \mathcal{P}_{SU(2)}\ [\boldsymbol{c}_i \cdot \sigma \otimes \boldsymbol{c}_i \cdot \sigma] \right)\ , \tag{B.31}$$

which outputs the two-qubit symmetric states. A general $SU(2)$ group transformation is $U(g) = e^{i\varphi \boldsymbol{r}\cdot\sigma}$, where $\boldsymbol{r} = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$. Considering each term individually,

$$\begin{aligned}
&\mathcal{P}_{SU(2)}(\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma)\\
&= \frac{1}{2\pi^2}\int_0^{2\pi} d\theta \int_0^{\pi} d\phi\ \sin\phi \int_0^{\pi} d\varphi\ \sin^2\varphi\ \left(e^{i\varphi\boldsymbol{r}\cdot\sigma}\boldsymbol{c}_i\cdot\sigma e^{-i\varphi\boldsymbol{r}\cdot\sigma}\otimes e^{i\varphi\boldsymbol{r}\cdot\sigma}\boldsymbol{c}_i\cdot\sigma e^{-i\varphi\boldsymbol{r}\cdot\sigma}\right) \qquad \text{(B.32)}\\
&= \frac{1}{2\pi^2}\int_0^{2\pi} d\theta \int_0^{\pi} d\phi\ \sin\phi \int_0^{\pi} d\varphi\ \sin^2\varphi\cos^2 2\varphi\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma\\
&\qquad + 2\sin^2\varphi\cos 2\varphi\sin 2\varphi\ [\boldsymbol{c}_i\cdot\sigma\otimes(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma + (\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma]\\
&\qquad + 2\cos 2\varphi\sin^4\varphi\ (\boldsymbol{r}\cdot\boldsymbol{c}_i)\ [\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma]\\
&\qquad + 2\sin 2\varphi\sin^4\varphi\ (\boldsymbol{r}\cdot\boldsymbol{c}_i)\ [(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \boldsymbol{r}\cdot\sigma\otimes(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma]\\
&\qquad + 4\sin^6\varphi\ (\boldsymbol{r}\cdot\boldsymbol{c}_i)^2\ \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \sin^2\varphi\sin^2 2\varphi\ (\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma\otimes(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma \qquad \text{(B.33)}\\
&= \frac{1}{8\pi}\int_0^{2\pi} d\theta \int_0^{\pi} d\phi\ \sin\phi\ \Big(\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma - 2(\boldsymbol{r}\cdot\boldsymbol{c}_i)\ [\boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma + \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma]\\
&\qquad\qquad + (\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma\otimes(\boldsymbol{c}_i\times\boldsymbol{r})\cdot\sigma + 5(\boldsymbol{r}\cdot\boldsymbol{c}_i)^2\ \boldsymbol{r}\cdot\sigma\otimes\boldsymbol{r}\cdot\sigma\Big)\ . \qquad \text{(B.34)}
\end{aligned}$$

Integrating over θ and ϕ, we find

$$\mathcal{P}_{SU(2)}(\rho) = \frac{1}{4}\Big[\ \mathbb{I}\otimes\mathbb{I} + \tau\sum_{i=1}^{3}\sigma_i\otimes\sigma_i\ \Big]. \qquad \text{(B.35)}$$

B.1.11 $\mathrm{Iso}(\rho) = \mathrm{K}_\infty$

These are states in $\hat{C}(K_\infty)$ but not in $C\big(SU(2)\big)$, namely

$$\rho = \frac{1}{4}\left[\mathbb{I}\otimes\mathbb{I} + \sum_i \tau_i\ \boldsymbol{c}_i\cdot\sigma\otimes\boldsymbol{c}_i\cdot\sigma\right] \qquad \text{(B.36)}$$

where $\tau_1 = \tau_2 \neq \tau_3$.

Bibliography

[1] Thomas Hebdige and David Jennings. On the classification of two-qubit group orbits and the use of coarse-grained 'shape' as a superselection property. Quantum, 3:119, February 2019.

[2] Nikolaos Koukoulekidis, Rhea Alexander, Thomas Hebdige, and David Jennings. The geometry of passivity for quantum systems and a novel elementary derivation of the Gibbs state, 2019.

