

**Coherence, thermodynamics  
and uncertainty relations**

Kamil Korzekwa

Imperial College London  
Department of Physics  
November 2016

Supervisors: David Jennings and Terry Rudolph

# Author's declaration

The content of this thesis is a result of research carried out during the last three years as a graduate student under the supervision of David Jennings and Terry Rudolph. I hereby declare that the presented material is my own, except where otherwise acknowledged and appropriately referenced.

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

The results presented in this thesis have been published in the following papers:

- K. Korzekwa, D. Jennings, T. Rudolph, “Operational constraints on state-dependent formulations of quantum error-disturbance trade-off relations,” *Phys. Rev. A* **89**, 052108 (2014).
- K. Korzekwa, M. Lostaglio, D. Jennings, T. Rudolph, “Quantum and classical entropic uncertainty relations,” *Phys. Rev. A* **89**, 042122 (2014).
- M. Lostaglio, K. Korzekwa, D. Jennings, T. Rudolph, “Quantum coherence, time-translation symmetry and thermodynamics,” *Phys. Rev. X* **5**, 021001 (2015).
- K. Korzekwa, M. Lostaglio, J. Oppenheim, D. Jennings, “The extraction of work from quantum coherence,” *New J. Phys.* **18**, 023045 (2016).
- K. Korzekwa, M. Lostaglio, “Classical noise and the structure of minimum uncertainty states,” *Phys. Rev. A* **93**, 062347 (2016).

# Abstract

The principle of superposition is one of the main building blocks of quantum physics and has tremendous consequences both for our fundamental understanding of nature and for technological applications. In particular, the existence of coherent superposition leads to the concept of unavoidable quantum uncertainty. The role played by this “coherent” uncertainty within thermodynamics, as well as its relationship to classical lack of knowledge, is the main subject of this thesis.

In Part I we study thermodynamic limitations of processing quantum coherence within a resource-theoretic framework. Using the time-translation symmetry that arises from the first law of thermodynamics, we find constraints on possible manipulations of coherence and prove their irreversibility due to the second law. We also generalise to the quantum domain Szilard’s concept of converting information into work. Namely, we show how, in the presence of a heat bath, coherence of a system can be exploited to perform mechanical work. Finally, we analyse the effect that coherence has on the structure of the thermodynamic arrow of time, i.e., on the set of states into which a given state can freely evolve under thermodynamic constraints.

In Part II we focus on the interplay between quantum and classical uncertainty manifested in uncertainty relations. We show that separating the total uncertainty into these two distinct components leads to a new type of “fixed-entropy” uncertainty relation. We also analyse how classical ignorance affects the structure of states that minimise the unavoidable uncertainty arising from the noncommutativity of two observables. Finally, we study error-disturbance trade-off relations and, by proving that quantum uncertainty can be simultaneously maximised for any two observables, we clarify the unphysical nature of state-dependent relations.

# Acknowledgements

First of all I would like to thank my supervisors for their guidance and support. In particular, I am grateful to David Jennings for his patience, rigorous reasoning and scientific responsibility; and to Terry Rudolph for his passion, intuitive reasoning and relaxed approach to non-scientific matters. I am also greatly indebted to my friends and closest collaborators, Matteo Lostaglio and Antony Milne. It was a pleasure to have been “uncovering the mysteries of the quantum realm” in your company for the last four years. I am thankful for all the amazing time that we have spent together, whether doing physics, enjoying London, or just fooling around.

I am also very grateful to all the members of the Centre for Doctoral Training in Controlled Quantum Dynamics for creating an incredibly friendly, stimulating and supporting working environment. I would like to express my gratitude, in particular, to the members of Cohort 4: Zachary Blunden-Codd (for being Zach all the time everywhere), Max Frenzel (for all the late nights out), David Murgia (for quite a few enlightening chats), Claudio Polisseni (for being the best cohort representative), Izzie Rabey (for your down-to-earth attitude), Chris Self (for your empathy and endless discussions about all things that matter), and Chris Sparrow (for your optimistic pessimism). Also, many thanks to all the PhD students for creating the unforgettable atmosphere of Level 12, especially to: Cristina Cirstoiu, Howard Dale, Ben Dive, Mercedes Gimeno-Segovia, David Newman and Viggie Venkataraman. Furthermore, I would like to acknowledge the helpful discussions on the topics presented in this thesis with the following researchers: Alvaro Alhambra, Janet Anders, Oscar Dahlsten, Martí Perarnau, Chris Perry, Jonathan Richens, Tommaso Tufarelli and Raam Uzdin.

Last, but not least, I would like to thank my family and friends. Special thanks to my brother Piotrek for amazing four years together in London, as well as to the other two phenomenal flatmates at 5 Glenridding: Gosia and Alicja. I also owe a debt of gratitude to my wonderful friends from Wrocław, especially to Michał “Węzu” Marczenko and Ewa “Zu” Kupińska. Finally, I would like to thank my parents for constantly supporting all the life choices that have led me to this point.

# Contents

<b>Introduction</b>	<b>11</b>
<b>Mathematical background and notation</b>	<b>16</b>
<b>I Quantum thermodynamics</b>	<b>23</b>
<b>Summary of results</b>	<b>24</b>
<b>1 Resource theory approach to thermodynamics</b>	<b>26</b>
1.1 Thermal operations . . . . .	29
1.1.1 Superset of Gibbs-preserving operations . . . . .	30
1.1.2 Superset of time-translation covariant operations . . . . .	31
1.2 Thermodynamics of classical states . . . . .	32
1.2.1 Equivalence of GP and thermal operations . . . . .	33
1.2.2 Thermodynamic ordering . . . . .	35
1.2.3 Catalysts . . . . .	40
1.2.4 Work extraction and the Szilard engine . . . . .	41
<b>2 Thermal operations and modes of coherence</b>	<b>51</b>
2.1 Modes of coherence . . . . .	52
2.2 Bounds on coherence transformations . . . . .	54
2.2.1 Upper bound for covariant operations . . . . .	54
2.2.2 Upper bound for thermal operations . . . . .	57
2.2.3 Lower bound on guaranteed coherence preservation . . . . .	58

2.3	Analysis and applications of the bounds . . . . .	60
2.3.1	Coherence shifting and merging . . . . .	60
2.3.2	Qubit transformations . . . . .	65
2.3.3	General remarks . . . . .	68
<b>3</b>	<b>Work extraction from coherence</b>	<b>71</b>
3.1	Coherence and work . . . . .	72
3.1.1	Work-locking . . . . .	74
3.1.2	Different thermodynamic regimes . . . . .	75
3.1.3	Individual processing regime . . . . .	76
3.2	Protocol . . . . .	80
3.2.1	Explicit steps of the work-extraction protocol . . . . .	82
3.2.2	Performance . . . . .	84
3.3	Fundamental limitations of coherence-to-work conversion . . . . .	86
3.3.1	Limitations of bounded references . . . . .	86
3.3.2	Approaching free energy limit of work extraction . . . . .	88
3.4	Extracting work with perfect repeatability and bounded references . . . . .	91
3.4.1	Average work extraction . . . . .	92
3.4.2	Single-shot work extraction . . . . .	94
<b>4</b>	<b>Order theory approach</b>	<b>96</b>
4.1	Order theory . . . . .	98
4.1.1	Partial order . . . . .	98
4.1.2	Lattice . . . . .	100
4.2	Ordering of classical states . . . . .	102
4.2.1	Infinite temperature and a lattice structure . . . . .	102
4.2.2	Finite temperatures and a broken lattice structure . . . . .	105
4.3	Ordering of quantum states . . . . .	109
4.3.1	Infinite temperature and a lattice structure . . . . .	109
4.3.2	Finite temperatures: qubit evidence for a lattice structure . . . . .	110

<b>II</b>	<b>Uncertainty relations</b>	<b>115</b>
	<b>Summary of results</b>	<b>116</b>
<b>5</b>	<b>Quantum and classical sources of uncertainty</b>	<b>118</b>
5.1	Splitting total uncertainty into quantum and classical parts . . . . .	119
5.1.1	Extension to general projective measurements . . . . .	122
5.2	Non-existence of state-independent bounds . . . . .	124
5.3	Quantum uncertainty relations and fixed-entropy lower bounds . . . . .	127
5.3.1	Strong fixed-entropy lower bounds are nonlinear for $d > 2$ . . . . .	129
5.3.2	Linear strong fixed-entropy lower bound for qubits . . . . .	131
5.4	Minimum uncertainty states with fixed von Neumann entropy . . . . .	135
<b>6</b>	<b>Minimum uncertainty states in the presence of classical noise</b>	<b>139</b>
6.1	General families of uncertainty measures . . . . .	140
6.1.1	Minimal requirement of Schur-concavity . . . . .	141
6.1.2	Requirement of context-independence . . . . .	142
6.1.3	Requirement of decidability . . . . .	143
6.2	Minimum uncertainty states . . . . .	145
6.2.1	Universal minimum uncertainty states . . . . .	146
6.2.2	No-go theorem for a pure universal MUS . . . . .	147
6.3	Noise and uncertainty . . . . .	150
6.3.1	Role of noise and $H_2$ in the classical case . . . . .	150
6.3.2	Role of noise and $H_2$ in the quantum case . . . . .	153
6.4	Existence of a noisy universal MUS . . . . .	155
6.4.1	General results . . . . .	155
6.4.2	Universal MUS for qubit systems . . . . .	157
6.4.3	Approximate notion of universality for higher dimensions . . . . .	158
6.5	Thermodynamic considerations . . . . .	161
<b>7</b>	<b>Error-disturbance relations</b>	<b>165</b>
7.1	General framework . . . . .	167
7.1.1	State-dependent notions . . . . .	167
7.1.2	Disturbance of a measurement . . . . .	169

7.2	Vanishing error and disturbance . . . . .	170
7.2.1	Mutually unbiased ZNZN states . . . . .	171
7.2.2	Existence of non-trivial ZNZN states . . . . .	173
7.2.3	Consequences for error-disturbance relations . . . . .	176
7.3	Analysis of the Ozawa error-disturbance relation . . . . .	177
<b>Conclusions</b>		<b>180</b>
<b>Appendices</b>		<b>185</b>
<b>A</b>	<b>Alternative proof of Theorem 2.1</b>	<b>186</b>
<b>B</b>	<b>Details of the proof of Theorem 3.1</b>	<b>188</b>
<b>C</b>	<b>Proof of Theorem 4.1</b>	<b>191</b>
<b>D</b>	<b>Existence of a universal MUS for qubit systems</b>	<b>194</b>
<b>Bibliography</b>		<b>199</b>

# List of Figures

1.1	Thermodynamic arrow of time . . . . .	28
1.2	Work extraction from an incoherent qubit state . . . . .	46
2.1	Guaranteed coherence preservation . . . . .	59
2.2	Irreversibility of coherence shift cycle . . . . .	63
2.3	Irreversibility of coherence merging cycle . . . . .	64
2.4	Extremal qubit states under constraint dynamics . . . . .	66
2.5	Future thermal cones for qubits (thermal operations) . . . . .	67
2.6	General structure . . . . .	69
3.1	Thermodynamic regimes . . . . .	76
3.2	Basic work-extraction protocol . . . . .	83
3.3	Coherence boost to average work extraction from $ \gamma\rangle$ . . . . .	93
3.4	Coherence boost to single-shot work extraction from $ \gamma\rangle$ . . . . .	95
4.1	Decomposition of preorder into partial order between equivalence classes . . . . .	99
4.2	Visualising the join and meet . . . . .	101
4.3	Examples of lattice and non-lattice partial orders . . . . .	103
4.4	Meet of a majorisation lattice . . . . .	104
4.5	Join of a majorisation lattice . . . . .	105
4.6	Thermo-majorisation order is not a lattice . . . . .	107
4.7	Future thermal cones for qubits (GP operations) . . . . .	111
4.8	Thermodynamic lattice for qubits . . . . .	113
5.1	Parametrisation of the Bloch sphere . . . . .	132
5.2	QC uncertainty plots for qubits . . . . .	134
5.3	MUSs with fixed von Neumann entropy for qubits . . . . .	136

5.4	QC uncertainty plots for qutrits . . . . .	137
6.1	Three families of uncertainty functions . . . . .	145
6.2	Emergence of a universal MUS for qubits . . . . .	159
6.3	Emergence of an approximate universal MUS for qutrits . . . . .	161
7.1	Uncertainty relations vs error-disturbance trade-off relations . . . . .	166
C.1	Possible orderings of eigenvalues . . . . .	192

# Introduction

Although quantum mechanics owes its name to the quantisation of energy, postulated at the beginning of the 20th century by the old quantum theory, its counter-intuitive “weirdness” comes from the superposition principle that was discovered a quarter of a century later. Concepts such as Schrödinger’s cat, quantum teleportation and quantum computing all stem from this fundamental rule of the quantum realm: given two valid states of a system, e.g., spin-up and spin-down, their coherent superposition is also a valid state. This principle leads to a new kind of purely quantum uncertainty that is unavoidable. In classical physics, if we fully specify the state of a system then we can predict with certainty the outcomes of all measurements performed on that system. However, in quantum physics even if we possess complete knowledge about a system state, there will exist measurements for which the outcomes are uncertain, and the best we can do is to predict their probabilities. For example, if a system is prepared in an equal superposition of spin-up and spin-down states, we know everything there is to know about the system, but we are still uncertain about the outcome of a spin direction measurement – it will yield the spin-up and spin-down results with 50% probability each. It is vital to emphasise that there is a crucial difference between this inability to provide a deterministic answer and classical uncertainty, when the system is in one of the two states but we just do not know which. The former cannot be avoided as it reflects the structure of nature, whilst the latter just describes our inability to get a full description of a system state.

This thesis studies the consequences of the interplay between these two different kinds of uncertainty. In Part I we focus on thermodynamics, where classical uncertainty represented by entropy is always present, and investigate the role played by coherence in thermodynamic processes. In Part II we study how uncertainty relations, originating purely from quantum coherence, are affected by the presence of classical ignorance.

## Quantum thermodynamics

Fundamental laws of nature often take the form of restrictions: nothing can move faster than light in vacuum, energy cannot be created from nothing, there exist no perpetual mobiles. It is due to these limitations that we can ascribe values to different objects and phenomena, e.g., energy would not cost us money if we could obtain it for free. The mathematical framework developed to study the influence of such constraints on the possible evolution of physical systems is known under the collective name of *resource theories*.

Perhaps the best known example of this approach was to formalise and harness the puzzling phenomenon of entanglement [1]. Insight into such quantum correlations has been obtained by considering the following two-party scenario: Alice and Bob each control a part of a bigger quantum system and can only perform operations locally on their parts. In addition they are allowed to communicate through a classical channel. With this restricted set of operations, known as *local operations and classical communication* (LOCC), entanglement cannot be created between Alice and Bob. However, if Alice and Bob were initially handed an entangled state, they could perform otherwise impossible operations, such as quantum teleportation [2] or superdense coding [3]. This way one may study the power of entanglement as a resource that enlarges the set of LOCC operations, and one can also quantify it, e.g., given a particular entangled bipartite state one may ask how many quantum bits can be teleported between Alice and Bob.

This approach encapsulates a more general idea that underlies all resource theories. These are formulated in terms of free objects and a set of rules specifying how we can transform them. Within quantum theory we thus specify a subset of quantum states and a subset of quantum operations that are considered free (these are represented by separable states and LOCC operations in the case of the resource theory of entanglement). Every state that is not free is considered to be a resource, and free transformations, which are assumed to be the only ones allowed within the theory, must map the set of free states into itself. This way no resources can be freely created and one can investigate how they can be exploited and manipulated under free transformations. This includes, e.g., studying the possibility of interconversion between two given states, partially ordering the set of all states based on resource measures, and studying the advantages of possessing a resource state in a given quantum information processing protocol.

When speaking of thermodynamics one inevitably thinks of concepts such as heat flows, thermal machines and pressure, which seem to be far removed from the ideas of quantum information theory. However, on a more abstract level, thermodynamics can be seen as a field studying the accessibility/inaccessibility of one physical state from another [4, 5]. The first and second laws of thermodynamics are fundamental constraints on state transformations, forcing thermodynamic processes to conserve the overall energy and forbidding free conversion of heat into work. Hence, the resource-theoretic machinery developed to study entanglement is also perfectly suited to shed light on thermodynamics. A natural question then arises: what amounts to a resource when we are restricted by the laws of thermodynamics? One answer is *athermality* – a property whereby a state has a distribution over energy levels that is not thermal [6]. This is because such a state, during the process of equilibration, can be exploited to perform work [7, 8], which in turn can be used to drive another system out of equilibrium. However, *coherence* can also be viewed as a second, independent resource in thermodynamics [9]. This stems from the fact that energy conservation, implied by the first law, restricts the thermodynamic processing of coherence. Hence, possessing a state with coherence allows one to implement otherwise impossible transformations.

Although the resource-theoretic approach has yielded new insights into the thermodynamics of individual quantum systems, most of them are limited to quantum states that do not possess coherence between different energy eigenspaces [7, 8, 10, 11]. Despite partial results [12–16] we still lack a complete understanding of coherence manipulations allowed by thermodynamic means. This foundational question may be of interest for future advancements in nanotechnology, as interference effects are particularly relevant [17, 18] at scales we are increasingly able to control [19–23]. Moreover, recent evidence suggests that biological systems may harness quantum coherence in relevant timescales [24–26]. Therefore, the main aim of Part I of this thesis is to study the problem of coherence processing within thermodynamics.

## Uncertainty relations

The proper assessment of uncertainties and measurement errors is at the very core of the scientific method, and thus probability has been an indispensable component of physics for centuries. On the one hand the use of statistical tools allows one to appropriately

account for uncontrollable factors in an experiment, thus avoiding false conclusions. On the other hand, the proper incorporation of uncertainties about the actual state of a large and complex system results in a probabilistic description of that system. This allows one to make statements and predictions about average global properties of the system and fluctuations about these averages. This way one can obtain incomplete but useful information about the system that would otherwise be completely intractable. As an example consider a macroscopic container with gas particles, a full description of which would require keeping track of the positions and momenta of the order of  $10^{23}$  particles, which is practically impossible. Instead, most useful information about the macroscopic state of the gas can be recovered without such complete knowledge, but from averaged quantities such as pressure and temperature. This approach lies at the heart of statistical physics that at the end of 19th century formalised (and actually replaced) traditional approaches to thermodynamics.

Hence, by the time quantum mechanics appeared on the stage in the early 1900s, physicists were already used to probabilistic models of reality.<sup>1</sup> However, until then the probabilistic approach was only considered a useful tool to overcome practical difficulties of being completely certain about a state of a physical system (as in the example with a gas container presented above). As such, uncertainty was not considered something fundamental – rather it was a statement about the imperfection of our tools and observation methods. The advent of quantum mechanics changed this perspective completely. We now know that while classical uncertainty arises from ignorance, quantum phenomena are irreducibly unpredictable: even for a single fixed measurement and a pure quantum state of maximal knowledge, we can typically only make probabilistic predictions.

The situation worsens when we consider two potential measurements of a system that do not commute – there exist fundamental constraints on our ability to make predictions about either possible set of outcomes. The most celebrated such constraint is the Heisenberg-Robertson uncertainty relation [28]; however, over time, the study of uncertainty relations has itself become a major topic and has played an important role in uncovering the mysteries of the quantum realm. After the original introduction by Heisenberg as an error-disturbance relation in his famous thought experiment [29], they

---

<sup>1</sup>It is worth noting that the idea of quantisation that has revolutionised our understanding of the probabilistic nature of Nature actually originated from Max Planck's work on the thermodynamic problem of black-body radiation. [27]

have been reformulated and refined in various ways over the last 90 years. Starting as a statement about the outcome statistics of independent measurements of position and momentum [30], they were quickly extended to generic observables [28]. Then, in the late 1940s, the link between uncertainty and information was pointed out by Shannon, resulting in the creation of a new field of science – information theory [31]. However, another 35 years had to pass,<sup>2</sup> until David Deutsch proposed the use of information entropy (Shannon entropy) of measurement outcome statistics to study uncertainty relations [33]. This not only brought a paradigm shift by leading to state-independent lower bounds on uncertainty [34], but also initiated the ongoing intimate relationship between uncertainty relations and the field of information theory.

Thus, with the rise of quantum information theory in the 1980s and its rapid development until this day, entropic uncertainty relations have become a widely studied subject within this new framework. Multiple entropic measures have been used to quantify uncertainty, allowing physicists to phrase a variety of quantum information problems in the language of uncertainty relations (for a complete review see Ref. [35, 36] and references therein). Let us just mention here the use of entropic uncertainty relations for quantum cryptographic security proofs [37, 38], quantitative wave-particle duality relations [39] and witnessing entanglement [38, 40, 41]. All these applications refocused interest in uncertainty relations, which was originally due to the fundamental limitations they impose on our ability to predict measurement outcomes. Previously, physicists were mostly interested in extreme scenarios of quantum systems prepared in pure states of maximal knowledge, when the origin of uncertainty lies in the coherence between eigenstates of a given observable. However, as soon as uncertainty relations are used to study quantum information problems, the investigated systems are not necessarily prepared in pure states. In fact, most quantum states arising in theoretical considerations (e.g., in entanglement theory) and the ones used to describe quantum systems prepared in laboratories are mixed. Therefore, a proper understanding of the interplay between classical lack of knowledge and purely quantum uncertainty coming from coherence is crucial for quantum information based technologies; deepening this understanding forms the main aim of Part II of this thesis.