[3] M.A. Nielsen and I.L. Chuang. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press, 2010.

[4] Mark M. Wilde. Quantum Information Theory. Cambridge University Press, 2013.

[5] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. Classical and Quantum Computation. American Mathematical Society, Boston, MA, USA, 2002.

[6] J. Gemmer, M. Michel, and G. Mahler. Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems. Lecture Notes in Physics. Springer Berlin Heidelberg, 2009.

[7] F. Binder, L.A. Correa, C. Gogolin, J. Anders, and G. Adesso. Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions. Fundamental Theories of Physics. Springer International Publishing, 2019.

[8] R. Augusiak, F. M. Cucchietti, and M. Lewenstein. Many-Body Physics from a Quantum Information Perspective, pages 245–294. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[9] Patrick Hayden. It from qubit: Simons collaboration on quantum fields, gravity, and information. http://web.stanford.edu/ phayden/simons/simons-proposal.pdf.

[10] Ryszard Horodecki, Pawel Horodecki, Michal Horodecki, and Karol Horodecki. Quantum entanglement. Rev. Mod. Phys., 81:865–942, Jun 2009.

[11] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio. Colloquium: Quantum coherence as a resource. Rev. Mod. Phys., 89:041003, Oct 2017.

[12] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reference frames, superselection rules, and quantum information. Rev. Mod. Phys., 79:555–609, Apr 2007.

[13] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. New Journal of Physics, 16(1):013009, 2014.

[14] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. A mathematical theory of resources. Information and Computation, 250(Supplement C):59 – 86, 2016.

[15] Eric Chitambar and Gilad Gour. Quantum resource theories. Rev. Mod. Phys., 91:025001, Apr 2019.

[16] O. Penrose. Foundations of statistical mechanics: a deductive treatment. International series of monographs in natural philosophy. Pergamon Press, 1969.

[17] E. T. Jaynes. Information theory and statistical mechanics. Phys. Rev., 106:620–630, May 1957.

[18] E. T. Jaynes. Information theory and statistical mechanics. ii. Phys. Rev., 108:171–190, Oct 1957.

[19] Christian Gogolin and Jens Eisert. Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems. Reports on Progress in Physics, 79(5):056001, May 2016.

[20] M A Cazalilla and M Rigol. Focus on dynamics and thermalization in isolated quantum many-body systems. New Journal of Physics, 12(5):055006, may 2010.

[21] Anatoli Polkovnikov, Krishnendu Sengupta, Alessandro Silva, and Mukund Vengalattore. Colloquium: Nonequilibrium dynamics of closed interacting quantum systems. Rev. Mod. Phys., 83:863–883, Aug 2011.

[22] W. Pusz and S. L. Woronowicz. Passive states and KMS states for general quantum systems. Comm. Math. Phys., 58(3):273–290, 1978.

[23] A. Lenard. Thermodynamical proof of the Gibbs formula for elementary quantum systems. Journal of Statistical Physics, 19(6):575–586, Dec 1978.

[24] Paul Skrzypczyk, Ralph Silva, and Nicolas Brunner. Passivity, complete passivity, and virtual temperatures. Phys. Rev. E, 91:052133, May 2015.

[25] O. Bratteli and D.W. Robinson. Operator Algebras and Quantum Statistical Mechanics: Equilibrium States. Models in Quantum Statistical Mechanics. Theoretical and Mathematical Physics. Springer Berlin Heidelberg, 2003.

[26] Martí Perarnau-Llobet, Karen V. Hovhannisyan, Marcus Huber, Paul Skrzypczyk, Jordi Tura, and Antonio Acín. Most energetic passive states. Phys. Rev. E, 92:042147, Oct 2015.

[27] Carlo Sparaciari, David Jennings, and Jonathan Oppenheim. Energetic instability of passive states in thermodynamics. Nature Communications, 8, Dec 2017.

[28] Raffaele Salvia and Vittorio Giovannetti. Energy upper bound for structurally stable n-passive states, 2019.

[29] Álvaro M. Alhambra, Georgios Styliaris, Nayeli A. Rodríguez-Briones, Jamie Sikora, and Eduardo Martín-Martínez. Fundamental limitations to local energy extraction in quantum systems. Phys. Rev. Lett., 123:190601, Nov 2019.

[30] Roger Balian and N.L. Balazs. Equiprobability, inference, and entropy in quantum theory. Annals of Physics, 179(1):97 – 144, 1987.

[31] Matteo Lostaglio, David Jennings, and Terry Rudolph. Thermodynamic resource theories, non-commutativity and maximum entropy principles. New Journal of Physics, 19(4):043008, apr 2017.

[32] Nicole Yunger Halpern, Philippe Faist, Jonathan Oppenheim, and Andreas Winter. Microcanonical and resource-theoretic derivations of the thermal state of a quantum system with noncommuting charges. Nature Communications, 7(1):12051, 2016.

[33] Yelena Guryanova, Sandu Popescu, Anthony J. Short, Ralph Silva, and Paul Skrzypczyk. Thermodynamics of quantum systems with multiple conserved quantities. Nature Communications, 7(1):12049, 2016.

[34] Paul Boes, Henrik Wilming, Jens Eisert, and Rodrigo Gallego. Statistical ensembles without typicality. Nature Communications, 9(1):1022, 2018.

[35] Manabendra Nath Bera, Arnau Riera, Maciej Lewenstein, Zahra Baghali Khanian, and Andreas Winter. Thermodynamics as a Consequence of Information Conservation. Quantum, 3:121, February 2019.

[36] Raam Uzdin and Saar Rahav. Global passivity in microscopic thermodynamics. Phys. Rev. X, 8:021064, Jun 2018.

[37] Raam Uzdin and Saar Rahav. The passivity deformation approach for the thermodynamics of isolated quantum setups. arXiv:1912.07922, 2019.

[38] Eric G Brown, Nicolai Friis, and Marcus Huber. Passivity and practical work extraction using Gaussian operations. New Journal of Physics, 18(11):113028, nov 2016.

[39] M. Planck and A. Ogg. Treatise on Thermodynamics. Dover books on physics. Dover Publications, 1945.

[40] Leon Mirsky. On the trace of matrix products. Mathematische Nachrichten, 20(36):171–174, 1959.

[41] G. H. Hardy, J. E. Littlewood, and George Polya. Inequalities. Cambridge, UK : Cambridge University Press, second edition ; first paperback edition edition, 1988.

[42] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. Inequalities: Theory of Majorization and Its Applications. Springer-Verlag New York, second edition edition, 2011.

[43] Rolf Schneider. Convex Bodies: The Brunn-Minkowski Theory. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2013.

[44] Terence Tao and Van H. Vu. Additive combinatorics., volume 105. Cambridge University Press, 2006.

[45] Wayne D. Blizard. Multiset theory. Notre Dame J. Formal Logic, 30(1):36–66, 12 1988.

[46] Rudolf Haag, Daniel Kastler, and Ewa B. Trych-Pohlmeyer. Stability and equilibrium states. Communications in Mathematical Physics, 38(3):173–193, Sep 1974.

[47] Carlo Sparaciari, Jonathan Oppenheim, and Tobias Fritz. Resource theory for work and heat. Phys. Rev. A, 96:052112, Nov 2017.

[48] H.B. Callen. Thermodynamics and an Introduction to Thermostatistics. Wiley, 1985.

[49] Robert Alicki and Mark Fannes. Entanglement boost for extractable work from ensembles of quantum batteries. Phys. Rev. E, 87:042123, Apr 2013.

[50] K. Riley, M. Hobson, and Ss Bence. Mathematical methods for physics and engineering - 3rd edition. American Journal of Physics, 67, 02 1999.

[51] A. E Allahverdyan, R Balian, and Th. M Nieuwenhuizen. Maximal work extraction from finite quantum systems. Europhysics Letters (EPL), 67(4):565–571, aug 2004.

[52] Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. Resource theory of quantum states out of thermal equilibrium. Phys. Rev. Lett., 111:250404, Dec 2013.

[53] Emmy Noether. Invariant variation problems. Transport Theory and Statistical Physics, 1(3):186–207, 1971.