---

<sup>2</sup>Let us point out that the paper by Białynicki-Birula and Mycielski [32] that introduced entropic uncertainty relations for continuous variables preceded the seminal paper by Deutsch [33].

# Mathematical background and notation

## Notation

### Operators

Unless otherwise specified all operators will act on  $d$ -dimensional Hilbert spaces. Appropriate subscripts will indicate the ancillary system that the operator acts on (e.g.,  $A_B$  acts on an ancillary bath system), whereas the operators acting on the main system under investigation will lack a subscript  $S$  unless necessary (e.g., for denoting a joint state  $\rho_{SB}$ ). The spectral decomposition of a general operator  $A$  will be given by

$$A = \sum_{i=1}^d a_i |a_i\rangle\langle a_i|,$$

i.e., the eigenvalues and eigenvectors will be labelled by lower-case letters corresponding to a given observable. An important exception to this rule is the energy observable. The eigenvectors and eigenvalues of a system Hamiltonian  $H$  will be denoted by  $\{\epsilon_i\}$ , of the bath Hamiltonian  $H_B$  by  $\{E_i\}$ , and of the ancillary system Hamiltonian  $H_A$  by  $\{\epsilon_i^A\}$ . The identity operator will be represented by  $\mathbb{I}$ .

The set of  $d \times d$  density operators, i.e., the set of positive semi-definite operators with unit trace that act on a  $d$ -dimensional Hilbert space, will be denoted by  $\mathcal{S}_d$ . A pure density matrix  $\rho = |\psi\rangle\langle\psi|$  will sometimes be identified with the vector  $|\psi\rangle$  from the corresponding Hilbert space. Completely positive trace-preserving maps (quantum channels) will be denoted by a calligraphic letter  $\mathcal{E}$ .

We will make extensive use of the following Pauli operators acting on two-level systems:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{M1})$$

Moreover, we will denote their eigenvectors by  $\{|+\rangle, |-\rangle\}$ ,  $\{|i\rangle, |-i\rangle\}$ ,  $\{|0\rangle, |1\rangle\}$ , respectively. We will also use the vector of Pauli matrices  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  to represent general qubit observables,  $A = \alpha_1 \mathbb{I} + \alpha_2 \mathbf{a} \cdot \boldsymbol{\sigma}$  (usually with  $\alpha_1 = 0$  and  $\alpha_2 = 1$ ), and general qubit states:

$$\rho = \frac{\mathbb{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}, \quad |\mathbf{r}| \leq 1. \quad (\text{M2})$$

### Probability distributions and measurements

The set of  $d$ -dimensional probability vectors will be denoted by  $\mathcal{P}_d$  and its elements by bold lower-case letters starting from  $\mathbf{p}$ . A maximally mixed probability vector will be denoted by  $\boldsymbol{\eta} = (1/d, \dots, 1/d)$ . Given a probability vector  $\mathbf{p}$ , we will denote by  $\mathbf{p}^\downarrow$  the same vector but with elements rearranged in a nonincreasing order. Probability distributions linked with the projective measurement of an observable  $A$  performed on a state  $\rho$  will be denoted by  $\mathbf{p}^A(\rho)$ ,

$$\mathbf{p}^A(\rho) = (p_1^A(\rho), \dots, p_d^A(\rho)), \quad p_i^A(\rho) = \text{Tr}(\rho |a_i\rangle\langle a_i|). \quad (\text{M3})$$

A map corresponding to a non-selective projective measurement of an observable  $A$ , i.e. a dephasing map in the eigenbasis of  $A$ , will be denoted by  $\mathcal{D}_A$ ,

$$\mathcal{D}_A(\rho) = \sum_{i=1}^d \langle a_i | \rho | a_i \rangle |a_i\rangle\langle a_i|. \quad (\text{M4})$$

The dephasing map in the energy eigenbasis will be simply denoted by  $\mathcal{D}$ .

### Colour-coding

Throughout the thesis the important content will be colour-coded in the following way:

**Definitions/Axioms/Requirements**

**Theorems/Lemmas/Observations/Corollaries**

**Examples/Remarks**

## List of abbreviations

CPTP	Completely positive trace-preserving
EUR	Entropic uncertainty relation
GP	Gibbs-preserving
LOCC	Local operations and classical communication
MUB	Mutually unbiased basis
MUS	Minimum uncertainty state
OR	Operational requirement
POVM	Positive operator valued measure
QC	Quantum-classical
SFEB	Strong fixed-entropy bound
UR	Uncertainty relation
TO	Thermal operation
ZNZD	Zero-noise zero-disturbance

## Mathematical background

Here, we collect the important definitions and results concerning probability distributions that will be extensively used throughout the thesis. In order to quantitatively assess how uncertain a given probability distribution  $\mathbf{p}$  is, we will use the following entropic measures of uncertainty:

### Definition M1: Rényi entropies

Rényi entropies of a real order  $\alpha \neq 0$  for a  $d$ -dimensional probability distribution  $\mathbf{p}$  are defined as [42, 43]:

$$H_\alpha(\mathbf{p}) := \frac{\text{sgn}(\alpha)}{1 - \alpha} \ln \left( \sum_{i=1}^d p_i^\alpha \right). \quad (\text{M5})$$

The  $\alpha \rightarrow \pm\infty$  and  $\alpha \rightarrow 1$  are defined by suitable limits,

$$H_1(\mathbf{p}) = - \sum_{i=1}^d p_i \ln p_i, \quad H_\infty(\mathbf{p}) = - \ln \max_i p_i, \quad H_{-\infty}(\mathbf{p}) = \ln \min_i p_i. \quad (\text{M6})$$

We define the  $\alpha = 0$  case by the Burg entropy,

$$H_0(\mathbf{p}) = \frac{1}{d} \sum_{i=1}^d \ln p_i, \quad (\text{M7})$$

and *not* by the limit of the Rényi entropy for  $\alpha \rightarrow 0^+$ , i.e.,  $H_0(\mathbf{p}) = \ln n$ , where  $n$  is the number of non-zero entries of  $\mathbf{p}$ . The  $\alpha = 1$  case is known as the Shannon entropy and sometimes simply denoted by  $H$  (not to be confused with the Hamiltonian).

The uncertainty of  $\mathbf{p}$  can also be measured by a “distance” between  $\mathbf{p}$  and a uniform distribution  $\boldsymbol{\eta}$  – the “closer”  $\mathbf{p}$  and  $\boldsymbol{\eta}$  are, the more uncertain  $\mathbf{p}$  is. Similarly, within thermodynamic scenarios we will be interested in “how far” a given distribution  $\mathbf{p}$  is from being a thermal distribution  $\boldsymbol{\gamma}$ . Thus, in order to quantify how much two probability distributions differ, we will use the following relative entropies:

### Definition M2: Rényi divergences

Given probability distributions  $\mathbf{p}$  and  $\mathbf{q}$ , the  $\alpha$ -Rényi divergence (or relative entropy) is defined as [42]:

$$S_\alpha(\mathbf{p}||\mathbf{q}) = \frac{\text{sgn}(\alpha)}{\alpha - 1} \ln \sum_i p_i^\alpha q_i^{1-\alpha}, \quad (\text{M8})$$

for  $\alpha \neq \{0, 1\}$ . The values at  $\alpha = \{0, 1, \pm\infty\}$  are defined through Eq. (M8) by the corresponding limits and read [8]:

$$S_0(\mathbf{p}||\mathbf{q}) = -\ln \sum_{i|p_i \neq 0} q_i, \quad S_1(\mathbf{p}||\mathbf{q}) = \sum_i p_i \ln \frac{p_i}{q_i}, \quad (\text{M9a})$$

$$S_\infty(\mathbf{p}||\mathbf{q}) = \ln \max_i \frac{p_i}{q_i}, \quad S_{-\infty}(\mathbf{p}||\mathbf{q}) = S_\infty(\mathbf{q}||\mathbf{p}). \quad (\text{M9b})$$

Given two probability distributions,  $\mathbf{p}$  and  $\mathbf{q}$ , we will often be interested in which of the two is less uncertain. Although we could compare their Rényi entropies for a particular  $\alpha$ , we will also be employing a more general approach based on the partial order known as majorisation, and defined as follows:

### Definition M3: Majorisation

Given two  $d$ -dimensional probability distributions  $\mathbf{p}$  and  $\mathbf{q}$ , we say that  $\mathbf{p}$  majorises  $\mathbf{q}$ , and write  $\mathbf{p} \succ \mathbf{q}$ , if and only if

$$\sum_{i=1}^n p_i^\downarrow \geq \sum_{i=1}^n q_i^\downarrow, \quad n = 1, \dots, d-1.$$

Within majorisation theory a probability distribution  $\mathbf{p}$  is said to be less uncertain than

$\mathbf{q}$  if  $\mathbf{p} \succ \mathbf{q}$ , which is justified by the following well-known result [44]:

**Remark M1: Majorisation and bistochastic matrices**

The relation  $\mathbf{p} \succ \mathbf{q}$  holds if and only if  $\mathbf{q} = \Lambda \mathbf{p}$  for some bistochastic matrix  $\Lambda$ ,

$$\Lambda_{ij} \geq 0, \quad \sum_{i=1}^d \Lambda_{ij} = 1, \quad \sum_{j=1}^d \Lambda_{ij} = 1.$$

An alternative way to order probability distributions according to their uncertainty is to use a modified version of majorisation known as trumping:

**Definition M4: Trumping**

Given two  $d$ -dimensional probability distributions  $\mathbf{p}$  and  $\mathbf{q}$ , we say that  $\mathbf{p}$  trumps  $\mathbf{q}$ , and denote it by  $\mathbf{p} \succ_T \mathbf{q}$ , when there exists a probability distribution  $\mathbf{r}$  such that

$$\mathbf{p} \otimes \mathbf{r} \succ \mathbf{q} \otimes \mathbf{r}.$$

The trumping partial order, also known as catalytic majorisation [45], is connected with the measures of uncertainty via the following result:

**Remark M2: Trumping and Rényi entropies**

The results of Refs. [46, 47] link trumping and Rényi entropies as follows. If  $\mathbf{p}^\downarrow \neq \mathbf{q}^\downarrow$  then

$$\mathbf{p} \succ_T \mathbf{q} \iff H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q}), \quad \forall \alpha \in \mathbb{R}. \quad (\text{M10})$$

In thermodynamic considerations instead of comparing which of the two probabilities is more uncertain, we will study which one is closer to the thermal equilibrium distribution. Thus, we will use the ordering with respect to a thermal Gibbs distribution defined as follows:

**Definition M5: Thermal Gibbs distribution**

Given a  $d$ -dimensional, nondecreasing set of energies  $\{\epsilon_1 \leq \epsilon_2 \leq \dots \leq \epsilon_d\}$  and an inverse temperature  $\beta \geq 0$ , the thermal Gibbs distribution is given by

$$\boldsymbol{\gamma} = \frac{1}{Z} (e^{-\beta\epsilon_1}, \dots, e^{-\beta\epsilon_d}), \quad Z = \sum_{i=1}^d e^{-\beta\epsilon_i}. \quad (\text{M11})$$

The connection with uncertainty scenarios is provided by the limiting case of infinite temperature:

### Remark M3

In the infinite temperature limit,  $\beta \rightarrow 0$ , the Gibbs state becomes a maximally mixed state:  $\gamma \rightarrow \boldsymbol{\eta} = (1/d, \dots, 1/d)$ .

In order to define the thermodynamic equivalent of majorisation partial order, we first must introduce the  $\beta$ -ordering that will replace nonincreasing ordering at finite temperatures:

### Definition M6: Gibbs rescaling and $\beta$ -ordering

Given a thermal Gibbs distribution  $\gamma$ , Gibbs-rescaled version of a probability distribution  $\mathbf{p}$  is defined by  $\mathbf{p}^\gamma = (p_1/\gamma_1, \dots, p_d/\gamma_d)$ . The  $\beta$ -ordering of a probability distribution  $\mathbf{p}$  is defined by a permutation  $\pi_{\mathbf{p}}$  that arranges  $\mathbf{p}^\gamma$  in a nonincreasing order, i.e.,

$$(\mathbf{p}^\gamma)^\downarrow = \left( p_{\pi_{\mathbf{p}}^{-1}(1)}^\gamma, \dots, p_{\pi_{\mathbf{p}}^{-1}(d)}^\gamma \right).$$

Now, the  $\beta$ -ordered version of a probability vector  $\mathbf{p}$  is given by

$$\mathbf{p}^\beta = \left( p_{\pi_{\mathbf{p}}^{-1}(1)}, \dots, p_{\pi_{\mathbf{p}}^{-1}(d)} \right). \quad (\text{M12})$$

### Remark M4

In the infinite temperature limit,  $\beta \rightarrow 0$ , the  $\beta$ -ordering becomes a nonincreasing ordering:  $\mathbf{p}^\beta \rightarrow \mathbf{p}^\downarrow$ .

Thermodynamic ordering will be defined with the use of thermo-majorisation curves:

### Definition M7: Thermo-majorisation curves

Given a thermal Gibbs distribution  $\gamma$  and a probability vector  $\mathbf{p}$ , its thermo-majorisation curve  $f_{\mathbf{p}}$  is composed of linear segments connecting the point  $(0, 0)$  and the points

$$\left( \sum_{i=1}^k \gamma_i^\beta, \sum_{i=1}^k p_i^\beta \right) = \left( \sum_{i=1}^k \gamma_{\pi_{\mathbf{p}}^{-1}(i)}, \sum_{i=1}^k p_{\pi_{\mathbf{p}}^{-1}(i)} \right), \quad (\text{M13})$$

for  $k \in \{1, \dots, d\}$ , where  $\pi_{\mathbf{p}}$  is a permutation that  $\beta$ -orders  $\mathbf{p}$ .

As we will justify in Sec. 1.2.2 of Chapter 1, a probability distribution  $\mathbf{p}$  is said to be further from the thermal equilibrium distribution than  $\mathbf{q}$  is if  $\mathbf{p}$  thermo-majorises  $\mathbf{q}$ . This is defined as follows:

**Definition M8: Thermo-majorisation**

Given a thermal Gibbs distribution  $\gamma$  and probability vectors  $\mathbf{p}$  and  $\mathbf{q}$ , we say that  $\mathbf{p}$  thermo-majorises  $\mathbf{q}$ , denoting it by  $\mathbf{p} \succ_{\beta} \mathbf{q}$ , if the thermo-majorisation curve  $f_{\mathbf{p}}$  is above  $f_{\mathbf{q}}$  everywhere, i.e.,  $f_{\mathbf{p}}(x) \geq f_{\mathbf{q}}(x)$ .

**Remark M5**

In the infinite temperature limit,  $\beta \rightarrow 0$ , the thermo-majorisation ordering of probability distributions is replaced by the majorisation ordering.

# Part I

## Quantum thermodynamics

# Summary of results

Chapter 1 serves as an introduction to the resource theory of thermodynamics, and thus consists mainly of previously known results. After formally introducing the set of free operations, known as *thermal operations*, we analyse their properties and point out that the resulting resource theory can be seen as a hybrid theory of coherence and athermality. The remainder of the chapter is then devoted to the restricted case of states without coherence, which we use to explain important thermodynamic notions and problems. We therefore present the concept of thermodynamic ordering between incoherent states, discuss the role of catalysts and thoroughly analyse the crucial thermodynamic task of work extraction.

In Chapter 2 we proceed to study the allowed thermodynamic transformations of general quantum states, in particular we focus on the thermodynamic limitations of coherence processing. The symmetry-constraints of thermal operations allow us to employ the recently developed tools from the resource theory of asymmetry [48, 49]. Using these tools we show that all quantum states can be decomposed into *mode* operators that quantify the coherence present in a state, and that each such mode transforms independently under thermal operations. We then establish general upper and lower bounds for the evolution of quantum coherence under arbitrary thermal operations, valid for any temperature. Using these bounds we identify primitive coherence manipulations and show that the transfer of coherence between energy levels exhibits thermodynamic irreversibility. Finally, we point out that the traditional Szilard result on the information-to-work conversion rate cannot be directly generalised to quantum states with coherence, which leads us to the following chapter.

In Chapter 3 we study the possibility of exploiting quantum coherence to perform work. Classically, using a thermal bath one can convert the knowledge about a system state  $A$  into the amount of work equal to the free energy of  $A$ . Within quantum formalism,

however, energy conservation enforces a modification of this traditional result: although coherence contributes to the free energy, it remains “locked” and cannot be extracted as work with the use of free operations [9, 13]. In order to overcome this limitation, we introduce an ancillary system with coherence, a *reference*, that allows us to “unlock” the free energy of the system. Through a careful account of all resources involved in the thermodynamic transformations within a fully quantum-mechanical treatment, we then show that there exist references extracting work from coherence arbitrarily well, while not being deteriorated. On the other hand, we show that for any reference with finite resources not all the coherence of a state can be extracted as work. However, we also prove that even bounded references can be reused infinitely many times in the process of work extraction from coherence.

Finally, in Chapter 4 we switch our focus from thermal operations to transformations that leave the thermal equilibrium state unchanged,<sup>1</sup> and study the structure of the induced thermodynamic ordering of states. More precisely, we investigate the structural properties of the partial order defined by reachability of one quantum state from the other via free thermodynamic operations. We analyse this ordering separately for the set of classical (incoherent) states and the full set of quantum states, and point out that in the infinite temperature limit it exhibits a lattice structure in both regimes. This means that when energy does not matter and the only thermodynamic resource is given by information, the thermodynamic arrow of time has a very particular property: for any two states there exists a unique state in the past consistent with considered states at present and with all possible joint pasts of those states. Similarly, there also exists a unique state in the future consistent with the considered states at present and with all possible joint futures of those states. We also show how the lattice structure in the classical regime is broken into parts at finite temperatures, i.e., when energy is a relevant thermodynamic resource. On the other hand, we prove that in the simplest scenario of a two-dimensional system, within the full quantum state space (when we include states with coherence) the lattice structure is preserved at finite temperatures.

---

<sup>1</sup>Note that the set of such transformations forms a superset of thermal operations.

# Chapter 1

## Resource theory approach to thermodynamics

In this work we will investigate the thermodynamics of finite dimensional systems of dimension  $d$  in the presence of a single heat bath at inverse temperature  $\beta = 1/k_B T$ , where  $k_B$  denotes the Boltzmann constant. The space of quantum states is then given by the set of density operators  $\mathcal{S}_d$ , and the most general evolution of a quantum state  $\rho$  of a system that is initially uncorrelated with its environment is given by a *completely positive trace-preserving* (CPTP) map  $\mathcal{E}$ , also known as a quantum channel. However, due to thermodynamic constraints encoded by the laws of thermodynamics, not all state transformations are thermodynamically allowed. For example, a pure ground state cannot be transformed into a pure excited state for free (due to the first law of thermodynamics), or a system cannot be taken out of thermal equilibrium without an additional cost (because of the second law of thermodynamics). The laws of thermodynamics only have to hold globally and thus locally we can observe their violation, e.g., using external source of work one can decrease the entropy of a system. In order to study fundamental limitations of thermodynamics, though, we assume that the system under investigation and its thermal environment form a closed system. This way any external sources of energy, negentropy or other thermodynamic resources (like coherence, as we shall see in Chapter 2) have to be explicitly modelled by ancillary systems.

In order to define a resource theory of thermodynamics we need to identify the free states and free operations of the theory. A state of a system that is in equilibrium with a thermal bath at inverse temperature  $\beta$  is considered to be a free state. Such a state is not

a resource, as according to the second law of thermodynamics it cannot be transformed into any other (out-of-equilibrium) state for free. As we will argue in Sec. 1.2.4, allowing for any other free state would trivialise the theory by, e.g., allowing one to perform infinite amount of work for free. Therefore, for a bath at given inverse temperature  $\beta$ , the only free state of the system described by a Hamiltonian  $H$  is given by the *thermal Gibbs state*,

$$\gamma = \frac{e^{-\beta H}}{Z}, \quad Z = \text{Tr} (e^{-\beta H}). \quad (1.1)$$

The set of free operations that is the main focus of Part I of this thesis is the set of *thermal operations* (TOs) [6, 8, 10, 50]. However, before we explicitly define it, let us make a few general comments that are independent of this specific choice. A set of free thermodynamic operations encodes the structure of the thermodynamic arrow of time by telling us which states can be reached from a given state (and which states can evolve into it) in accordance with the laws of thermodynamics. Specifically, we can introduce the following:

**Definition 1.1: Thermal cones**

The set of states  $\mathcal{T}_+(\rho)$  that a quantum state  $\rho$  can be mapped to via free thermodynamic operations is called the *future thermal cone* of  $\rho$ . The set of states  $\mathcal{T}_-(\rho)$  that can be mapped to  $\rho$  via free thermodynamic operations is called the *past thermal cone* of  $\rho$ .

Thermal cones induce ordering along the thermodynamic arrow of time within the state space, see Fig. 1.1. We will discuss its structure in more detail in Chapter 4. In order to characterise this ordering in a quantitative way one can introduce *thermodynamic monotone* functions that tell us how far a given state is from being a free (thermal equilibrium) state. As free operations, by definition, cannot increase the amount of resources present in a state, we have the following definition:

**Definition 1.2: Thermodynamic monotone**

A function  $\phi$  mapping the set of quantum states  $\mathcal{S}_d$  to non-negative real numbers  $\mathbb{R}_+ \cup \{0\}$  is a thermodynamic monotone if and only if:

1.  $\phi(\mathcal{E}(\rho)) \leq \phi(\rho)$  for all states  $\rho$  and all free thermodynamic operations  $\mathcal{E}$ .
2.  $\phi(\gamma) = 0$ .

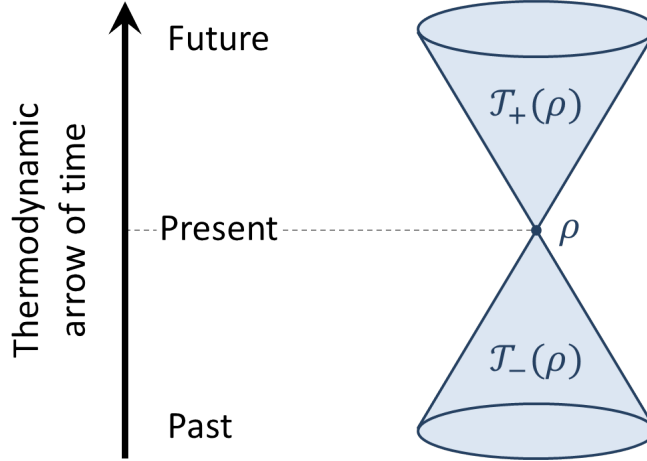


Figure 1.1: *Thermodynamic arrow of time*. The reachability of one state from another via free thermodynamic operations introduces the ordering of states along the thermodynamic arrow of time. States that can be reached from a given state  $\rho$  form its future thermal cone  $\mathcal{T}_+(\rho)$ , whereas states that can be transformed into  $\rho$  form its past thermal cone  $\mathcal{T}_-(\rho)$ .

A function  $\phi$  such that  $\phi(\rho) = 0$  for all  $\rho$  is an example of a trivial thermodynamic monotone. A non-trivial example, the operational meaning of which will be explained in Sec. 1.2.4 and investigated in Chapter 3, is given by the following:

### Example 1.1: Free energy monotone

Quantum relative entropy between a given state  $\rho$  and a thermal Gibbs state is a thermodynamic monotone,  $\phi(\rho) = S(\rho||\gamma)$ , because for every free operation  $\mathcal{E}$  we have

$$\phi(\rho) = S(\rho||\gamma) \geq S(\mathcal{E}(\rho)||\mathcal{E}(\gamma)) = S(\mathcal{E}(\rho)||\gamma) = \phi(\mathcal{E}(\rho)),$$

where the first inequality comes from the fact that quantum relative entropy satisfies the data processing inequality, and the second equality from the fact that free operations must map a set of free states into itself. We can now write down  $S(\rho||\gamma)$  explicitly,

$$S(\rho||\gamma) = -\text{Tr}(\rho \ln \gamma) + \text{Tr}(\rho \ln \rho) = \beta \text{Tr}(\rho H_S) - S(\rho) + \ln Z, \quad (1.2)$$

where  $S(\cdot)$  is the von Neumann entropy. We thus see that  $S(\rho||\gamma)/\beta$  is a generalised free energy difference between  $\rho$  and a thermal Gibbs state  $\gamma$  (compare with the

classical expression  $F = U - TS$ , where  $F$  is the free energy,  $U$  is the average energy,  $S$  is the thermodynamic entropy, and recall that the free energy of an equilibrium state is equal to  $-\ln Z/\beta$ .

## 1.1 Thermal operations

Thermal operations are defined with the use of the following general thermodynamic setting. A quantum system, previously isolated and characterised by a Hamiltonian  $H$ , is brought into thermal contact with a bath described by a Hamiltonian  $H_B$ . The joint system evolves unitarily and after some time the system is decoupled from the bath. The only assumption made is that this interaction conserves energy overall (note that heat can still flow from and to the bath), according to the first law of thermodynamics. Mathematically this can be formalised in the following way:

### Definition 1.3: Thermal operations

The set of thermal operations  $\{\mathcal{E}_T\}$  consists of CPTP maps that act on a system  $\rho$  in the following way:

$$\mathcal{E}_T(\rho) = \text{Tr}_B (U(\rho \otimes \gamma_B)U^\dagger), \quad (1.3)$$

where  $U$  is a joint unitary commuting with the total Hamiltonian of the system and bath,  $[U, H + H_B] = 0$ , and  $\gamma_B$  is a thermal Gibbs state of the bath at some fixed inverse temperature  $\beta$ ,  $\gamma_B = e^{-\beta H_B} / \text{Tr} (e^{-\beta H_B})$ .

More generally the partial trace in Eq. (1.3) can be performed over any subsystem. However, within this thesis we will only focus on the case where the input and output spaces of a TO are the same.

Thermal operations have two crucial properties that we will extensively use to investigate the structure of the resource theory of thermodynamics. These are:

1.  $\{\mathcal{E}_T\}$  preserve the Gibbs state,

$$\mathcal{E}_T(\gamma) = \gamma. \quad (1.4)$$

2.  $\{\mathcal{E}_T\}$  are time-translation covariant [49], i.e.,

$$\mathcal{E}_T (e^{-iHt} \rho e^{iHt}) = e^{-iHt} \mathcal{E}_T(\rho) e^{iHt}. \quad (1.5)$$

The first property incorporates the core physical principle of the second law of thermodynamics: the non-existence of a machine able to run a cycle in which thermal energy is converted into work. The second property reflects energy conservation, a consequence of the first law, and the fact that the thermal bath is an incoherent mixture of energy states. We will now comment on these two properties in more detail.

### 1.1.1 Superset of Gibbs-preserving operations

Intuitively, based on the second law of thermodynamics, the thermodynamic arrow of time enforces all states to evolve towards the thermal equilibrium state  $\gamma$ . As the set of free states must be mapped under free operations into itself, these operations should preserve  $\gamma$ . The property of having a fixed point  $\gamma$  is thus a minimal requirement on any set of free thermodynamic operations, and Eq. (1.4) ensures that one cannot bring a thermal state out of equilibrium at no work cost. Indeed, if this was not the case we could equilibrate it back and extract work, thus constructing a perpetuum mobile of the second kind and breaking the second law of thermodynamics. We will show how such perpetuum mobile could be constructed in Sec. 1.2.4. Here, we conclude that the maximal set of operations that are consistent with the thermodynamic arrow of time must preserve the Gibbs state and is defined as follows:

#### Definition 1.4: Gibbs-preserving operations

We say that a CPTP map  $\mathcal{E}$  is *Gibbs-preserving*, or GP for short, if the thermal Gibbs state  $\gamma$  is its fixed point:  $\mathcal{E}(\gamma) = \gamma$ .

Although GP operations may not have as clear operational interpretation as thermal operations (whose explicit Stinespring dilation gives us an algorithm to realise them), they are conceptually an interesting object to investigate. On the one hand, a GP condition is a simple requirement of having a particular fixed point, and thus it is much easier to verify whether some operation can be performed with a GP map than to check if it can be performed using a TO. On the other hand, the condition is strong enough to ensure the second law of thermodynamics, so that no work can be performed for free. One can, therefore, study GP operations independently from TOs, as a maximal set of thermodynamically meaningful operations (as we will do in Chapter 4); or, since GP operations form a strict superset of thermal operations, one can use it to find constraints

on what is possible to be achieved with thermal operations (as we will do in Chapter 2). In other words, everything that cannot be realised with the use of GP operations, also cannot be realised with TOs. Moreover, as we will show in Sec. 1.2.1, if one restricts to transformations between states without coherence then, actually, every transformation allowed with the use of GP maps is also achievable via TOs.

### 1.1.2 Superset of time-translation covariant operations

The fact that every thermal operation  $\mathcal{E}_T$  is time-translation covariant, i.e., satisfies Eq. (1.5), has been first highlighted in Ref. [9]. To see this, let us denote the total Hamiltonian of the system and bath by  $H_T = H + H_B$ , and note the following:

$$\begin{aligned} \mathcal{E}_T(e^{-iHt}\rho e^{iHt}) &= \text{Tr}_B(U(e^{-iHt}\rho e^{iHt} \otimes \gamma_B)U^\dagger) = \text{Tr}_B(Ue^{-iH_T t}(\rho \otimes \gamma_B)e^{iH_T t}U^\dagger) \\ &= \text{Tr}_B(e^{-iH_T t}U(\rho \otimes \gamma_B)U^\dagger e^{iH_T t}) = e^{-iHt}\mathcal{E}_T(\rho)e^{iHt}, \end{aligned}$$

where we have used that  $\gamma_B$  is diagonal in the eigenbasis of  $H_B$ ,  $[U, H_T] = 0$  and the fact that trace is cyclic.

The set of time-translation covariant operations is a special case of a more general set of  $G$ -covariant operations, defined for arbitrary group  $G$ .  $G$ -covariant operations form the set of free operations in the resource theory of asymmetry [51], also known as the theory of quantum reference frames [52]. In this theory, states asymmetric with respect to a symmetry group  $G$  are treated as resources under the limitation of the symmetric evolution, i.e., being restricted to symmetric evolution one cannot create an asymmetric state, but having one allows for otherwise forbidden operations. More formally, the set of free states and the set of free operations in the resource theory of asymmetry are given by the following (here we also assume that the input and output spaces of  $\mathcal{E}$  are the same):

**Definition 1.5:  $G$ -invariant states and  $G$ -covariant operations**

Given a group  $G$  and its unitary representation  $\{U(g), \forall g \in G\}$ , a quantum state  $\rho$  is  $G$ -invariant if and only if

$$\forall g \in G : U(g)\rho U^\dagger(g) = \rho. \quad (1.6)$$

CPTP map  $\mathcal{E}$  is said to be  $G$ -covariant if and only if

$$\forall g \in G : \mathcal{E}[U(g)(\cdot)U^\dagger(g)] = U(g)\mathcal{E}(\cdot)U^\dagger(g). \quad (1.7)$$

Similarly to GP operations, the set of time-translation covariant operations is a superset of TOs. To see this note that the Stinespring dilation of a general time-translation covariant map is given by Eq. (1.3), however with the dilating system thermal state  $\gamma_B$  replaced by a general invariant state  $\sigma_B$  satisfying  $[\sigma_B, H_B] = 0$  [53–55]. Hence, all the restrictions on state transformations that are due to covariance structure, will also hold for TOs. Let us then note that a resource theory of thermodynamics can be seen as a hybrid theory coming from two independent restrictions: Gibbs-preserving property and covariance. As we will discuss in Chapter 2, the first restriction is linked with athermality resource (the distance of the distribution over the energy levels from the thermal distribution), whereas the second one is linked with coherence resource (the distance of a state from its dephased version).

## 1.2 Thermodynamics of classical states

Among all the bases of  $\mathcal{S}_d$  the energy eigenbasis  $\{|\epsilon_m\rangle\langle\epsilon_n|\}$ , defined by the eigenstates of the system Hamiltonian  $H = \sum_n \epsilon_n |\epsilon_n\rangle\langle\epsilon_n|$ , is distinguished by the evolution of the system. Namely, states diagonal in this basis evolve trivially in time under the free evolution of the system, i.e., they are time-translation invariant. Such states are usually referred to as *classical states* because their energy is well-defined up to observer’s classical lack of knowledge and they do not exhibit quantum features of superposition. This is in strict contrast to generic quantum states with coherence between energy eigenstates, when even the possession of complete knowledge about a state may leave one with uncertainty about its energy (the classical and quantum origins of uncertainty will be discussed in Chapter 5). Formally, we thus define the following:

### Definition 1.6: Classical states

Consider a  $d$ -dimensional system described by a Hamiltonian  $H = \sum_i \epsilon_i |\epsilon_i\rangle\langle\epsilon_i|$ . A state  $\rho$  of a system will be called classical if it is diagonal in the energy eigenbasis, i.e., if  $\langle\epsilon_i|\rho|\epsilon_j\rangle = 0$  for all  $i \neq j$ . Such a state will be equivalently represented by a  $d$ -dimensional probability vector  $\mathbf{p} \in \mathcal{P}_d$  with  $p_i = \langle\epsilon_i|\rho|\epsilon_i\rangle$ .

The transformations between classical states, which we will also refer to as *incoherent* states, can be represented by a stochastic matrix  $\Lambda$  transforming the diagonal of a state

when represented as a vector (recall that a stochastic matrix describes the most general linear mapping between probability distributions; it satisfies  $\Lambda_{ij} \geq 0$  and  $\sum_i \Lambda_{ij} = 1$ ). Also note that, since the thermal Gibbs state  $\gamma$  is diagonal in the energy eigenbasis, it is a classical state and thus can be represented by a probability vector  $\gamma$ .

We will now present the prior results on the resource theory of thermodynamics restricted to classical states. Note that with such a restriction we only focus on one aspect of the hybrid theory – on athermality. As incoherent states are invariant under time-translations, i.e., they are free states in the resource theory of asymmetry, the covariance restriction plays no role in defining classically restricted thermal cones. Therefore, the results presented in this section can be seen as the “classical” component of the full quantum thermodynamics theory. In the remainder of Part I we will focus on the “quantum” component of the theory that involves the study of general quantum states with coherence.

### 1.2.1 Equivalence of GP and thermal operations

We will start from showing that the thermodynamic ordering of classical states under the restriction of thermal operations is the same as under GP operations. Therefore, by solving the problem of interconversion of states for significantly simpler case of GP maps, we will then obtain classically restricted thermal cones for TOs. The following theorem (in a different, but equivalent form) and the presented proof have been first stated in Ref. [10].

#### **Theorem 1.1: Equivalence of classically restricted thermal cones**

Restricting to classical states, the future thermal cone defined by the set of thermal operations,  $\mathcal{T}_+^T$ , and the future thermal cone defined by GP operations,  $\mathcal{T}_+^{GP}$ , coincide. In other words, there exists a thermal operation  $\mathcal{E}_T$  mapping a classical state  $\rho$  into a classical state  $\sigma$  if and only if there exists a GP operation  $\mathcal{E}_{GP}$  such that  $\mathcal{E}_{GP}(\rho) = \sigma$ .

*Proof.* First of all we know that every thermal operation is a GP operation. Hence, we only need to prove that every state transition achievable via a GP operation can be achieved by some TO. We will prove this by showing that with an appropriate choice of the bath system and energy-conserving unitary interaction, one can construct a thermal

operation whose action on the diagonal is described by an arbitrary GP stochastic matrix.

Consider a system described by a Hamiltonian  $H = \sum_i \epsilon_i |\epsilon_i\rangle\langle\epsilon_i|$  and prepared in a classical state  $\rho = \sum_i p_i |\epsilon_i\rangle\langle\epsilon_i|$ , represented by a probability vector  $\mathbf{p}$ . Also, denote the thermal Gibbs state of the system by  $\gamma = \sum_i \gamma_i |\epsilon_i\rangle\langle\epsilon_i|$ , represented by a probability vector  $\boldsymbol{\gamma}$ . Note that for simplicity we assumed that the system has non-degenerate spectrum, however the extension to the degenerate case is immediate. Now, we take a bath described by the following Hamiltonian

$$H_B = \sum_j E_j \sum_{g=1}^{g(E_j)} |E_j, g\rangle\langle E_j, g|, \quad (1.8)$$

with  $g(E_j)$  denoting the degeneracy of the energy level  $E_j$ . Moreover, we require that the energy spectrum of the bath satisfies the following conditions: for any two energy levels of the system,  $\epsilon_i$  and  $\epsilon_j$ , there exist  $E_k$  and  $E_l$  such that  $\epsilon_i - \epsilon_j = E_k - E_l$ ; and that the degeneracy scales exponentially  $g(E + \epsilon) = g(E)e^{\beta\epsilon}$ . The first requirement is satisfied for a bath with continuous spectrum and can be approached arbitrarily well with energy splittings of the bath being small enough compared to the energy gaps of the system. The second requirement is satisfied when the average energy of the bath is much bigger than that of the system. To see this note that the entropy of a fixed energy shell  $E$  of the bath is defined by  $S(E) := \ln g(E)$  and for  $\epsilon \ll E$  we have:

$$g(E + \epsilon) = e^{S(E+\epsilon)} \approx e^{S(E) + \epsilon \frac{\partial S(E)}{\partial E}} = g(E)e^{\beta\epsilon}, \quad (1.9)$$

where the inverse temperature is traditionally defined by  $\beta = \frac{\partial S(E)}{\partial E}$  [56].

The thermal Gibbs state of the bath is given by

$$\gamma_B = \frac{1}{Z_B} \sum_j e^{-\beta E_j} \sum_{g=1}^{g(E_j)} |E_j, g\rangle\langle E_j, g|, \quad (1.10)$$

where  $Z_B = \sum_j e^{-\beta E_j} g(E_j)$ . We thus have that the joint system-bath state  $\rho \otimes \gamma_B$  has the following form

$$\sum_i \sum_j \sum_{g=1}^{g(E_j)} p_i \frac{e^{-\beta E_j}}{Z_B} |\epsilon_i\rangle\langle\epsilon_i| \otimes |E_j, g\rangle\langle E_j, g| = \sum_E \sum_i \sum_{g=1}^{g(E-\epsilon_i)} p_i \frac{e^{-\beta(E-\epsilon_i)}}{Z_B} |E, \epsilon_i, g\rangle\langle E, \epsilon_i, g|,$$

where we have changed the summation: instead of independently summing over the energies of the system,  $\epsilon_i$ , and bath,  $E_j$ , we sum over the total energy  $E = \epsilon_i + E_j$  and the energy of the system  $\epsilon_i$ . By introducing normalisation factors  $N(E) = e^{-\beta E} g(E)/Z_B$  and

multiplicity factors  $D_i = g(E)e^{-\beta\epsilon_i}$ , we can rewrite the above as follows:

$$\rho \otimes \gamma_B = \sum_E N(E) |E\rangle\langle E| \otimes \sum_i \sum_{g=1}^{D_i} \frac{p_i}{D_i} |\epsilon_i, g\rangle\langle \epsilon_i, g| := \sum_E N(E) |E\rangle\langle E| \otimes \rho_E,$$

where  $\rho_E$  are normalised states that encode relative degrees of freedom for a fixed total energy  $E$ .

Now, in each fixed energy block  $E$  we have a freedom to perform any unitary  $U_E$ , as the only requirement for  $U$  being energy-preserving is that it has a block-diagonal structure in total energies:  $U = \bigoplus_E U_E$ . Let us focus on one such block with fixed  $E$ . The spectrum of a state  $\rho_E$  within this block can be divided into groups labelled by the energy of the system  $\epsilon_i$ , consisting of  $D_i$  copies of eigenvalues  $p_i/D_i$ . Let us consider unitaries  $U_E$  that are permutations between these groups, moving  $n_{i|j}$  eigenvalues from group  $j$  to group  $i$ . The occupations  $p'_i$  of the system energy states  $\epsilon_i$  for the transformed state,  $\rho'_E = U_E \rho_E U_E^\dagger$ , are given by

$$p'_i = \sum_j n_{i|j} \frac{p_j}{D_j} = \sum_j \frac{n_{i|j}}{D_j} p_j. \quad (1.11)$$

In other words, the probability vector  $\mathbf{p}$  representing the initial classical state is transformed by a stochastic matrix  $\Lambda$  with  $\Lambda_{ij} = \frac{n_{i|j}}{D_j}$ . Hence, ensuring high enough degeneracy of the bath spectrum ( $D_j \gg 1$  for all  $j$ ) we can get arbitrarily close to any matrix element  $\Lambda_{ij}$ , by appropriately choosing the coefficient  $n_{i|j}$ . These coefficients are only constrained by

$$\sum_i n_{i|j} = D_j, \quad \sum_j n_{i|j} = D_i,$$

which is equivalent to

$$\sum_i \Lambda_{ij} = 1, \quad \sum_j \Lambda_{ij} \gamma_j = \gamma_i, \quad (1.12)$$

i.e.,  $\Lambda$  is a GP stochastic matrix. Hence, within the block of fixed total energy  $E$  we can perform any GP map. By making a similar choice of  $U_E$  for every other block, we can obtain a transformation of the diagonal of the initial state described by an arbitrary GP stochastic matrix.  $\square$

## 1.2.2 Thermodynamic ordering

One of the main questions within a resource theory approach is the interconversion problem: given two states,  $\rho$  and  $\sigma$ , we want to know whether there exists a free operation

mapping  $\rho$  to  $\sigma$ . In the resource theory of thermodynamics the question is: given an initial state  $\rho$ , does a state  $\sigma$  lie in the future thermal cone  $\mathcal{T}_+(\rho)$ , in the past thermal cone  $\mathcal{T}_-(\rho)$ , or in neither of them? Here, we will present an answer to this question in the restricted case when both  $\rho$  and  $\sigma$  are classical states, represented by probability vectors  $\mathbf{p}$  and  $\mathbf{q}$ . Using Theorem 1.1, instead of looking at TOs, we will consider the set of free operations to be given by GP operations acting on incoherent states as stochastic GP matrices. Further, in Chapter 2, we will investigate additional restrictions that come from time-translation covariance of TOs when one considers transformations between generic quantum states.