[54] H.F. Jones. Groups, Representations and Physics. CRC Press, 1998.

[55] M. Tinkham. Group theory and quantum mechanics. International series in pure and applied physics. McGraw-Hill, 1964.

[56] D. Grier and L. Schaeffer. The Classification of Stabilizer Operations over Qubits. ArXiv e-prints, March 2016.

[57] D. Gottesman. Stabilizer codes and quantum error correction. PhD thesis, California Institute of Technology, 1997.

[58] D S Franca and A K Hashagen. Approximate randomized benchmarking for finite groups. Journal of Physics A: Mathematical and Theoretical, 51(39):395302, aug 2018.

[59] Robert Raussendorf. Cohomological framework for contextual quantum computations. Quantum Information & Computation, 19(13&14):1141–1170, 2019.

[60] Markus Frembs, Sam Roberts, and Stephen D Bartlett. Contextuality as a resource for measurement-based quantum computation beyond qubits. New Journal of Physics, 20(10):103011, oct 2018.

[61] Shalev Ben-David and Supartha Podder. How symmetric is too symmetric for large quantum speedups? arXiv e-prints, page arXiv:2001.09642, Jan 2020.

[62] Iman Marvian, Robert W. Spekkens, and Paolo Zanardi. Quantum speed limits, coherence, and asymmetry. Phys. Rev. A, 93:052331, May 2016.

[63] E. P. Wigner. Die messung quantenmechanischer operatoren. Zeitschrift fur Physik A Hadrons and nuclei, 133(1):101–108, Sep 1952.

[64] Huzihiro Araki and Mutsuo M. Yanase. Measurement of quantum mechanical operators. Phys. Rev., 120:622–626, Oct 1960.

[65] L. Loveridge and P. Busch. Measurement of quantum mechanical operators revisited. The European Physical Journal D, 62(2):297–307, Apr 2011.

[66] Mehdi Ahmadi, David Jennings, and Terry Rudolph. The Wigner-Araki-Yanase theorem and the quantum resource theory of asymmetry. New Journal of Physics, 15(1):013057, 2013.

[67] I. Marvian and R. W. Spekkens. An information-theoretic account of the Wigner-Araki-Yanase theorem. ArXiv e-prints, December 2012.

[68] Miguel Navascues and Sandu Popescu. How energy conservation limits our measurements. Phys. Rev. Lett., 112:140502, Apr 2014.

[69] Dominik Safranek, Mehdi Ahmadi, and Ivette Fuentes. Quantum parameter estimation with imperfect reference frames. New Journal of Physics, 17(3):033012, 2015.

[70] I. Marvian and R. W. Spekkens. Extending Noether's theorem by quantifying the asymmetry of quantum states. Nature Communications, 5:3821, May 2014.

[71] Cristina Cirstoiu, Kamil Korzekwa, and David Jennings. Robustness of Noether's principle: Maximal disconnects between conservation laws and symmetries in quantum theory, 2019.

[72] Matteo Lostaglio, Kamil Korzekwa, David Jennings, and Terry Rudolph. Quantum coherence, time-translation symmetry, and thermodynamics. Phys. Rev. X, 5:021001, Apr 2015.

[73] Gilad Gour, David Jennings, Francesco Buscemi, Runyao Duan, and Iman Marvian. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics. Nature Communications, 9(1):5352, 2018.

[74] M. Lostaglio, D. Jennings, and T. Rudolph. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. Nature Communications, 6:6383, March 2015.

[75] M. Hamermesh. Group Theory and Its Application to Physical Problems. Addison Wesley Series in Physics. Dover Publications, 1989.

[76] I. Bengtsson and K. Zyczkowski. Geometry of Quantum States: An Introduction to Quantum Entanglement. Cambridge University Press, 2017.

[77] A. Arvanitogeorgos. An Introduction to Lie Groups and the Geometry of Homogeneous Spaces. Student mathematical library. American Mathematical Society, 2003.

[78] G.E. Bredon. Introduction to Compact Transformation Groups. Pure and Applied Mathematics. Elsevier Science, 1972.

[79] C. Cirstoiu and D. Jennings. Irreversibility and quantum information flow under global and local gauge symmetries. ArXiv e-prints, Dec 2017.