Let us start with the infinite temperature limit that can also be seen as an information-theoretic limit of thermodynamics. This is because when the inverse temperature  $\beta \rightarrow 0$ , the thermal Gibbs state is described by a uniformly distributed probability vector  $\boldsymbol{\eta} := (1/d, 1/d, \dots, 1/d)$ . Hence, in this limit we effectively deal with energy states all being degenerate and the negentropy (or information) is the only thermodynamic resource. The set of GP transformations acting on classical states is then given by bistochastic matrices  $\{\Lambda\}$ , i.e., by stochastic matrices such that  $\Lambda\boldsymbol{\eta} = \boldsymbol{\eta}$ . From Definition M3 and Remark M1 we know that the existence of such  $\Lambda$  mapping  $\mathbf{p}$  to  $\mathbf{q}$  is equivalent to  $\mathbf{p} \succ \mathbf{q}$ , where  $\succ$  denotes majorisation [44]. We thus obtain that, in the infinite temperature limit, a classical state  $\mathbf{p}$  can be transformed into  $\mathbf{q}$  via a TO if and only if  $\mathbf{p} \succ \mathbf{q}$ .

We will now show how to use the above result for the infinite temperature limit to understand thermodynamic ordering of incoherent states at finite temperatures. In order to do this we will use the notion of an *embedding map*. To define it, we first need to approximate the entries of the Gibbs state  $\boldsymbol{\gamma}$  with rational numbers (the irrational values can be approached with arbitrarily high accuracy):

$$\boldsymbol{\gamma} = \left( \frac{D_1}{D}, \dots, \frac{D_d}{D} \right), \quad D = \sum_{i=1}^d D_i. \quad (1.13)$$

We now define the following transformation:

**Definition 1.7: Embedding map**

The embedding map  $\Gamma$  is sending a  $d$ -dimensional probability vector,  $\mathbf{p} \in \mathcal{P}_d$ , to a

$D$ -dimensional one,  $\Gamma(\mathbf{p}) \in \mathcal{P}_D$ , in the following way:

$$\Gamma(\mathbf{p}) = \left( \underbrace{\frac{p_1}{D_1}, \dots, \frac{p_1}{D_1}}_{D_1 \text{ times}}, \underbrace{\frac{p_2}{D_2}, \dots, \frac{p_2}{D_2}}_{D_2 \text{ times}}, \dots, \underbrace{\frac{p_d}{D_d}, \dots, \frac{p_d}{D_d}}_{D_d \text{ times}} \right). \quad (1.14)$$

**Remark 1.2**

The embedding map sends a thermal Gibbs state  $\gamma$  to a maximally mixed state:

$$\Gamma(\gamma) = \boldsymbol{\eta} = (1/D, \dots, 1/D).$$

By writing down the embedding matrix  $\Gamma$  explicitly, we can also easily find its left inverse  $\Gamma^{-1}$ :

$$\Gamma = \begin{pmatrix} \frac{1}{D_1} & 0 & \dots & 0 \\ \vdots & \vdots & \dots & 0 \\ \frac{1}{D_1} & 0 & \dots & 0 \\ 0 & \frac{1}{D_2} & \dots & 0 \\ \vdots & \vdots & \dots & 0 \\ 0 & \frac{1}{D_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{D_d} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{D_d} \end{pmatrix}, \quad \Gamma^{-1} = \begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 1 \end{pmatrix},$$

where in each column  $i$  of  $\Gamma$  and in each row  $i$  of  $\Gamma^{-1}$  there is  $D_i$  non-zero entries. We now note that the following statements are equivalent:

$$\Lambda \mathbf{p} = \mathbf{q} \Leftrightarrow \Gamma \Lambda \mathbf{p} = \Gamma \mathbf{q} \Leftrightarrow \Gamma \Lambda \Gamma^{-1} \Gamma \mathbf{p} = \Gamma \mathbf{q},$$

which results in

$$\Lambda \mathbf{p} = \mathbf{q} \Leftrightarrow \tilde{\Lambda} \tilde{\mathbf{p}} = \tilde{\mathbf{q}}, \quad (1.15)$$

where  $\tilde{\Lambda} = \Gamma \Lambda \Gamma^{-1}$  is the embedded version of  $\Lambda$ , whereas  $\tilde{\mathbf{p}} = \Gamma \mathbf{p}$  and  $\tilde{\mathbf{q}} = \Gamma \mathbf{q}$  are the embedded versions of  $\mathbf{p}$  and  $\mathbf{q}$ . The last thing required to use the infinite temperature result to study ordering at finite temperatures is the following lemma concerning embedded GP stochastic matrices:

**Lemma 1.2: Embedded GP stochastic matrix**

Given a Gibbs-preserving stochastic matrix  $\Lambda$  acting on  $\mathcal{P}_d$  ( $\Lambda\gamma = \gamma$ ), an embedded matrix  $\tilde{\Lambda} = \Gamma\Lambda\Gamma^{-1}$  is a bistochastic matrix acting on  $\mathcal{P}_D$ :  $\tilde{\Lambda}\eta = \eta$ .

Before presenting the proof let us state that as a result of the above lemma, together with Eq. (1.15) and Remark M1, a state  $\mathbf{p}$  can be mapped to  $\mathbf{q}$  via a GP transformation if and only if the embedded probability vector  $\tilde{\mathbf{p}}$  majorises  $\tilde{\mathbf{q}}$ .

*Proof.* The matrix elements of  $\tilde{\Lambda}$  are given by

$$\tilde{\Lambda}_{ij} = \sum_{k,l=1}^d \Gamma_{ik} \Lambda_{kl} \Gamma_{lj}^{-1},$$

so the conditions for bistochasticity of  $\tilde{\Lambda}$  yield

$$\forall i : \sum_{j=1}^D \sum_{k,l=1}^d \Gamma_{ik} \Lambda_{kl} \Gamma_{lj}^{-1} = 1, \quad \forall j : \sum_{i=1}^D \sum_{k,l=1}^d \Gamma_{ik} \Lambda_{kl} \Gamma_{lj}^{-1} = 1. \quad (1.16)$$

Using the explicit forms of  $\Gamma$  and  $\Gamma^{-1}$  we note the following. First, for all  $l$  we have  $\sum_j \Gamma_{lj}^{-1} = D_l$ . And also, for all  $k$  we have  $\sum_i \Gamma_{ik} = 1$ . Thus, we can simplify the conditions specified by Eq. (1.16) to obtain the following:

$$\forall i : \sum_{k,l=1}^d \Gamma_{ik} \Lambda_{kl} D_l = 1, \quad \forall j : \sum_{k,l=1}^d \Lambda_{kl} \Gamma_{lj}^{-1} = 1.$$

Finally, taking into account that for a fixed  $i$  and  $j$  there is just one non-zero element of  $\Gamma$ ,  $\Gamma_{ik} = 1/D_k$ , and one non-zero element of  $\Gamma^{-1}$ ,  $\Gamma_{lj}^{-1} = 1$ , we get

$$\sum_{l=1}^d \Lambda_{kl} D_l = D_k, \quad \sum_{k=1}^d \Lambda_{kl} = 1. \quad (1.17)$$

The first condition is fulfilled because  $\Lambda$  is preserving a Gibbs state (recall that  $\gamma_l = D_l/D$ ), and the second condition is fulfilled because  $\Lambda$  is a stochastic matrix.  $\square$

Another way to quickly verify that a uniform distribution  $\eta$  is a fixed point of an embedded GP stochastic map (but not that such a map is stochastic), is by the following string of equalities:

$$\tilde{\Lambda}\eta = \Gamma\Lambda\Gamma^{-1}\eta = \Gamma\Lambda\Gamma^{-1}\Gamma\gamma = \Gamma\Lambda\gamma = \Gamma\gamma = \eta. \quad (1.18)$$

We will now show that majorisation between embedded vectors coincides exactly with the notion of *thermo-majorisation* (see Definition M8). Thus, the main result concerning

the interconversion of classical states under GP operations (so also, using Theorem 1.1, under thermal operations) can be phrased as follows:

**Theorem 1.3: Thermodynamic ordering under GP maps**

At finite temperatures a classical state  $\mathbf{p}$  can be mapped via a GP stochastic map  $\Lambda$  into  $\mathbf{q}$  if and only if  $\mathbf{p}$  thermo-majorises  $\mathbf{q}$ :  $\mathbf{p} \succ_{\beta} \mathbf{q}$  [10, 57].

*Proof.* We know that there exists a GP transformation  $\Lambda$  transforming  $\mathbf{p}$  into  $\mathbf{q}$  if and only if  $\tilde{\mathbf{p}} \succ \tilde{\mathbf{q}}$ . This is equivalent to saying that

$$\sum_{i=1}^k \tilde{p}_i^{\downarrow} \geq \sum_{i=1}^k \tilde{q}_i^{\downarrow}, \quad (1.19)$$

for every  $k \in \{1, \dots, D\}$ . Now note that rearranging an embedded vector in a non-increasing order is equivalent to the  $\beta$ -ordering of the original vector (see Definition M6). More precisely, the vector  $\tilde{\mathbf{p}}^{\downarrow}$  consists of groups of entries equal to  $p_i/D_i$  arranged in a non-increasing order, each group consisting of  $D_i$  elements. If one replaced each group with its sum, equal to  $p_i$ , then one would obtain a  $\beta$ -ordered version of  $\mathbf{p}$ , a vector  $\mathbf{p}^{\beta}$ . Satisfying Eq. (1.19) can be equivalently seen as a curve  $f_{\tilde{\mathbf{p}}}$ , consisting of points  $(k, \sum_{i=1}^k \tilde{p}_i^{\downarrow})$ , lying above a curve  $f_{\tilde{\mathbf{q}}}$ , consisting of points  $(k, \sum_{i=1}^k \tilde{q}_i^{\downarrow})$ . However, as the embedded vector  $\tilde{\mathbf{p}}$  consists of groups of repeated elements, the corresponding curve is composed of linear segments connecting the extremal points  $(\sum_{i=1}^k D_i^{\beta}, \sum_{i=1}^k p_i^{\beta})$ , where  $D_i^{\beta}$  is a vector  $(D_1, \dots, D_d)$  rearranged according to  $\beta$ -ordering of  $\mathbf{p}$ . Analogously, a curve  $f_{\tilde{\mathbf{q}}}$  is made of linear segments connecting the points  $(\sum_{i=1}^k D_i^{\beta}, \sum_{i=1}^k q_i^{\beta})$ , where now  $D_i^{\beta}$  is rearranged according to  $\beta$ -ordering of  $\mathbf{q}$ . Recalling that  $\gamma_i = D_i/D$ , we see that  $f_{\tilde{\mathbf{p}}}$  and  $f_{\tilde{\mathbf{q}}}$  correspond exactly to the thermo-majorisation curves of  $\mathbf{p}$  and  $\mathbf{q}$  (see Definition M7). Therefore,  $\mathbf{p} \succ_{\beta} \mathbf{q}$  is equivalent to  $\tilde{\mathbf{p}} \succ \tilde{\mathbf{q}}$  and means that there exists a stochastic GP map transforming  $\mathbf{p}$  into  $\mathbf{q}$ .  $\square$

The equivalence of embedded majorisation and thermo-majorisation can be illustrated with the following example:

**Example 1.3: Embedded majorisation and thermo-majorisation**

Consider the following thermal Gibbs distribution  $\gamma$  and two probability vectors  $\mathbf{p}$

and  $\mathbf{q}$ :

$$\boldsymbol{\gamma} = \left( \frac{3}{6}, \frac{2}{6}, \frac{1}{6} \right), \quad \mathbf{p} = \left( \frac{1}{2}, \frac{3}{8}, \frac{1}{8} \right), \quad \mathbf{q} = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{3} \right).$$

The embedding map  $\Gamma$  sends  $\mathbf{p}$  and  $\mathbf{q}$  into

$$\tilde{\mathbf{p}} = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{3}{16}, \frac{3}{16}, \frac{1}{8} \right), \quad \tilde{\mathbf{q}} = \left( \frac{1}{18}, \frac{1}{18}, \frac{1}{18}, \frac{1}{12}, \frac{1}{12}, \frac{2}{3} \right),$$

which, after rearranging in a nonincreasing order, results in

$$\tilde{\mathbf{p}}^\downarrow = \left( \overbrace{\frac{3}{16}, \frac{3}{16}}^{\frac{3}{8}}, \overbrace{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}}^{\frac{1}{2}}, \frac{1}{8} \right), \quad \tilde{\mathbf{q}}^\downarrow = \left( \frac{2}{3}, \overbrace{\frac{1}{12}, \frac{1}{12}}^{\frac{1}{6}}, \overbrace{\frac{1}{18}, \frac{1}{18}, \frac{1}{18}}^{\frac{1}{6}} \right).$$

At the same time the  $\beta$ -ordering of vectors  $\mathbf{p}$  and  $\mathbf{q}$  yields

$$\mathbf{p}^\beta = \left( \frac{3}{8}, \frac{1}{2}, \frac{1}{8} \right), \quad \mathbf{q}^\beta = \left( \frac{2}{3}, \frac{1}{6}, \frac{1}{6} \right),$$

so that  $\mathbf{q} \succ_\beta \mathbf{p}$ , and it is equivalent to  $\tilde{\mathbf{q}} \succ \tilde{\mathbf{p}}$ .

### 1.2.3 Catalysts

Within a resource theory every state that is not free has to be modelled explicitly in order to avoid consuming resources that are not accounted for. However, there exists a possibility that an ancillary system is present and aids a transformation, but its resources are not used up. This phenomenon is known as *catalysis*: a resource state allows us to perform otherwise impossible transformations, but it is not modified itself, i.e., at the end of the process it is returned unchanged. Within a thermodynamic scenario such an ancillary system can be thought of as a thermal machine that works in a cyclic process. Within the field of quantum information catalysis was first discovered while studying entanglement theory [58]. In particular, it has been shown that it may happen that neither  $\mathbf{p}$  majorises  $\mathbf{q}$ , nor  $\mathbf{q}$  majorises  $\mathbf{p}$ , but there exists  $\mathbf{r}$  such that

$$\mathbf{p} \otimes \mathbf{r} \succ \mathbf{q} \otimes \mathbf{r}, \tag{1.20}$$

where  $\mathbf{r}$  is considered a catalyst. When such a catalytic majorisation occurs we say that  $\mathbf{p}$  *trumps*  $\mathbf{q}$  and denote it  $\mathbf{p} \succ_T \mathbf{q}$  (see Definition M4). From Remark M2 we know that for  $\mathbf{p}^\downarrow \neq \mathbf{q}^\downarrow$ , the probability distribution  $\mathbf{p}$  trumps  $\mathbf{q}$  if and only if all the  $\alpha$ -Rényi entropies (see Definition M1) of  $\mathbf{p}$  are smaller than for  $\mathbf{q}$  [46, 47].

Since the concepts of majorisation and thermo-majorisation are closely related, one can expect that a similar result holds in the resource theory of thermodynamics, where probability distributions  $\mathbf{p}, \mathbf{q}, \mathbf{r}$  represent classical states. Indeed, the authors of Ref. [8] performed the analysis, based mainly on the use of the embedding map, and found the following. The existence of a classical catalyst  $\mathbf{r}$  that allows for a transformation of a classical state  $\mathbf{p} \otimes \mathbf{r}$  into  $\mathbf{q} \otimes \mathbf{r}$  via TO is equivalent to

$$\forall \alpha \geq 0 : \quad S_\alpha(\mathbf{p}||\boldsymbol{\gamma}) \geq S_\alpha(\mathbf{q}||\boldsymbol{\gamma}), \quad (1.21)$$

where  $S_\alpha(\mathbf{p}||\boldsymbol{\gamma})$  is  $\alpha$ -Rényi divergence between a state  $\mathbf{p}$  and a thermal Gibbs state  $\boldsymbol{\gamma}$  (see Definition M2). Note that, unlike in Remark M2, here we only require the above equation to hold for  $\alpha \geq 0$  and not for  $\alpha \in \mathbb{R}$ . This results from the fact that the authors of Ref. [8] showed that the conditions for  $\alpha < 0$  can be relaxed if one is allowed to invest arbitrary small amount of work in the process. Let us notice that Eq. (1.21) tells us that  $S_\alpha(\mathbf{p}||\boldsymbol{\gamma})$  is nonincreasing under thermal operations and, since it is equal to 0 for  $\mathbf{p} = \boldsymbol{\gamma}$ , it is a thermodynamic monotone. Therefore, we see that a necessary condition of nonincreasing of all thermodynamic monotones is also a sufficient one for thermal operations with catalysts.

### 1.2.4 Work extraction and the Szilard engine

After describing the thermodynamic ordering between classical states we will now focus on another central concept of classical thermodynamics – work. We will first explain the important role of *free energy* in traditional thermodynamic transformations between equilibrium states: acting as a lower/upper bound on the invested/extracted work during the process. We will then argue how this concept can be extended to general (non-equilibrium) classical states, emphasising the energetic value of information. This will lead us to the description of a resource-theoretic treatment of the work extraction process, where we will explain the concepts of average, deterministic and single-shot work extraction protocols. Finally, we will use the results presented in this section to argue for the uniqueness of the thermal Gibbs state as the only free state allowed in the resource theory of thermodynamics. Let us stress that the results presented here only concern classical states, whilst Chapter 3 will be devoted to study the problem of work extraction for general quantum states with coherence.

### Traditional equilibrium result

Traditionally [59], one considers a system in contact with a thermal equilibrium bath and asks: how much work is performed on a system when some external parameters (e.g., volume or a magnetic field) are being changed in time? Let us denote by  $\mathbf{z} = (\mathbf{q}, \mathbf{p})$  a point in the position-momentum phase space of our classical system, and by  $H(\mathbf{z}, \lambda)$  its Hamiltonian, with the value of  $\lambda \in [0, 1]$  parametrising the path  $\Gamma$  in the external parameter space. Without loss of generality we may assume a constant switching rate  $\dot{\lambda} = 1/\tau$ , where  $\tau$  denotes the total time of the process. Now, if the initial state is given by  $\mathbf{z}_0$  then the work performed on the system is defined by its energy change along the trajectory:

$$W_{\mathbf{z}_0} := \int_0^\tau dt \dot{\lambda} \frac{\partial}{\partial \lambda} H(\mathbf{z}(t), \lambda). \quad (1.22)$$

A system at thermal equilibrium is described by the thermal Gibbs distribution  $p(\mathbf{z}, \lambda) = e^{-\beta H(\mathbf{z}, \lambda)} / Z(\lambda)$ , where  $Z(\lambda)$  is the partition function given by

$$Z(\lambda) = \int d\mathbf{z} e^{-\beta H(\mathbf{z}, \lambda)}. \quad (1.23)$$

Hence, for a system initially at equilibrium the average amount of work is given by:

$$\langle W \rangle = \int d\mathbf{z}_0 \frac{e^{-\beta H(\mathbf{z}_0, 0)}}{Z(0)} W_{\mathbf{z}_0}. \quad (1.24)$$

Let us now consider the limit of infinitely slow switching,  $\tau \rightarrow \infty$ , when the system is approximately at thermal equilibrium at all times. The average work performed on the system is then given by

$$\langle W \rangle = \int_0^1 d\lambda \int d\mathbf{z} \frac{e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} \frac{\partial}{\partial \lambda} H(\mathbf{z}, \lambda), \quad (1.25)$$

which can be rewritten in the following way:

$$\begin{aligned} \langle W \rangle &= \int_0^1 d\lambda \int d\mathbf{z} \frac{-\frac{1}{\beta} \frac{d}{d\lambda} (e^{-\beta H(\mathbf{z}, \lambda)})}{Z(\lambda)} = -\frac{1}{\beta} \int_0^1 d\lambda \frac{\frac{d}{d\lambda} \int d\mathbf{z} e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} = -\frac{1}{\beta} \int_0^1 d\lambda \frac{\frac{d}{d\lambda} Z(\lambda)}{Z(\lambda)} \\ &= -\frac{1}{\beta} \int_0^1 d\lambda \frac{d}{d\lambda} \ln Z(\lambda) = -\frac{1}{\beta} (\ln Z(1) - \ln Z(0)) = F(1) - F(0). \end{aligned}$$

Here we have introduced the free energy function  $F(\lambda) = -\ln Z(\lambda)/\beta$ , which in traditional thermodynamics is defined as  $F = U(\lambda) - S(\lambda)/\beta$ , with  $U(\lambda)$  denoting the average energy of the system and  $S(\lambda)$  its entropy. Indeed, the two definitions coincide, as can be quickly

verified:

$$U(\lambda) = \int d\mathbf{z} \frac{e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} H(\mathbf{z}, \lambda) = \frac{-\frac{\partial}{\partial \beta} \int d\mathbf{z} e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} = -\frac{\partial}{\partial \beta} \ln Z(\lambda), \quad (1.26a)$$

$$\frac{S(\lambda)}{\beta} = -\frac{1}{\beta} \int d\mathbf{z} \frac{e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} \ln \left( \frac{e^{-\beta H(\mathbf{z}, \lambda)}}{Z(\lambda)} \right) = U(\lambda) + \frac{\ln Z(\lambda)}{\beta}. \quad (1.26b)$$

We thus see that in the infinitely slow switching process the amount of work performed on the system is equal to the free energy difference between the final and initial state:  $\Delta F := F(1) - F(0)$ . Moreover, it can be shown<sup>1</sup> [59] that for finite switching speed the process becomes irreversible, leading to the dissipation of heat and the following inequality:

$$\langle W \rangle \geq \Delta F. \quad (1.27)$$

Note that when work is performed on the system,  $\langle W \rangle \geq 0$ , then the amount of invested work required for the transformation is bigger or equal to  $\Delta F$ . On the other hand, when the system performs work,  $\langle W \rangle \leq 0$ , then the amount of *extracted* work  $|\langle W \rangle|$ , is bounded by  $|\Delta F|$ , i.e., the free energy change is the upper limit of extracted work. This result can also be seen as one of the manifestations of the second law of thermodynamics.

### Non-equilibrium generalisation

Free energy  $F(\lambda)$  has been originally defined only for states at thermal equilibrium. However, taking into account its operational meaning, one can extend the definition to all non-equilibrium states. More precisely, we will show that the traditional expression  $F(\lambda) = U(\lambda) - S(\lambda)/\beta$ , when appropriately generalised, quantifies the maximum amount of work that can be extracted on average from the system in an out-of-equilibrium state. Instead of considering classical continuous systems, whose states are represented by points  $\mathbf{z} = (\mathbf{q}, \mathbf{p})$  in the phase space, we will now start linking with the resource theoretic framework that deals with finite-dimensional systems. The state space will thus be discrete with states labelled by the set  $\{i\}$ , the Hamiltonian parametrised by some external parameter  $\lambda$  will take the values  $\epsilon_i(\lambda)$ , and the distribution over energy states will be given by a probability vector  $\mathbf{p}$ . This corresponds exactly to a quantum system described by a Hamiltonian  $H(\lambda) = \sum_i \epsilon_i(\lambda) |\epsilon_i(\lambda)\rangle \langle \epsilon_i(\lambda)|$  when we restrict to states diagonal in the

<sup>1</sup>One of the ways to prove this is to use the Jarzynski equality:  $e^{-\beta \Delta F} = \langle e^{-\beta W} \rangle$  [59]. Then, using Jensen's inequality, which states that for any convex function  $f$  and a random variable  $X$  we have  $f(\langle X \rangle) \leq \langle f(X) \rangle$ , one arrives at Eq. (1.27).

energy eigenbasis, with the diagonal given by  $\mathbf{p}$ . The thermal Gibbs state is, as usual, represented by a probability vector  $\boldsymbol{\gamma}(\lambda)$  with  $\gamma_i(\lambda) = e^{-\beta\epsilon_i(\lambda)}/Z(\lambda)$ .

The average energy and entropy of a state  $\mathbf{p}$  are now given by

$$U(\mathbf{p}, \lambda) = \sum_i p_i \epsilon_i(\lambda), \quad S(\mathbf{p}, \lambda) = S(\mathbf{p}) = - \sum_i p_i \ln p_i. \quad (1.28)$$

Hence, we can extend the free energy to non-equilibrium states as follows:

$$F(\mathbf{p}, \lambda) = \sum_i p_i \epsilon_i(\lambda) + \frac{1}{\beta} \sum_i p_i \ln p_i = \frac{1}{\beta} (S(\mathbf{p}||\boldsymbol{\gamma}(\lambda)) - \ln Z(\lambda)), \quad (1.29)$$

where we have used the information-theoretic notion of relative entropy (see Definition M2):

$$S(\mathbf{p}||\boldsymbol{\gamma}(\lambda)) := \sum_i p_i \ln \frac{p_i}{\gamma_i(\lambda)} = \sum_i p_i \ln p_i + \beta \sum_i p_i \epsilon_i(\lambda) + \ln Z(\lambda). \quad (1.30)$$

Since the relative entropy  $S(\mathbf{p}||\boldsymbol{\gamma}(\lambda))$  is always positive and equal to 0 if and only if  $\mathbf{p} = \boldsymbol{\gamma}(\lambda)$ , we see that the thermal equilibrium state is minimising free energy.<sup>2</sup> The claim now is that the maximum amount of work that can be extracted while transforming a system with an initial Hamiltonian  $H(0)$  and in a classical state  $\mathbf{p}$ , into a system described by the final Hamiltonian  $H(1)$  and a classical state  $\mathbf{q}$ , is given by

$$\langle W \rangle = F(\mathbf{p}, 0) - F(\mathbf{q}, 1). \quad (1.31)$$

In a particular case when the initial and final Hamiltonians are the same (an equivalent of a cyclic process), and when we are interested in maximising the extracted work (so that the final state is thermal, achieving the minimum of free energy) we get:

$$\langle W \rangle(\mathbf{p}) = \Delta F(\mathbf{p}) := F(\mathbf{p}, 0) - F(\boldsymbol{\gamma}, 0) = \frac{1}{\beta} S(\mathbf{p}||\boldsymbol{\gamma}). \quad (1.32)$$

To see how this can be achieved, let us use a simple example of a qubit system, which is also illustrated in Fig. 1.2.

#### Example 1.4: Work extraction from incoherent qubit state

Consider a qubit system initially described by a Hamiltonian

$$H(0, \epsilon) = 0 |\epsilon_0\rangle\langle\epsilon_0| + \epsilon |\epsilon_1\rangle\langle\epsilon_1|,$$

<sup>2</sup>A more traditional way to show this is to use Lagrange multipliers: extremise  $F(\mathbf{p})$  with the constraint that  $\mathbf{p}$  is a normalised probability distribution.

and prepared in an excited state  $|\epsilon_1\rangle\langle\epsilon_1|$ , i.e., in a classical state  $\mathbf{p} = (0, 1)$ . The maximum free energy change of the system is then given by:

$$\Delta F(\mathbf{p}) := F(\mathbf{p}, 0, \epsilon) - F(\boldsymbol{\gamma}(0, \epsilon), 0, \epsilon) = \epsilon + \frac{1}{\beta} \ln Z(0, \epsilon).$$

To extract the amount of work given by the above expression one can perform the following steps. First, we need to change the Hamiltonian  $H(0, \epsilon)$  infinitely quickly (or at arbitrary speed, but with the system isolated from the thermal bath) into  $H(0, 0)$ , with the same eigenstates, but degenerate eigenvalues equal to 0. The state of the system will still be given by  $\mathbf{p}$ , but its energy will be decreased by  $\epsilon$ . This is the energetic part of the extracted work (see the upper red panel of Fig. 1.2). Secondly, still in isolation from the bath, we change the Hamiltonian to  $H(\infty, 0)$  with  $\infty$  describing an eigenvalue that is much bigger than  $1/\beta$ . As we have only changed the energy of an unoccupied level and have not transformed the distribution, no energy change occurs in the system. Finally, we connect the system to the thermal bath and infinitely slowly change the Hamiltonian  $H(E, 0)$  from  $E = \infty$  to  $E = \epsilon$ . This yields the following amount of extracted work (compare with Eq. (1.25)):

$$- \int_{\infty}^{\epsilon} \gamma_0(E, 0) dE = \int_{\epsilon}^{\infty} \frac{e^{-\beta E}}{1 + e^{-\beta E}} dE = \frac{1}{\beta} \ln Z(0, E).$$

This is the information part of the extracted work (see the lower blue panel of Fig. 1.2).

As we have seen in the above example, in the presence of a heat bath the information about a system state, even if it is a ground state, has an energetic value. This has already been observed in the seminal work by Szilard [60], who devised an engine that can change 1 bit of information, i.e., information about which of the two degenerate states is occupied, into  $\ln 2/\beta$  amount of work. The link between thermodynamics and information is even more evident in Eq. (1.32), where the amount of work is proportional to the relative entropy that quantifies the amount of information lost when thermal distribution  $\boldsymbol{\gamma}$  is used to approximate the system state  $\mathbf{p}$ .

Finally, the qubit result presented in Example 1.4 can be generalised to arbitrary  $d$ -dimensional systems. In Ref. [11] it has been shown that a sequence of level transformations (while the system is detached from the bath) and full thermalisations leads to

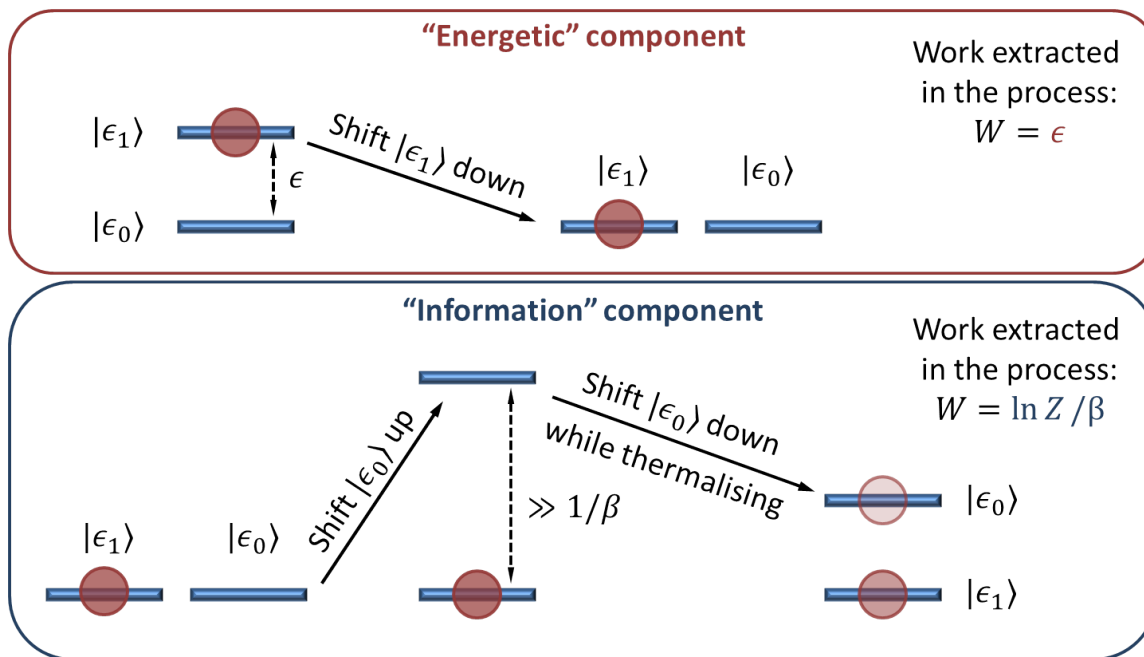


Figure 1.2: Work extraction from an incoherent qubit state as described in Example 1.4.

the extraction of work approaching the free energy change. More precisely, given initial Hamiltonian  $H(0)$ , the corresponding thermal Gibbs state  $\gamma(0)$  and an initial state of the system  $\mathbf{p}$ , one needs to perform the following steps. First, we need to transform the Hamiltonian into  $H(1)$  so that the new thermal state  $\gamma(1)$  coincides with the initial state:  $\gamma(1) = \mathbf{p}$ . Then, we connect the system to the thermal bath and infinitely slowly change the Hamiltonian back to  $H(0)$ . At the end of the process the system will be in the thermal state of the initial Hamiltonian and we will extract  $S(\mathbf{p}||\gamma(0))/\beta$  of work. Note that these are exactly the steps that we have described in Example 1.4, up to the constant shift of the energy spectrum.

### Resource-theoretic treatment and the work system

So far we have discussed the importance of generalised free energy as a way to measure the maximum amount of extractable work from the system. The work extraction protocols described were based on controlling and changing the external parameters defining the Hamiltonian of the system. Within a resource theoretic formulation of thermodynamics, however, we try to avoid using any external systems (so that we do not bring in resources for free) and we want to model everything explicitly. Therefore, it is not immediately clear, whether one can extract the amount of work equal to the free energy change of

the system within a resource theory formalism. We will now briefly discuss how this can be achieved for average work extraction and introduce the concepts of deterministic and single-shot work extractions.

First of all, in order to study the extraction of work from quantum systems we explicitly model the ancillary work body (that we will also call the *battery*), and our aim is to transform it from initially pure energy state to another pure state with higher energy. The work system is usually described by a continuous Hamiltonian  $H_W = \int dw w |w\rangle\langle w|$ , but we can as well choose a Hamiltonian with the discrete spectrum, as long as its energy differences coincide with the amount of work we want to extract. Without loss of generality let us focus on two states of the work body:  $|W\rangle\langle W|$  and  $|0\rangle\langle 0|$ . The possibility of extracting the amount of work equal to  $W$  from a system described by a Hamiltonian  $H^3$  and prepared in a state  $\rho$  is then equivalent to the existence of a thermal operation  $\mathcal{E}_T$  such that:

$$\mathcal{E}_T(\rho \otimes |0\rangle\langle 0|) = \gamma \otimes |W\rangle\langle W|. \quad (1.33)$$

The maximum value of  $W$  for which the above thermal operation exists is the maximum amount of work that can be extracted from  $\rho$ .

The traditional result of Eq. (1.32) has been recovered within a resource-theoretic framework for classical states in two different ways. First, the authors of Ref. [6] considered the extraction of work from  $N$  copies of an incoherent state  $\rho^{\otimes N}$ , with  $\rho$  being represented by a probability vector  $\mathbf{p}$ . They proved that in the asymptotic limit  $N \rightarrow \infty$  there exists a collective thermal operation on the joint system extracting the amount of work approaching  $N \cdot \Delta F(\mathbf{p})$ , i.e., reaching the limit of free energy change per copy. On the other hand, the authors of Ref. [13] have studied thermal operations applied to a single copy of  $\rho$ , however they allowed the work system to be transformed into a mixed state. This way they have shown that the amount of work that can be extracted on average from a state  $\rho$  is given by  $\Delta F(\mathbf{p})$ . By repeating their protocol on  $N \rightarrow \infty$  copies of the system (but, crucially, transforming each copy individually) one ends up with a work system in a mixed state, but whose distribution over energy eigenstates is strongly peaked around  $N \cdot \Delta F(\mathbf{p})$ . Hence, a result of  $\Delta F(\mathbf{p})$  extractable work per copy is recovered in the asymptotic limit.

---

<sup>3</sup>Note that, since we do not allow for any external controls, the Hamiltonian of the system is fixed, and only the state is allowed to change as a result of applying a thermal operation to it.

What if we are not interested in an average work extraction or processing many copies of a system? Then we deal with a concept of deterministic work extraction, when Eq. (1.33) has to be strictly satisfied. To answer how much work can be deterministically extracted, the authors of Ref. [8] used the necessary and sufficient condition for the existence of a catalytic thermal operation between two classical states, specified by Eq. (1.21). Namely, for a two-level work system with a Hamiltonian  $H_W = 0|0\rangle\langle 0| + W|W\rangle\langle W|$  initialised in a state  $|0\rangle\langle 0|$  [represented by probability distribution  $(1, 0)$ ] one has:

$$\forall \alpha \geq 0 : \quad S_\alpha(\mathbf{p} \otimes (1, 0) || \boldsymbol{\gamma} \otimes \boldsymbol{\gamma}_W) \geq S_\alpha(\boldsymbol{\gamma} \otimes (0, 1) || \boldsymbol{\gamma} \otimes \boldsymbol{\gamma}_W),$$

where  $\boldsymbol{\gamma}$  and  $\boldsymbol{\gamma}_W$  denote thermal states of the system and battery, respectively. Taking into account that

$$S_\alpha(\mathbf{p} \otimes \mathbf{p}' || \mathbf{q} \otimes \mathbf{q}') = S_\alpha(\mathbf{p} || \mathbf{q}) + S_\alpha(\mathbf{p}' || \mathbf{q}'),$$

one obtains the following condition:

$$\forall \alpha \geq 0 : \quad S_\alpha(\mathbf{p} || \boldsymbol{\gamma}) \geq \beta W. \quad (1.34)$$

Using the fact that  $S_\alpha(\mathbf{p} || \mathbf{q})$  is non-decreasing with  $\alpha$  one finds that the maximum amount of extractable work is given by

$$W = \frac{1}{\beta} S_0(\mathbf{p} || \boldsymbol{\gamma}) \leq \frac{1}{\beta} S_1(\mathbf{p} || \boldsymbol{\gamma}) = \Delta F(\mathbf{p}). \quad (1.35)$$

Finally, the problem of work extraction can also be studied in a the so-called single-shot regime [11] that lies somewhere in between the average and deterministic work extraction. Extracted work is then seen as a random variable, maximising the average of which yields Eq. (1.32). What one is interested in the single-shot regime is the maximum amount of deterministic (i.e., fluctuation-free) work that can be extracted during a single instance of the protocol, while allowing for the failure probability  $\varepsilon$  (so that in the limit  $\varepsilon = 0$  we recover deterministic work extraction). In Ref. [11] it was shown that for a classical state  $\mathbf{p}$  this quantity is given by

$$W_{ss}^\varepsilon(\mathbf{p}) = F_0^\varepsilon(\mathbf{p}) - F(\boldsymbol{\gamma}), \quad F_0^\varepsilon(\mathbf{p}) = -\frac{1}{\beta} \log Z_\varepsilon, \quad (1.36)$$

where  $F_0^\varepsilon$  is a single-shot free energy defined as follows. Given a subset  $\Lambda$  of the indices  $\{i\}$  labelling the energy levels of the system, we have

$$Z(\Lambda) = \sum_{i \in \Lambda} e^{-\beta \epsilon_i}, \quad Z_\varepsilon := \min\{Z(\Lambda) : \sum_{i \in \Lambda} p_i > 1 - \varepsilon\}, \quad (1.37)$$

where  $\epsilon_i$  are the eigenvalues of  $H$ . This result is also in agreement with other work extraction models based on thermal operations [10].

## The uniqueness of the free thermal state

Within this section we have argued, using several frameworks, that the optimal amount of work one can extract while transforming a classical state  $\mathbf{p}$  into  $\mathbf{q}$  is given by the difference between their corresponding free energies. As a thermal Gibbs state  $\gamma$  minimises the free energy, no work can be extracted from it, and hence it can be treated as a free state. However, if we were to enlarge the set of free states beyond  $\gamma$ , by adding to it any other state  $\mathbf{f}$ , we would trivialise the resource theory of thermodynamics for incoherent states. Indeed, if  $\mathbf{f}$  were free we could bring infinitely many copies,  $N \rightarrow \infty$ , of it and extract the amount of work  $N \cdot \Delta F(\mathbf{f}) \rightarrow \infty$ . This would not only allow us to build a perpetuum mobile (as, using a thermal equilibrium bath, we would be able to get infinite amount of work for free), but would also mean that any incoherent state can be transformed into any other for free. The latter comes from the fact that with enough copies of  $\mathbf{f}$  one can generally satisfy all conditions for catalytic thermo-majorisation (given by Eq. (1.21)).