[80] Marvian Mashhad, Iman. Symmetry, Asymmetry and Quantum Information. PhD thesis, 2012.

[81] Sania Jevtic, David Jennings, and Terry Rudolph. Quantum mutual information along unitary orbits. Physical Review A, 85(5):052121, 2012.

[82] Sania Jevtic, David Jennings, and Terry Rudolph. Maximally and minimally correlated states attainable within a closed evolving system. Physical review letters, 108(11):110403, 2012.

[83] Ryszard Horodecki and Michal Horodecki. Information-theoretic aspects of inseparability of mixed states. Phys. Rev. A, 54:1838–1843, Sep 1996.

[84] A. Gilchrist, D. R. Terno, and C. J. Wood. Vectorization of quantum operations and its use. ArXiv e-prints, November 2009.

[85] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th Beth. Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law. International Journal of Theoretical Physics, 39(12):2717–2753, 2000.

[86] Erick Hinds Mingo, Yelena Guryanova, Philippe Faist, and David Jennings. Quantum Thermodynamics with Multiple Conserved Quantities, pages 751–771. Springer International Publishing, Cham, 2018.

[87] J. A. Vaccaro, F. Anselmi, H. M. Wiseman, and K. Jacobs. Tradeoff between extractable mechanical work, accessible entanglement, and ability to act as a reference system, under arbitrary superselection rules. Phys. Rev. A, 77:032114, Mar 2008.

[88] Gilad Gour, Iman Marvian, and Robert W. Spekkens. Measuring the quality of a quantum reference frame: The relative entropy of frameness. Phys. Rev. A, 80:012307, Jul 2009.

[89] Gilad Gour and Robert W Spekkens. The resource theory of quantum reference frames: manipulations and monotones. New Journal of Physics, 10(3):033023, 2008.

[90] Iman Marvian and Robert W Spekkens. The theory of manipulations of pure state asymmetry: I. basic tools, equivalence classes and single copy transformations. New Journal of Physics, 15(3):033001, 2013.

[91] S. Burris and H.P. Sankappanavar. A course in universal algebra. Graduate texts in mathematics. Springer-Verlag, 1981.

[92] R. Schmidt. Subgroup Lattices of Groups. De Gruyter Expositions in Mathematics. De Gruyter, 1994.

[93] M. Keyl and R. F. Werner. Optimal cloning of pure states, testing single clones. Journal of Mathematical Physics, 40(7):3283–3299, 1999.

[94] Dennis Kretschmann, Dirk Schlingemann, and Reinhard F. Werner. A continuity theorem for Stinespring's dilation. Journal of Functional Analysis, 255(8):1889 – 1904, 2008.

[95] S. Amari and H. Nagaoka. Methods of Information Geometry. Translations of mathematical monographs. American Mathematical Society, 2007.

[96] Oleg Kabernik. Quantum coarse graining, symmetries, and reducibility of dynamics. Phys. Rev. A, 97:052130, May 2018.

[97] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. The Journal of Chemical Physics, 21(6):1087–1092, 1953.

[98] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1):97–109, 04 1970.

[99] Francis Sullivan and Jack Dongarra. Guest editors' introduction: The top 10 algorithms. Computing in Science & Engineering, 2:22–23, jan 2000.

[100] D.J.C. MacKay. Information Theory, Inference and Learning Algorithms. Cambridge University Press, 2003.

[101] M. S. Keane and George L. O'Brien. A Bernoulli factory. ACM Trans. Model. Comput. Simul., 4(2):213–219, April 1994.

[102] Jean-Michel Marin, Pierre Pudlo, Christian P. Robert, and Robin J. Ryder. Approximate Bayesian computational methods. Statistics and Computing, 22(6):1167–1180, Nov 2012.

[103] J Lintusaari, P Blomstedt, T Sivula, MU Gutmann, S Kaski, and J Corander. Resolving outbreak dynamics using approximate Bayesian computation for stochastic birth-death models [version 1; peer review: 1 approved with reservations]. Wellcome Open Research, 4(14), 2019.

[104] Alan R. Templeton. Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. Mol Ecol, 18(2):319–331, Jan 2009. 19192182[pmid].