We thus see that bringing a free state whose distribution over energy eigenstates differs from the thermal distribution would lead to a trivial and unphysical resource theory. However, one could still wonder, whether the set of free states could be enlarged by adding to it a general quantum state  $\rho$  with coherence, such that  $\langle \epsilon_i | \rho | \epsilon_i \rangle = \gamma_i$ . To see that this is not the case and would also bring an unlimited source of free work, let us consider the following example:

### Example 1.5: Work extraction from many copies of $|\gamma\rangle\langle\gamma|$

Consider a qubit system described by a Hamiltonian  $H = \epsilon_0 |\epsilon_0\rangle\langle\epsilon_0| + \epsilon_1 |\epsilon_1\rangle\langle\epsilon_1|$  and in a “coherent Gibbs state”  $|\gamma\rangle\langle\gamma|$ , where

$$|\gamma\rangle = \sqrt{\gamma_0} |\epsilon_0\rangle + \sqrt{\gamma_1} |\epsilon_1\rangle, \quad (1.38)$$

with  $\gamma = (\gamma_0, \gamma_1)$  denoting a thermal Gibbs distribution. If  $|\gamma\rangle\langle\gamma|$  were a free state we could also bring two copies of it:

$$|\gamma\gamma\rangle = \gamma_0 |\epsilon_0\epsilon_0\rangle + \gamma_1 |\epsilon_1\epsilon_1\rangle + \sqrt{2\gamma_0\gamma_1} \frac{|\epsilon_0\epsilon_1\rangle + |\epsilon_1\epsilon_0\rangle}{\sqrt{2}}.$$

However, now within a fixed energy subspace spanned by  $\{|\epsilon_0\epsilon_1\rangle, |\epsilon_1\epsilon_0\rangle\}$  any unitary  $U$  is a thermal operation, hence in particular we can perform the following

transformation:

$$U |\gamma\gamma\rangle = \gamma_0 |\epsilon_0\epsilon_0\rangle + \gamma_1 |\epsilon_1\epsilon_1\rangle + \sqrt{2\gamma_0\gamma_1} |\epsilon_0\epsilon_1\rangle.$$

After dephasing (which is also a thermal operation), the above state becomes classical and is described by a non-thermal distribution  $(\gamma_0^2, 2\gamma_0\gamma_1, 0, \gamma_1^2)$ . Since one can create such a state for free, one can also extract arbitrary amount of work for free.

We thus see that the reason for which a state with thermal distribution over energy eigenstates, but with coherence, cannot be a free state is the following. When bringing multiple copies of such a state we obtain a joint state that is not diagonal in fixed energy subspaces. This in turn allows us to modify this distribution at no cost using energy-preserving unitaries, resulting in a non-thermally distributed final state.

## Chapter 2

# Thermal operations and modes of coherence

The resource theory of thermodynamics was developed to study thermal processes in the quantum regime. However, the results we have presented so far only concerned classical states – we actually dealt with stochastic processes transforming classical probability distributions. Therefore, we have not investigated the effects of quantum interference coming from the principle of superposition, i.e., we have not studied thermodynamic transformations between states with coherence. In fact, the only scenario involving coherence we looked at, Example 1.5, indicates it is likely to be very interesting. In this chapter we will study this problem and focus on the limitations of general quantum state transformations under the restriction of thermal operations. We will thus study the fully quantum theory of thermodynamics, as a general state will be both a resource of athermality and of coherence. On the other hand, our results can also be seen as the analysis of possible coherence manipulations under thermodynamically free operations.

We will use the time-translation covariant structure of thermal operations, that comes from the underlying energy-conservation within thermodynamics, to constrain free thermodynamic transformations of a given state, i.e., to bound the future thermal cone of a general quantum state. The covariant structure of TOs allows us to make use of harmonic analysis techniques, developed in Ref. [48], to track the evolution of the system under thermodynamic transformations in terms of *modes of coherence* of the system state. This constitutes a natural framework to understand coherence, thus allowing us to separate out the constraints that stem solely from symmetry arguments from those particular

to thermodynamics. The approach presented in this chapter also implies that the existing thermo-majorisation results for classical states, discussed in Chapter 1, can be viewed as a particular case of the symmetry-based analysis when only the zero-mode is present. Beyond this classical regime we will show that every non-zero mode obeys independent constraints and that a rich dynamics is allowed, in which coherence can be transferred among different energy levels within a given mode. Moreover, similarly to heat flows, coherence flows show directionality due to the limitations imposed by the second law, so that this new kind of irreversibility adds up to the ones identified in work extraction [10] and coherence distillation [9].

## 2.1 Modes of coherence

The most important fact that we will use in this chapter is that time-translation covariant operations, i.e., maps satisfying Eq. (1.5), naturally decompose quantum states into modes. Mode decomposition can be seen as a generalisation of Fourier analysis to the context of operators [48]. Physically, mode operators identify components within a quantum state that transform independently as a consequence of the underlying symmetry of the dynamics. They have been introduced to the field of quantum information in Ref. [48], and here we will adapt them to our thermodynamic framework. We consider a  $d$ -dimensional system described by a Hamiltonian  $H$ ,

$$H = \sum_{i=1}^d \epsilon_i |\epsilon_i\rangle\langle\epsilon_i| := \sum_{i=1}^d \hbar\omega_i |\epsilon_i\rangle\langle\epsilon_i|, \quad (2.1)$$

and expand its state  $\rho$  in the energy eigenbasis,

$$\rho = \sum_{i,j} \rho_{ij} |\epsilon_i\rangle\langle\epsilon_j|.$$

We only focus on non-degenerate Hamiltonians, as thermal operations allow us to perform any unitary on a fixed energy subspace. This means that coherence between degenerate energy eigenstates is not a resource, and hence there are no limitations on transferring it between levels within fixed energy eigenspace.

Let us now denote the set of all differences between eigenfrequencies of  $H$  by  $\{\omega\}$ . We then have the following definition:

**Definition 2.1: Modes of coherence**

Consider the following decomposition of a state  $\rho$ :

$$\rho = \sum_{\omega} \rho^{(\omega)}, \quad \rho^{(\omega)} = \sum_{\substack{i,j \\ \omega_i - \omega_j = \omega}} \rho_{ij} |\epsilon_i\rangle\langle\epsilon_j| := \sum_{i,j}^{(\omega)} \rho_{ij} |\epsilon_i\rangle\langle\epsilon_j|, \quad (2.2)$$

where we have introduced the symbol  $\sum^{(\omega)}$  to indicate the sum over indices  $i, j$  such that  $\omega_i - \omega_j = \omega$ . The operators  $\rho^{(\omega)}$  are known as the *modes of coherence* of a state  $\rho$ .

**Remark 2.1: Modes as irreps**

A given mode  $\omega$  is characterised by its transformation property under the symmetry group:

$$e^{-iHt} \rho^{(\omega)} e^{iHt} = e^{-i\omega t} \rho^{(\omega)}, \quad (2.3)$$

and is therefore composed of multiple copies of a fixed one-dimensional irrep of the time-translation group.

We now note that for a time-translation covariant CPTP map  $\mathcal{E}$  such that  $\mathcal{E}(\rho) = \sigma$ , we have

$$e^{-iHt} \mathcal{E}(\rho^{(\omega)}) e^{iHt} = \mathcal{E}(e^{-iHt} \rho^{(\omega)} e^{iHt}) = e^{-i\omega t} \mathcal{E}(\rho^{(\omega)}).$$

We thus see that each mode  $\rho^{(\omega)}$  of the initial state  $\rho$  is independently mapped by a covariant operation to the corresponding mode  $\sigma^{(\omega)}$  of the final state:

$$\forall \omega : \quad \mathcal{E}(\rho^{(\omega)}) = \sigma^{(\omega)}. \quad (2.4)$$

The above equation allows us to introduce natural measures of coherence for each mode [48]. Since the 1-norm is contractive under general quantum operations, for any bounded linear operator  $A$  we have the following

$$\|\mathcal{E}(A)\|_1 \leq \|A\|_1, \quad \|A\|_1 := \text{Tr} \left( \sqrt{AA^\dagger} \right). \quad (2.5)$$

Now, Eqs. (2.5) and (2.4) together imply that the total amount of coherence in each mode is nonincreasing under time-translation covariant operations. Thus, for all  $\omega$  we have

$$\sum_{i,j}^{(\omega)} |\sigma_{ij}| \leq \sum_{i,j}^{(\omega)} |\rho_{ij}|. \quad (2.6)$$

Since thermal operations are a subset of time-translation covariant operations, the constraints discussed above also affect all transformations performed with TOs.

## 2.2 Bounds on coherence transformations

In Sec. 1.2.2 of Chapter 1 we have discussed the problem of thermodynamic interconversion between states diagonal in the energy eigenbasis. Here we will investigate this issue for generic quantum states with coherence. Recall that given an initial incoherent state  $\rho = \rho^{(0)}$ , a final state  $\sigma = \sigma^{(0)}$  is thermodynamically accessible if and only if

$$\rho^{(0)} \succ_{\beta} \sigma^{(0)}, \quad (2.7)$$

where  $\succ_{\beta}$  denotes thermo-majorisation (see Definition M8). Now, the constraint specified by Eq. (2.4) shows that given two general quantum states,  $\rho$  and  $\sigma$ , for  $\sigma$  to be thermally accessible from  $\rho$  a set of independent equations must be simultaneously fulfilled, one for each mode. The thermo-majorisation condition of Eq. (2.7) only ensures that Eq. (2.4) is satisfied for the  $\omega = 0$  mode. Thus, the question of thermodynamic constraints on transformations of coherence for all non-zero modes is left open. Here, we will address it by developing upper and lower bounds on how the modes of coherence evolve under general thermodynamic transformations. Then, in Sec. 2.3, we will discuss how we can apply our results to study coherence manipulations allowed within thermodynamics.

### 2.2.1 Upper bound for covariant operations

As the modes transform independently, for any time-translation covariant operation  $\mathcal{E}$  we can introduce a stochastic GP matrix  $\Lambda$ , whose matrix elements are given by

$$\Lambda_{i'i} = p_{i'|i} = \langle \epsilon'_i | \mathcal{E}(|\epsilon_i\rangle\langle\epsilon_i|) | \epsilon'_i \rangle, \quad (2.8)$$

and the evolution of the  $\omega = 0$  mode will be completely described by  $\Lambda$ . More precisely, by identifying the zeroth mode  $\rho^{(0)}$  with a classical state  $\mathbf{p}$ , we can identify  $\mathcal{E}(\rho^{(0)})$  with  $\Lambda\mathbf{p}$ . Now, instead of directly investigating when a given state  $\rho$  can be transformed into  $\sigma$  via a thermal operation, we will address this problem indirectly by focusing on the following related question: given a covariant map  $\mathcal{E}$  that transforms the diagonal elements of  $\rho$  according to the stochastic (but not necessarily GP) matrix  $\Lambda$ , how much coherence can

be preserved [15, 61]? In physical terms, we can think of this problem as follows: knowing the classical action of  $\mathcal{E}$ , i.e., the energy flows it induces, we ask what is the optimal rate of preserving the coherence content of the system.

We provide the answer in form of the following theorem:

**Theorem 2.1: Coherence preservation under covariant operations**

Let  $\mathcal{E}$  be a time-translation covariant map such that  $\sigma = \mathcal{E}(\rho)$ , and whose classical action (i.e., the action transforming diagonal elements of a density matrix) is given by the stochastic matrix  $\Lambda$  with  $\Lambda_{i'i} = p_{i'|i}$ . Then  $|\sigma_{i'j'}|$  is bounded by:

$$|\sigma_{i'j'}| \leq \sum_{i,j}^{(\omega_{i'j'})} \sqrt{p_{i'|i} p_{j'|j}} |\rho_{ij}|, \quad (2.9)$$

where  $\omega_{i'j'} = \omega_{i'} - \omega_{j'}$ .

*Proof.* First, let us parametrise the action of  $\mathcal{E}$  on the matrix element  $|\epsilon_i\rangle\langle\epsilon_j|$  as follows:

$$\mathcal{E}(|\epsilon_i\rangle\langle\epsilon_j|) = \sum_{i',j'}^{(\omega_{ij})} c_{j'|j}^{i'|i} |\epsilon_{i'}\rangle\langle\epsilon_{j'}|, \quad (2.10)$$

where  $\omega_{ij} = \omega_i - \omega_j$ . For  $i \neq j$  the coefficients  $c_{j'|j}^{i'|i}$  describe how much the initial coherence  $|\epsilon_i\rangle\langle\epsilon_j|$  contributes to the final coherence  $|\epsilon_{i'}\rangle\langle\epsilon_{j'}|$ . For  $i = j$  we have  $c_{i'|i}^{i'|i} = p_{i'|i}$ , i.e., the coefficients describe the classical action of  $\mathcal{E}$  given by a stochastic matrix  $\Lambda$ . Now, we need to recall the Jamiołkowski-Choi isomorphism [62], which maps a quantum channel  $\mathcal{E}$  into the state  $J[\mathcal{E}] := [\mathcal{E} \otimes \mathcal{I}] (|\phi^+\rangle\langle\phi^+|)$ , where  $|\phi^+\rangle = \sum_{i=1}^d |\epsilon_i\epsilon_i\rangle / \sqrt{d}$  is the maximally entangled state, and  $\mathcal{I}$  denotes the identity superoperator. In Ref. [63] it was shown that for a time-translation covariant channel  $\mathcal{E}$  the Choi state  $J[\mathcal{E}]$  satisfies

$$e^{-i\tilde{H}t} J[\mathcal{E}] e^{i\tilde{H}t} = J[\mathcal{E}], \quad (2.11)$$

where  $\tilde{H} = H \otimes \mathbb{I} - \mathbb{I} \otimes H^*$ . In other words, the Choi state is time-translation invariant under the evolution generated by  $\tilde{H}$ .

In order to ensure that  $\mathcal{E}$  is a time translation covariant CPTP map, we can equivalently demand that  $J[\mathcal{E}]$  is positive [64, 65] and satisfies Eq. (2.11). The latter condition means that the Choi state is block diagonal in the eigenbasis of  $\tilde{H}$ . Hence, positivity of  $J[\mathcal{E}]$  is equivalent to positivity of each block. Let us identify these blocks by writing down

$J[\mathcal{E}]$  explicitly:

$$J[\mathcal{E}] = \frac{1}{d} \sum_{i,j} \sum_{i',j'}^{(\omega_{ij})} c_{j'|j}^{i'|i} |\epsilon_{i'}\rangle\langle\epsilon_{j'}| \otimes |\epsilon_i\rangle\langle\epsilon_j| = \frac{1}{d} \sum_{i',i} \sum_{j',j}^{(\omega_{i'i})} c_{j'|j}^{i'|i} |\epsilon_{i'}\epsilon_i\rangle\langle\epsilon_{j'}\epsilon_j|, \quad (2.12)$$

where we have rearranged the expression to emphasise the block-diagonal structure. Each block consists of matrix elements  $c_{j'|j}^{i'|i}$  for which  $\omega_{i'} - \omega_i = \omega_{j'} - \omega_j = \omega$  and can thus be labelled by  $\omega$ . A necessary condition for the positivity of block  $\omega$  is that for all  $i, j$  and  $i', j'$  within, one has

$$|c_{j'|j}^{i'|i}| \leq \sqrt{p_{i'|i} p_{j'|j}}. \quad (2.13)$$

Now note that from Eq. (2.10) and the triangle inequality, we have

$$|\sigma_{i'j'}| \leq \sum_{i,j}^{(\omega_{i'j'})} |c_{j'|j}^{i'|i}| |\rho_{ij}|. \quad (2.14)$$

Using the above together with Eq. (2.13) immediately yields Eq. (2.9).  $\square$

In Appendix A we also provide an alternative proof based on the Cauchy-Schwarz inequality, which is more elementary (it does not require using the Jamiołkowski-Choi isomorphism for a covariant channel), but does not shed much light on the origin and attainability of the bound. On the other hand, we will now show that the construction used in the proof above allows us to study when the bound can be attained. First of all, let us choose each block  $\omega$  of the Choi state  $J[\mathcal{E}]$  to be proportional to an unnormalised pure state  $|\psi_\omega\rangle\langle\psi_\omega|$ , given by

$$|\psi_\omega\rangle = \sum_{i',i}^{(\omega)} \varphi_{i'i} \sqrt{p_{i'|i}} |\epsilon_{i'}\epsilon_i\rangle, \quad (2.15)$$

where  $\varphi_{i'i}$  are phase factors, i.e.,  $|\varphi_{i'i}| = 1$ . This way we ensure positivity of the Choi state for any given classical action  $\Lambda$ . The corresponding quantum channel  $\mathcal{E}$  is described, according to Eq. (2.10), by the following parameters:

$$c_{j'|j}^{i'|i} = \varphi_{i'i} \varphi_{j'j}^* \sqrt{p_{i'|i} p_{j'|j}}. \quad (2.16)$$

We now note the bound from Eq. (2.9) can always be attained for a particular coherence term  $\sigma_{i'j'}$ . This can be achieved by choosing for all  $i$  and  $j$  such that  $\omega_{ij} = \omega_{i'j'}$  the following phase factors:  $\varphi_{j'j} = 1$  and  $\varphi_{i'i} = \rho_{ij}^* / |\rho_{ij}|$ . Moreover, if for the initial state  $\rho$  there exist a set of phase factors  $\{\phi_i\}$  such that

$$\forall i, j : \quad \frac{\rho_{ij}}{|\rho_{ij}|} = \phi_i \phi_j^*, \quad (2.17)$$

then the bound can be simultaneously attained for all coherence terms by choosing  $\varphi_{i'i} = \phi_i^*$  for all  $i'$ . Such a set of phase factors exists, e.g., when  $\rho$  is a pure state or when all off-diagonal elements of  $\rho$  are positive.

### 2.2.2 Upper bound for thermal operations

The bound of Theorem 2.1 applies to all time-translation covariant operations. However, thermal operations are not only covariant, but also Gibbs-preserving. This allows us to further refine this bound by noting that the GP condition puts restrictions on the transition probabilities  $p_{i'|i}$ . First of all, for a GP stochastic matrix  $\Lambda$  with elements  $\Lambda_{i'j} = p_{i'|j}$  we have

$$\gamma_{i'} = \sum_j p_{i'|j} \gamma_j. \quad (2.18)$$

After simple transformations one obtains that for every  $i'$

$$p_{i'|i'} = 1 - \sum_{j \neq i'} p_{i'|j} \frac{\gamma_j}{\gamma_{i'}}.$$

Taking into account that  $p_{i'|i'}$  is positive (as it represents transition probability) yields for every  $i \neq i'$

$$p_{i'|i} \leq \frac{\gamma_{i'}}{\gamma_i} - \sum_{j \neq i, i'} p_{i'|j} \frac{\gamma_j}{\gamma_i} \leq \frac{\gamma_{i'}}{\gamma_i} = e^{\beta \hbar (\omega_i - \omega_{i'})}. \quad (2.19)$$

Hence, if the energy of the final state  $\hbar \omega_{i'}$  is higher than the energy of the initial state  $\hbar \omega_i$ , the transition probability is bounded by  $e^{-\beta \hbar (\omega_{i'} - \omega_i)}$ .

Let us now split the bound given by Eq. (2.9):

$$|\sigma_{i'j'}| \leq \sum_{\substack{i,j \\ \omega_i \leq \omega_{i'}}}^{(\omega_{i'j'})} |\rho_{ij}| \sqrt{p_{i'|i} p_{j'|j}} + \sum_{\substack{i,j \\ \omega_i > \omega_{i'}}}^{(\omega_{i'j'})} |\rho_{ij}| \sqrt{p_{i'|i} p_{j'|j}}. \quad (2.20)$$

We can use the inequality given by Eq. (2.19) in the first sum, and use the time-translation covariance condition  $\omega_i - \omega_j = \omega_{i'} - \omega_{j'}$  that implies  $\omega_{j'} - \omega_j = \omega_{i'} - \omega_i \geq 0$ . Simple manipulations lead then to the final result given by:

$$|\sigma_{i'j'}| \leq \sum_{\substack{i,j \\ \omega_i \leq \omega_{i'}}}^{(\omega_{i'j'})} |\rho_{ij}| e^{-\beta \hbar (\omega_{i'} - \omega_i)} + \sum_{\substack{i,j \\ \omega_i > \omega_{i'}}}^{(\omega_{i'j'})} |\rho_{ij}|. \quad (2.21)$$

This bound on coherence transformations by thermal operations can be easily interpreted physically. Time-translation covariance implies that the contributions to  $\sigma_{i'j'}$  can only

come from elements within the same mode. The Gibbs-preserving condition (necessary for the non-existence of perpetuum mobiles) imposes an asymmetry in the contributions to the final coherence. The initial low-energy coherences, when contributing to the final high-energy coherences, are exponentially damped by the factor  $e^{-\beta\hbar(\omega_{j'} - \omega_i)}$ . On the other hand, our bound does not constrain the possibility of transforming high-energy coherences into coherences between lower energy levels.

### 2.2.3 Lower bound on guaranteed coherence preservation

We now turn to a lower bound on the guaranteed fraction of coherence that can be preserved during a transformation induced by a thermal operation. More precisely, we will study the following scenario. Given an initial state  $\rho$  we want to transform its classical component, the zeroth mode  $\rho^{(0)}$ , into  $\sigma^{(0)}$  using a thermal operation. From Eq. (2.7) we know that there exists a TO performing this transformation if and only if  $\rho^{(0)} \succ_{\beta} \sigma^{(0)}$ . However, we now also want to know how much quantum coherence can be preserved in such a process. We will thus establish a lower bound on the guaranteed coherence that relies only on the known results about thermodynamic transformations among incoherent states (see Sec. 1.2.2 of Chapter 1).

In order to achieve this, we will first prove the following lemma:

#### Lemma 2.2: Thermal operations are convex

Thermal operations form a convex set, i.e., for any two given TOs,  $\mathcal{E}_1$  and  $\mathcal{E}_2$ , a map  $\lambda\mathcal{E}_1 + (1 - \lambda)\mathcal{E}_2$  is also a TO for all  $\lambda \in [0, 1]$ .

*Proof.* Let  $\mathcal{E}_1$  and  $\mathcal{E}_2$  be two thermal maps acting on a system  $S$  defined as in Eq. (1.3):  $\mathcal{E}_1$  by  $(U_{SB_1}, \gamma_{B_1})$  and  $\mathcal{E}_2$  by  $(U_{SB_2}, \gamma_{B_2})$ , where

$$\gamma_{B_1} = \frac{e^{-\beta H_{B_1}}}{Z_{B_1}}, \quad \gamma_{B_2} = \frac{e^{-\beta H_{B_2}}}{Z_{B_2}},$$

are the thermal Gibbs states of two baths,  $B_1$  and  $B_2$ , and  $U_{SB_i}$  is an energy preserving unitary on  $S + B_i$ :

$$[U_{SB_i}, H + H_{B_i}] = 0, \quad i = 1, 2.$$

Let us introduce a  $D$ -dimensional ancillary system  $A$ , described by a Hamiltonian  $H_A = \mathbb{I}_A$  and prepared in a thermal Gibbs state  $\gamma_A = \mathbb{I}_A/D$ . Now, the total Hamiltonian of the

system, the two baths and the ancilla is  $H_T = H + H_{B_1} + H_{B_2} + H_A$ . We define the following unitary acting on this total joint system:

$$U := \Pi_A^{(1)} \otimes U_{SB_1} + \Pi_A^{(2)} \otimes U_{SB_2}, \quad (2.22)$$

where  $\Pi_A^{(1)}$  and  $\Pi_A^{(2)}$  are respectively rank  $k$ , and rank  $D - k$  projectors onto the subspaces of the degenerate ancillary system  $A$ , and  $\Pi_A^{(1)} + \Pi_A^{(2)} = \mathbb{I}_A$ . We can check that, for  $i = 1, 2$ ,

$$[\Pi_A^{(i)} \otimes U_{SB_i}, H_T] = [\Pi_A^{(i)} \otimes U_{SB_i}, H + H_{B_i}] = \Pi_A^{(i)} \otimes [U_{SB_i}, H + H_{B_i}] = 0,$$

so that  $U$  is energy-preserving on  $S + B_1 + B_2 + A$ . We finally have

$$\begin{aligned} \text{Tr}_{AB_1B_2} \left( U(\rho \otimes \gamma_A \otimes \gamma_{B_1} \otimes \gamma_{B_2}) U^\dagger \right) &= \frac{1}{D} \sum_{i=1}^2 \text{Tr}_{AB_i} \left( \Pi_A^{(i)} \otimes U_{SB_i} (\rho \otimes \mathbb{I} \otimes \gamma_{B_i}) \Pi_A^{(i)} \otimes U_{SB_i}^\dagger \right) \\ &= \frac{k}{D} \mathcal{E}_1(\rho) + \left( 1 - \frac{k}{D} \right) \mathcal{E}_2(\rho). \end{aligned}$$

Thus  $(U, \gamma_A \otimes \gamma_{B_1} \otimes \gamma_{B_2})$  defines a thermal operation equivalent to any rational convex combination of  $(U_{SB_1}, \gamma_{B_1})$  and  $(U_{SB_2}, \gamma_{B_2})$ . Irrational convex combinations are approached with arbitrary accuracy.  $\square$

We are now ready to present an algorithm to find the guaranteed amount of coherence that can always be preserved. We illustrate the following reasoning in Fig. 2.1. Assume

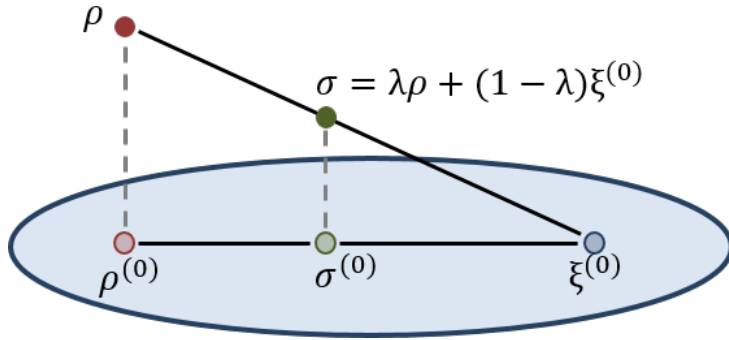


Figure 2.1: *Guaranteed coherence preservation.* The shaded region represents the set of incoherent states. By convexity of the set of thermal operations, if a state  $\zeta^{(0)}$  is achievable from  $\rho$  through a thermal operation, then also all states of the form  $\lambda\rho + (1 - \lambda)\zeta^{(0)}$  for  $\lambda \in [0, 1]$  are achievable. Thus, finding a state  $\zeta^{(0)}$  along the line of  $\rho^{(0)}$  and  $\sigma^{(0)}$  (which only requires verifying the thermo-majorisation condition as explained in the main text), means that  $\lambda$  fraction of coherence can be preserved, while transforming the diagonal of a state from  $\rho^{(0)}$  to  $\sigma^{(0)}$ .

that there exists a thermal operation mapping  $\rho^{(0)}$  into  $\sigma^{(0)}$ . Define  $\Sigma$  as the set of quantum states with the zeroth mode equal to  $\sigma^{(0)}$  and denote by  $\mathcal{T}_+(\rho)$  the set of states accessible from  $\rho$  through thermal maps. It is easy to see that  $\Sigma \cap \mathcal{T}_+(\rho) \neq \emptyset$ , because the dephasing operation  $\mathcal{D}(\rho) = \rho^{(0)}$  is a thermal operation. We now want to show how, among these states, one can find a state that preserves a guaranteed fraction of coherence of the initial state  $\rho$ , and what that fraction is. First, consider the set  $\mathcal{T}_+(\rho^{(0)})$ , which is contained in  $\mathcal{T}_+(\rho)$  and is completely characterised by thermo-majorisation. Within this set, consider the set of incoherent states  $\{\xi = \xi^{(0)} : \sigma^{(0)} = \lambda\rho^{(0)} + (1 - \lambda)\xi^{(0)}\}$  along the line of  $\rho^{(0)}$  and  $\sigma^{(0)}$ . For any member of the set we can define a state  $\sigma = \lambda\rho + (1 - \lambda)\xi^{(0)}$ . One can check that  $\sigma \in \Sigma$  and that  $\sigma$  is a convex combination of two states in  $\mathcal{T}_+(\rho)$  [ $\rho \in \mathcal{T}_+(\rho)$  trivially and by definition  $\xi^{(0)} \in \mathcal{T}_+(\rho^{(0)}) \subseteq \mathcal{T}_+(\rho)$ ]. Moreover, since the set of thermal maps is a convex set we have that  $\mathcal{T}_+(\rho)$  is also convex. This immediately implies  $\sigma \in \mathcal{T}_+(\rho)$ . Now note that the modes  $\sigma^{(\omega)}$  of the final state  $\sigma$  (as defined in Eq. (2.2)) can only come from the initial state  $\rho$ , as  $\xi^{(0)}$  has zero coherence. Therefore, we conclude that the fraction  $\lambda$  gives a lower bound on the coherence that can be preserved in each mode, as  $\sigma^{(\omega)} = \lambda\rho^{(\omega)}$ . By extremising  $\xi^{(0)}$  within the set  $\mathcal{T}_+(\rho^{(0)})$  one can obtain the optimal fraction  $\lambda = \lambda_*$  of guaranteed coherence in each mode:

$$\sigma^{(\omega)} = \lambda_*\rho^{(\omega)}. \quad (2.23)$$

Thus, the extremal value of  $\lambda$  for which the state  $\zeta^{(0)} = (\sigma^{(0)} - \lambda\rho^{(0)})/(1 - \lambda)$  is thermo-majorised by  $\rho^{(0)}$  also yields the guaranteed fraction of coherence that can be preserved while transforming the diagonal of a state from  $\rho^{(0)}$  to  $\sigma^{(0)}$  via thermal operations.

## 2.3 Analysis and applications of the bounds

### 2.3.1 Coherence shifting and merging

One of the simplifying assumption that can be made while studying coherence transformations under thermal maps is that all energy differences in the Hamiltonian of the system are distinct [15, 16]. This results in all non-zero modes being composed of just a single off-diagonal element, and their evolution being limited to a simple decay. Such approach, however, does not capture all the physics of ubiquitous systems, such as harmonic oscillators or spin- $j$  particles in a magnetic field, where modes are composed of more than one

off-diagonal element. Here, we want to show how our results that bound only the overall coherence in a mode, and not each single element separately, allow us to study non-trivial dynamics within each composite mode.

In order to study this problem we will consider the simplest system with non-trivial mode structure – a qutrit in a state  $\rho$  described by the following Hamiltonian:

$$H = \sum_{j=0}^2 j \hbar \omega_0 |\epsilon_j\rangle\langle\epsilon_j|.$$

Using Eq. (2.2) we easily identify that the mode  $\omega_0$  is composed of two off-diagonal elements:

$$\rho^{(\omega_0)} = \rho_{10} |\epsilon_1\rangle\langle\epsilon_0| + \rho_{21} |\epsilon_2\rangle\langle\epsilon_1|,$$

while, e.g.,  $\rho^{(2\omega_0)} = \rho_{20} |\epsilon_2\rangle\langle\epsilon_0|$  consists of a single term. We will consider two primitive operations on mode  $\omega_0$  that may be used as building blocks in general coherence processing for higher-dimensional systems. One of them is *coherence shifting*: shifting up or down in energy the coherence between two given energy levels, preserving as much of it as we can (e.g.  $\rho_{10}$  can be shifted “up” to  $\rho_{21}$ , which can be then shifted “down” to  $\rho_{10}$ ). The second primitive that we will consider is *coherence merging*: given two coherence terms (e.g.  $\rho_{10}$  and  $\rho_{21}$ ) we want to optimally merge them into a single one (e.g.,  $\rho_{10}$ ). We will first study the limitations imposed by time-translation covariance and then show how the situation changes in thermodynamics due to the Gibbs-preserving condition.

### Coherence shifting

Assume that the only non-vanishing coherence term is  $|\rho_{10}| = c$  and that we want to transfer it inside mode  $\omega_0$  up, so that for the final state  $\sigma$  we have a non-zero  $|\sigma_{21}|$ , i.e., we want to transform the coherence between energy levels  $|\epsilon_0\rangle$  and  $|\epsilon_1\rangle$  into coherence between  $|\epsilon_1\rangle$  and  $|\epsilon_2\rangle$ . From Theorem 2.1 we know that if such transformation is performed via a time-translation covariant map, we have the following bound:

$$|\sigma_{21}| \leq c \sqrt{p_{1|0} p_{2|1}} \leq c. \quad (2.24)$$

If Eq. (2.24) is tight, a perfect shift can be obtained. It is easy to check that this is actually the case: a covariant map described by Kraus operators

$$M_1 = |\epsilon_1\rangle\langle\epsilon_0| + |\epsilon_2\rangle\langle\epsilon_1|, \quad M_2 = |\epsilon_2\rangle\langle\epsilon_2|, \quad (2.25)$$

perfectly shifts the coherence from  $|\epsilon_1\rangle\langle\epsilon_0|$  to  $|\epsilon_2\rangle\langle\epsilon_1|$ . The situation would be analogous if we started with a coherence term  $|\rho_{21}|$  and wanted to move it down in energy to  $|\rho_{10}|$ . Therefore, coherence transfer within a mode through a time-translation covariant operation is completely reversible.

This reversibility breaks down for thermal operations, where the second law of thermodynamics, encoded by GP condition given in Eq. (1.4), has to hold. We now need to distinguish two situations: either we start with a coherence term  $|\rho_{10}| = c$  and we move it up in energy to  $|\sigma_{21}|$ , or we perform the reverse task. From Eq. (2.21) we immediately obtain bounds for the final magnitude of the transferred coherence:

$$|\sigma_{10}| \leq c \text{ for shifting down,} \quad |\sigma_{21}| \leq ce^{-\beta\hbar\omega_0} \text{ for shifting up.} \quad (2.26)$$

Also in this case these bounds are tight, i.e., there are thermal operations achieving the above limits. To see this, consider a harmonic oscillator bath with thermal state given by:

$$\gamma_B = \frac{1}{Z_B} \sum_{j=0}^{\infty} e^{-\beta j\hbar\omega_0} |E_j\rangle\langle E_j|, \quad Z_B = (1 - e^{-\beta\hbar\omega_0})^{-1}.$$

Now, consider the following joint unitary acting on the qutrit system and the bath:

$$U = \sum_{i=2}^{\infty} (|\epsilon_1 E_{i-1}\rangle\langle\epsilon_2 E_{i-2}| + |\epsilon_0 E_i\rangle\langle\epsilon_1 E_{i-1}| + |\epsilon_2 E_{i-2}\rangle\langle\epsilon_0 E_i|) \\ + |\epsilon_0 E_0\rangle\langle\epsilon_0 E_0| + |\epsilon_0 E_1\rangle\langle\epsilon_1 E_0| + |\epsilon_1 E_0\rangle\langle\epsilon_0 E_1|.$$

It is straightforward to verify the above unitary is energy conserving, as it only mixes states with the same total energy. By direct calculation we can also verify that

$$\text{Tr}_B (U(|\epsilon_2\rangle\langle\epsilon_1| \otimes \gamma_B)U^\dagger) = |\epsilon_1\rangle\langle\epsilon_0|, \quad \text{Tr}_B (U^\dagger(|\epsilon_1\rangle\langle\epsilon_0| \otimes \gamma_B)U) = e^{-\beta\hbar\omega_0} |\epsilon_2\rangle\langle\epsilon_1|.$$

Hence both bounds from Eq. (2.26), for shifting down and up in energy, are achievable via the presented thermal operations. This also proves that the irreversibility (directionality) within each mode suggested by Eq. (2.21) is not just an artefact due to the bound being not tight. It is actually possible to perfectly transfer coherence down in energy, whereas the opposite task is exponentially damped due to the second law. Fig. 2.2 presents a “shift cycle”, in which coherence between high energy levels is transferred down to lower energies and then up again. Due to the GP condition this thermodynamic process is irreversible.

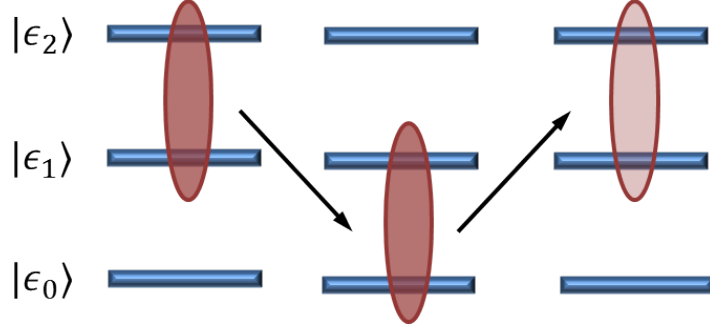


Figure 2.2: *Irreversibility of coherence shift cycle.* Coherence between high energy levels is transferred down to low energy levels and then up again. The magnitude of the coherence terms is proportional to the intensity of the blobs. The first operation can be achieved perfectly, whereas the second results in damping of coherence. This directionality imposed by the second law implies that coherence transfers, similarly to heat transfers, are generally irreversible.

### Coherence merging

Let us now analyse a second primitive operation, coherence merging. Assume we are given a state  $\rho$  with two non-vanishing coherence terms in mode  $\omega_0$ :  $|\rho_{10}| = a$  and  $|\rho_{21}| = b$ , and we want to merge them into a single coherence term  $\sigma_{10}$  (the results for merging into  $\sigma_{21}$  are analogous for time-translation covariant maps). The bound from Theorem 2.1 yields:

$$|\sigma_{10}| \leq \sqrt{p_{1|1}p_{0|0}}a + \sqrt{p_{1|2}p_{0|1}}b \leq \sqrt{p_{1|1}}a + \sqrt{p_{0|1}}b \leq \sqrt{p_{1|1}}a + \sqrt{1 - p_{1|1}}b.$$

One can easily prove that the above bound is maximised for  $p_{1|1} = a^2/(a^2 + b^2)$ , so ultimately

$$|\sigma_{10}| \leq \sqrt{a^2 + b^2}. \quad (2.27)$$

A time-translation covariant map achieving the above bound can be constructed in the following way. Consider a CPTP map specified by its Kraus operator decomposition:

$$M_j = \frac{1}{\sqrt{3}} \left[ |\epsilon_0\rangle \left( e^{i\frac{2\pi j}{3}} \langle \epsilon_0| + x \langle \epsilon_1| \right) \right] + |\epsilon_1\rangle \left( e^{i\frac{2\pi j}{3}} \sqrt{1-x^2} \langle \epsilon_1| + \langle \epsilon_2| \right),$$

with  $x \in [0, 1]$  and  $j = \{0, 1, 2\}$ . One can verify that this map is time-translation covariant by checking that each mode is mapped into itself. By direct calculation we can also show that

$$|\sigma_{10}| = \langle \epsilon_1| \left( \sum_{j=0}^2 M_j (a |\epsilon_1\rangle\langle \epsilon_0| + b |\epsilon_2\rangle\langle \epsilon_1|) M_j^\dagger \right) |\epsilon_0\rangle = \sqrt{1-x^2}a + xb.$$

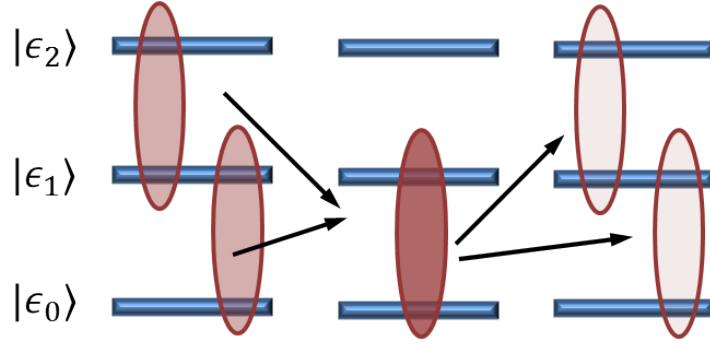


Figure 2.3: *Irreversibility of coherence merging cycle.* Merging of coherence that are sharing an energy level always results in irreversible losses, even if we merge into the lower energy term. The second law, however, imposes additional irreversibility that exponentially damps the contribution to high energy coherence coming from low energy coherence.

The choice  $x = b/\sqrt{a^2 + b^2}$  thus saturates the bound given by Eq. (2.27). Let us also note that coherence merging at the maximum rate  $a + b$  cannot be achieved (see Fig. 2.3), as inevitable losses arise when the two coherence terms have an overlap, i.e., both correspond to the coherence between state  $|\epsilon_1\rangle$  and one of the other two states. This property distinguishes merging from shifting.

Let us now switch to the thermodynamic scenario. The bound for merging two coherences into a single coherence term now depends on whether one merges into high energy coherence or into low energy coherence. By applying a similar reasoning as in the case of covariant operations we obtain bounds for coherence merging under thermal operations:

$$|\sigma_{10}| \leq \sqrt{a^2 + b^2} \text{ for merging down, } |\sigma_{21}| \leq \sqrt{e^{-\beta\hbar\omega_0} a^2 + b^2} \text{ for merging up.} \quad (2.28)$$

Finally, let us note that the qutrit example does not exhaust all the merging scenarios. The reason is that the non-trivial mode in the case analysed is composed of two off-diagonal elements that are overlapping. For higher dimensional systems one can imagine a situation in which elements of the same mode are not overlapping, e.g.,  $|\epsilon_1\rangle\langle\epsilon_0|$  and  $|\epsilon_3\rangle\langle\epsilon_2|$  for a system with equidistant spectrum. In contrast to the overlapping case for time-translation covariant operations one can then perform a perfect merging using the shift operation from the previous section, see Eq. (2.25).