[105] Radu Herbei and L. Mark Berliner. Estimating ocean circulation: An MCMC approach with approximated likelihoods via the Bernoulli factory. Journal of the American Statistical Association, 109(507):944–954, 2014.

[106] Lov K. Grover. Quantum mechanics helps in searching for a needle in a haystack. Phys. Rev. Lett., 79:325–328, Jul 1997.

[107] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In Proceedings 35th Annual Symposium on Foundations of Computer Science, pages 124–134, Nov 1994.

[108] Ewin Tang. A quantum-inspired classical algorithm for recommendation systems. In Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing,

STOC 2019, page 217228, New York, NY, USA, 2019. Association for Computing Machinery.

[109] Ewin Tang. Quantum-inspired classical algorithms for principal component analysis and supervised clustering. CoRR, abs/1811.00414, 2018.

[110] András Gilyén, Seth Lloyd, and Ewin Tang. Quantum-inspired low-rank stochastic regression with logarithmic dependence on the dimension. CoRR, abs/1811.04909, 2018.

[111] Howard Dale, David Jennings, and Terry Rudolph. Provable quantum advantage in randomness processing. Nature Communications, 6(1):8203, 2015.

[112] Johan Wastlund. Functions arising by coin flipping, 1999.

[113] John von Neumann. Various techniques used in connection with random digits. In A.S. Householder, G.E. Forsythe, and H.H. Germond, editors, Monte Carlo Method, pages 36–38. National Bureau of Standards Applied Mathematics Series, 12, Washington, D.C.: U.S. Government Printing Office, 1951.

[114] Ronen Shaltiel. An introduction to randomness extractors. In Luca Aceto, Monika Henzinger, and Jiri Sgall, editors, Automata, Languages and Programming, pages 21–41, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[115] Noam Nisan and Amnon Ta-Shma. Extracting randomness: A survey and new constructions. Journal of Computer and System Sciences, 58(1):148 – 173, 1999.

[116] Yuval Peres. Iterating von Neumann's procedure for extracting random bits. Ann. Statist., 20(1):590–597, 03 1992.

[117] Mikhail J. Atallah and Marina Blanton. Algorithms and Theory of Computation Handbook. Chapman & Hall/CRC, 2nd edition, 2009.

[118] L. Breiman. Probability. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1968.

[119] Raj B. Patel, Terry Rudolph, and Geoff J. Pryde. An experimental quantum Bernoulli factory. Science Advances, 5(1), 2019.

[120] Kenneth M. Levasseur. A probabilistic proof of the Weierstrass approximation theorem. The American Mathematical Monthly, 91(4):249–250, 1984.

[121] Larry Wasserman. All of Statistics: A Concise Course in Statistical Inference. Springer Publishing Company, Incorporated, 2010.

[122] Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum Amplitude Amplification and Estimation. Quantum Computation and Quantum Information, 305:53–74, 2002.

[123] Theodore J. Yoder. Building and bounding quantum Bernoulli factories. https://www.scottaaronson.com/6s899/tedyoder.pdf.

[124] Jiaqing Jiang, Jialin Zhang, and Xiaoming Sun. Quantum-to-quantum Bernoulli factory problem. Phys. Rev. A, 97:032303, Mar 2018.

[125] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum Metropolis sampling. Nature, 471(7336):87–90, 2011.

[126] Maris Ozols, Martin Roetteler, and Jérémie Roland. Quantum rejection sampling. In Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, ITCS 12, page 290308, New York, NY, USA, 2012. Association for Computing Machinery.

[127] Alexei Kitaev and William A. Webb. Wavefunction preparation and resampling using a quantum computer. arXiv e-prints, page arXiv:0801.0342, January 2008.

[128] William Cottrell, Ben Freivogel, Diego M. Hofman, and Sagar F. Lokhande. How to build the thermofield double state. Journal of High Energy Physics, 2019(2):58, Feb 2019.

[129] Michael Park. Quantum Circuits for Scale Transformations. arXiv e-prints, page arXiv:1903.08257, Mar 2019.

[130] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. arXiv e-prints, page arXiv:0208112, August 2002.