### 2.3.2 Qubit transformations

We will now show how our upper bound for coherence transformation can be used in order to solve the qubit interconversion problem for time-translation covariant and thermal operations. In other words, we will find future thermal cones for a general qubit state  $\rho$  and, on the way, the set of states achievable from  $\rho$  via time-translation covariant operations. Let us first parametrise the initial state of the qubit system  $\rho$  and its final state  $\sigma$ , written in the energy eigenbasis, in the following way

$$\rho = \begin{pmatrix} p & c \\ c & 1-p \end{pmatrix}, \quad \sigma = \begin{pmatrix} q & d \\ d & 1-q \end{pmatrix},$$

where  $c$  and  $d$  are assumed real without loss of generality, as a phase change in coherence terms is both covariant and conserves energy. From Eq. (2.9) for time-translation covariant operations we get

$$d \leq c\sqrt{p_{0|0}p_{1|1}}. \quad (2.29)$$

To obtain a distribution  $\mathbf{q} = (q, 1-q)$  from  $\mathbf{p} = (p, 1-p)$  the transition matrix  $\Lambda$ , defined by transition probabilities  $p_{i|j}$  with  $i, j \in \{0, 1\}$  [see Eq. (2.8)], must fulfil  $\Lambda\mathbf{p} = \mathbf{q}$ . This condition together with the stochasticity of  $\Lambda$  gives

$$p_{0|0} = \frac{(p_{1|1} - 1)(1-p) + q}{p} \leq \frac{q}{p}, \quad p_{1|1} = \frac{(p_{0|0} - 1)p + 1 - q}{1-p} \leq \frac{1-q}{1-p}.$$

Note that for  $q < p$  only the first inequality is non-trivial, whereas for  $q > p$  only the second inequality is non-trivial. Using these conditions in Eq. (2.29) gives:

$$d \leq c\sqrt{\alpha}, \quad (2.30)$$

where  $\alpha = \min\left(\frac{q}{p}, \frac{1-q}{1-p}\right)$ . One can check that the time-translation covariant CPTP map given by the following Kraus operators:

$$M_1 = |\epsilon_0\rangle\langle\epsilon_0| + \sqrt{\alpha}|\epsilon_1\rangle\langle\epsilon_1|, \quad M_2 = \sqrt{1-\alpha}|\epsilon_0\rangle\langle\epsilon_1|, \quad (2.31)$$

saturates this bound for  $q \geq p$ , whereas a CPTP map given by  $\{\sigma_x M_1 \sigma_x, \sigma_x M_2 \sigma_x\}$ , with  $\sigma_x$  being the Pauli  $x$  operator, saturates the bound for  $q < p$ . Of course if we can saturate the bound, we can also obtain all states with coherence smaller than maximal, simply by partially dephasing the optimal final state (partial dephasing is a time-translation covariant operation). This shows that the bound of Eq. (2.30) captures all the constraints

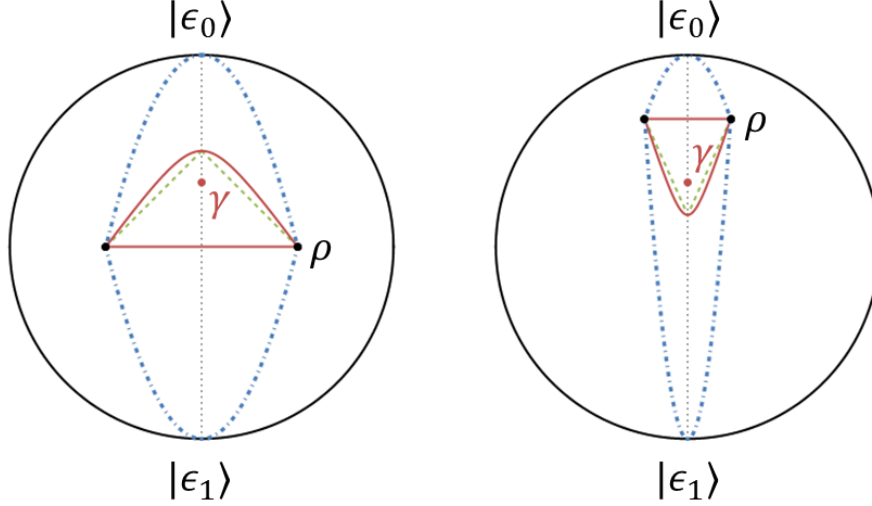


Figure 2.4: *Extremal qubit states under constraint dynamics.* Extremal achievable states from a given initial state  $\rho$  (black points) under time-translation covariant (blue dot-dashed lines) and thermal (solid red lines) operations presented in the Bloch sphere. Red points correspond to thermal occupation of the ground state, here chosen to be  $\gamma_0 = 2/3$ . Dashed green lines correspond to the set of states obtained using the bound for guaranteed coherence preservation. Note that the presented results are symmetric under the rotations around the dotted line connecting Hamiltonian eigenstates.

imposed by time-translation covariance on the evolution of qubit states. In Fig. 2.4 we depict the extremal set of obtainable states via covariant dynamics in the Bloch sphere for exemplary initial states (blue dot-dashed lines).

We will now focus on thermal maps and see how the GP condition  $\Lambda\gamma = \gamma$ , changes the picture in thermodynamics [ $\gamma = (\gamma_0, 1 - \gamma_0)$  stands here, as usual, for the vector of thermal occupation probabilities]. The choice of  $\mathbf{p}$  and  $\mathbf{q}$ , together with the Gibbs-preserving condition, completely fixes  $\Lambda$ :

$$p_{0|0} = \frac{q(1 - \gamma_0) - \gamma_0(1 - p)}{p - \gamma_0}, \quad p_{1|1} = \frac{\gamma_0(1 - q) - p(1 - \gamma_0)}{\gamma_0 - p}. \quad (2.32)$$

Hence, from Eq. (2.29) we obtain

$$d \leq c \frac{\sqrt{(q(1 - \gamma_0) - \gamma_0(1 - p))(p(1 - \gamma_0) - \gamma_0(1 - q))}}{|p - \gamma_0|}. \quad (2.33)$$

In Ref. [15] it has been shown how a thermal operation saturating the above bound can be constructed. Therefore, the above equation, together with the thermo-majorisation condition, completely specify the future thermal cone  $\mathcal{T}_+$  for a given state  $\rho$ . In Fig. 2.4

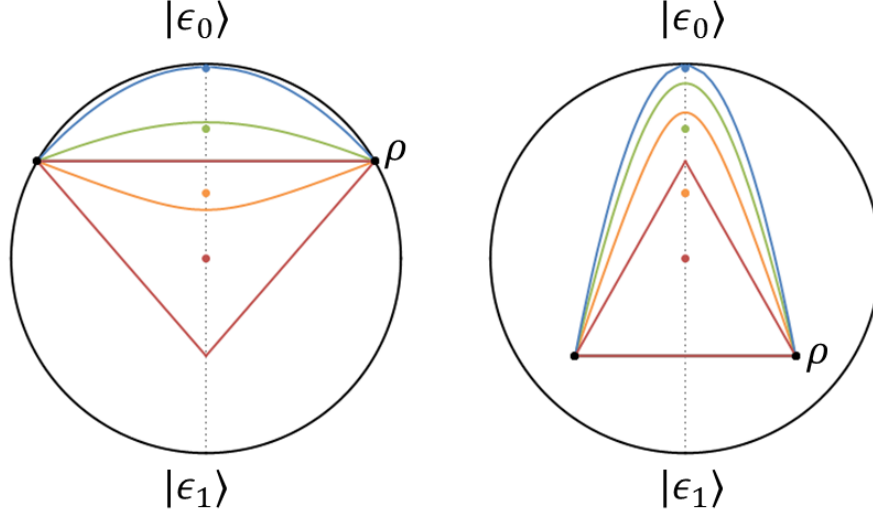


Figure 2.5: *Future thermal cones for qubits (thermal operations)*. Extremal achievable states from a given initial state  $\rho$  (black points) under thermal operations at different temperatures (coloured lines) presented in the Bloch sphere. Points on the  $z$  axis represent thermal states for the same set of temperatures (the red point in the centre corresponds to infinite-temperature bath, whereas the blue one at the boundary corresponds to the low-temperature limit bath). Note that the presented results are symmetric under the rotations around the dotted line connecting Hamiltonian eigenstates.

we depict the extremal set of obtainable states via thermal operations on the Bloch sphere for exemplary initial states [all states lying inside the region bounded by solid red lines belong to  $\mathcal{T}_+(\rho)$ ].

Let us now proceed to the guaranteed coherence bound. Using the thermo-majorisation condition for a qubit we find that the extremal achievable incoherent state, characterised by a probability distribution  $\tilde{\mathbf{q}} = (\tilde{q}, 1 - \tilde{q})$ , is given by

$$\tilde{q} = 1 - \frac{1 - \gamma_0}{\gamma_0} p. \quad (2.34)$$

Then  $\lambda_*$ , specified in Eq. (2.23), is given by

$$\lambda_* = \frac{q - \tilde{q}}{p - \tilde{q}} = \frac{\gamma_0(1 - q) - (1 - \gamma_0)p}{\gamma_0 - p}, \quad (2.35)$$

so that it is always possible to preserve at least  $d = \lambda_* c$  coherence, while thermodynamically transforming the probability distribution over the energies from  $\mathbf{p}$  to  $\mathbf{q}$ . The set of states obtained using the bound for guaranteed coherence preservation is depicted in Fig. 2.4 (green dashed lines).

Finally, in Fig. 2.5 we compare the set of obtainable states for different thermal distributions, i.e., for different temperatures. We make two interesting observations concerning thermal dependence of optimal coherence preservation. Firstly, note that as  $\gamma_0$  approaches 1 (the temperature goes to zero, which is the limit studied in Ref. [16]) the set of states obtainable via thermal operations coincides with “half” of the set of states obtainable via time-translation covariant operations - as long as  $q > p$  one can preserve the same amount of coherence. This suggests that the limitations of low-temperature thermodynamics can be inferred from the limitations on covariant operations, so from the resource theory of asymmetry. Secondly, let us distinguish between heating processes (when  $q < p$ ) and cooling processes (when  $q > p$ ). Then one can check that in the heating scenario the higher the temperature of the bath, the more coherence one can preserve, whereas for cooling processes, the lower temperature ensures better coherence preservation. This shows that for general thermodynamic state transformations to optimally preserve coherence it is necessary to use baths of different temperatures.

### 2.3.3 General remarks

We would like to finish this chapter with some general remarks concerning the resource theory of thermodynamics, coming from the simple example of a qubit system that we have just analysed. The core question of a resource theory is: given a state  $\rho$  that possesses quantum coherence, what is the set of quantum states  $\mathcal{T}_+(\rho)$  accessible from  $\rho$  under thermal operations? In the qubit case it is a rotationally symmetric (about the  $z$ -axis in the Bloch sphere), convex set of states that is presented in Fig. 2.6. The boundary surface of the red region denotes the states that preserve the maximal amount of coherence while having a given final energy distribution. What can we learn from this picture about the thermodynamics of a general  $d$ -dimensional system?

First of all one might expect that, due to the intrinsically dissipative interactions of a quantum system with the heat bath, coherence is only playing a passive role in the process. If this was the case all possible transformations would be attainable just by a combination of dephasings and thermal operations on incoherent states. However, already for qubits, the set of states achievable in this way is limited to the green triangle in Fig. 2.6 and clearly does not coincide with  $\mathcal{T}_+(\rho)$ . We conclude that coherence is actively contributing to enlarge the set of thermodynamically accessible states.

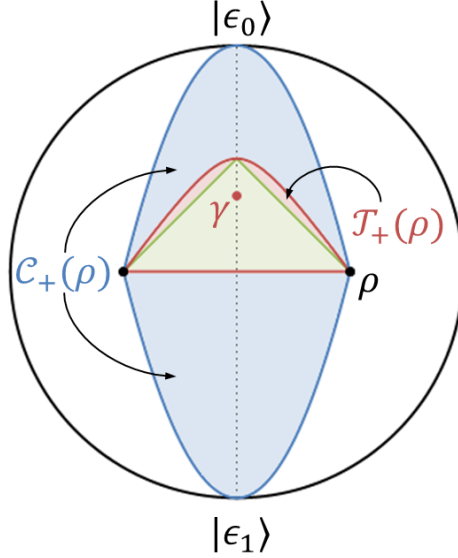


Figure 2.6: *General structure.* Future thermal cone  $\mathcal{T}_+(\rho)$  for a qubit system is given in the Bloch sphere by the region bounded by the solid red line. If thermal operations on coherent states were trivial, i.e., if they were equivalent to dephasings and operations on incoherent states, then this set would reduce to the green triangle. Moreover, even if one has access to arbitrary amount of work (but not coherence), the set of achievable states is extended to the blue region  $\mathcal{C}_+(\rho)$ , but not to the whole Bloch sphere.

One might also ask the following question: if we are given an unbounded source of work, would coherence still be a resource? If the answer was no, all constraints could be lifted by a sufficiently large work source. Work would be *the* universal resource of thermodynamics. However, this is not the case. To see this suppose the unbounded amount of work is given in the form of arbitrary number of copies of pure, zero-coherence states, say  $|\epsilon_0\rangle\langle\epsilon_0|^{\otimes N}$  (as explained in Sec. 1.2.4 of Chapter 1 one can extract  $\ln 2/\beta$  amount of work from a pure ground state of the system, so that when  $N \rightarrow \infty$  we indeed obtain an unbounded source of work). The assumption of zero coherence means that within the resulting theory we are restricted to a subset of time-translation covariant operations (because incoherent states are time-translation invariant they cannot lift the covariant constraint). Conversely, any time-translation covariant operation  $\mathcal{E}$  possesses a Stinespring dilation  $\mathcal{E}(\rho) = \text{Tr}_B(U(\rho \otimes \sigma_B)U^\dagger)$  where  $\sigma_B$  is an invariant state and  $U$  is an energy-preserving unitary on the joint system [53]. Since thermal operations on  $|\epsilon_0\rangle\langle\epsilon_0|^{\otimes N}$  (for arbitrary large  $N$ ) allow us to create any incoherent state, a state  $\sigma_B$  can be formed. Therefore, when an arbitrary amount of work is provided the resource theory

of thermodynamics is trivialised to the theory of time-translation covariant operations. For a qubit system this means that the thermal cone  $\mathcal{T}_+(\rho)$  is extended to the set of states  $\mathcal{C}_+(\rho)$  accessible under time-translation covariant evolution (blue region in Fig. 2.6). As the set  $\mathcal{C}_+(\rho)$  is a strict subset of the full Bloch sphere, we can conclude that work is not a universal resource and coherence resources should be carefully accounted for.

Finally, as we will discuss in detail in the following chapter, the standard Szilard argument linking classical information and thermodynamic work does not simply carry over to the quantum regime. Classically, we know that in the presence of a heat bath at inverse temperature  $\beta$  a single bit has an energetic value of  $\ln 2/\beta$  (see Sec. 1.2.4 of Chapter 1). Thus, one might also expect that a single pure qubit state can be used to perform  $\ln 2/\beta$  of mechanical work. However, consider a qubit system in a “coherent Gibbs state”  $|\gamma\rangle\langle\gamma|$  with  $|\gamma\rangle \propto e^{-\beta\epsilon_0/2} |\epsilon_0\rangle + e^{-\beta\epsilon_1/2} |\epsilon_1\rangle$ . As we will shortly show in Chapter 3, due to total energy conservation implied by the first law of thermodynamics it is fundamentally impossible to distinguish this state from the thermal Gibbs state  $\gamma$ . Therefore, no work can be extracted from such a pure state unless we have access to *coherence resources* [14].

# Chapter 3

## Work extraction from coherence

Since the concept of Maxwell’s demon was devised [66], we have known that information constitutes a thermodynamic resource: in the presence of a heat bath we can exploit the knowledge about the state of a system to perform work [60]. In Sec. 1.2.4 of Chapter 1 we discussed this process for classical states, i.e., states incoherent in the energy eigenbasis, and we have shown that the amount of extractable work is upper bounded by the change of the system’s free energy  $\Delta F$ . Thus, we have seen how the classical information about which energy state of the system is occupied can be traded for work. With quantum information, however, the situation gets more complicated due to the superposition principle resulting in the existence of states that have coherence between different energy eigenspaces. The existence of such states means that possessing full knowledge about the system, i.e., having a pure quantum state, does not necessarily mean that one can predict the outcome of an energy measurement. Since the classical result depends on the information about the distribution over energy eigenstates, it is not clear how the quantum information associated with a state with coherence can be converted into work.

In this chapter, we will analyse the process of work extraction from coherence (between states belonging to different energy eigenspaces) within the framework of the resource theory of thermodynamics, focusing particularly on the extraction from individual qubit systems. After setting the scene and introducing important notions, we will present an existing approach to the problem of work extraction and argue that, within the regime of individually processed systems, it fails to account for all the resources used during the process. The reason for this lies in a typical assumption that a classical external field can be used in the process, while experiencing no back-reaction [67, 68]. Such an assumption,

however, does not allow for a full accounting of the thermodynamic cost of maintaining the field. Although this cost may be small in a single use, it has to be accounted for since the work gain will also be small. On the other hand, a resource-theoretic treatment of the problem allows us for a careful book-keeping of all the resources used during a process. In particular, instead of employing an external field we will use the notion of a *reference* – an ancillary quantum system carrying coherence. It will allow us to address a crucial question of this chapter: how to extract work from coherence in a repeatable way, i.e., without deteriorating ancillary coherence resources. We will show how one can come arbitrarily close to extracting  $\Delta F$  amount of work from a state with coherence, by choosing the “size” of the reference system appropriately large. We will also prove that even with a reference of bounded size one can exploit the coherences of individual quantum systems to enhance the performance of work extraction protocols (both in the average and the single-shot sense), but not to the extent that could be expected in the classical limit.

### 3.1 Coherence and work

Let us start by introducing the framework and the core assumptions that our results rest upon. There are two ways in which coherence can enter the thermodynamics of the systems under consideration. This can happen either explicitly, by transferring it from an external system with quantum coherence (a trivial example being a swap operation between the system and an ancillary coherent state); or implicitly, by allowing operations that do not conserve energy (e.g.,  $|\epsilon_0\rangle \rightarrow (|\epsilon_0\rangle + |\epsilon_1\rangle)/\sqrt{2}$ ) or conserve it only on average (e.g.,  $|\epsilon_1\rangle \rightarrow (|\epsilon_0\rangle + |\epsilon_2\rangle)/\sqrt{2}$ , with  $\epsilon_2 = 2\epsilon_1$ ). As we want keep track of any coherence resources being used, we will only allow for those transformations that do not implicitly introduce coherence, and hence our set of free operations will be given by thermal operations (recall Definition 1.3). However, we will also take a closer look at an alternative approach in Sec. 3.1.3 and explain in more detail why we find it not satisfactory.

In this chapter we want to study how quantum coherence can be used to perform work. However, we do not aim here to settle the long-standing issue of what is an appropriate definition of work in quantum thermodynamics [69, 70]. We will thus simply assume that the following holds for classical (incoherent) states:

**Assumption 3.1: Work output from incoherent states**

Let  $\rho$  be a quantum state of a system described by a Hamiltonian  $H$ , with  $\rho$  being incoherent in the energy eigenbasis and thus described by a probability vector  $\mathbf{p}$ . Then, in the presence of a heat bath at inverse temperature  $\beta$ , an average amount of work  $\langle W \rangle(\rho)$  equal to the change of free energy of a state can be extracted from it:

$$\langle W \rangle(\mathbf{p}) = \Delta F(\mathbf{p}), \quad (3.1)$$

where  $\Delta F(\mathbf{p})$  is defined by Eq. (1.32). Moreover, using a single-shot protocol one can extract a sharp amount of work  $W_{ss}^\varepsilon$  with failure probability  $\varepsilon$ :

$$W_{ss}^\varepsilon(\mathbf{p}) = \Delta F_0^\varepsilon(\mathbf{p}) := F_0^\varepsilon(\mathbf{p}) - F(\gamma), \quad (3.2)$$

where  $F_0^\varepsilon$  is a single-shot free energy defined in Eqs. (1.36)-(1.37).

The results that we will present in this chapter can be applied to any definition of work satisfying the above assumption. As discussed in Sec. 1.2.4 of Chapter 1, it is consistent with traditional thermodynamics and has been obtained using work extraction models that differ in details, but agree on the result [6, 11, 67]. Our aim here is to extend Eq. (3.1) and Eq. (3.2) to quantum states with coherence. In particular, we will study whether the change of free energy generalised to quantum states by

$$\Delta F(\rho) = F(\rho) - F(\gamma), \quad F(\rho) = \text{Tr}(\rho H) - \frac{1}{\beta} S(\rho), \quad (3.3)$$

where  $S(\cdot)$  is the von Neumann entropy, properly quantifies the average amount of extractable work. Let us note that

$$\Delta F(\rho) = \frac{1}{\beta} S(\rho||\gamma) = \frac{1}{\beta} [S(\rho||\mathcal{D}(\rho)) + S(\mathcal{D}(\rho)||\gamma)] = \frac{1}{\beta} [S(\rho||\mathcal{D}(\rho)) + S(\mathbf{p}||\gamma)],$$

where  $S(\cdot||\cdot)$  denotes both relative entropy and quantum relative entropy, the latter defined by  $S(\rho||\sigma) = \text{Tr}(\rho \ln \rho) - \text{Tr}(\rho \ln \sigma)$ , whereas  $\mathcal{D}$  is a dephasing channel that removes all coherences between different energy eigenspaces from a quantum state,

$$\mathcal{D}(\rho) := \sum_i \text{Tr}(\Pi_i \rho) \Pi_i,$$

with  $\Pi_i$  being the projectors on the energy eigenspaces of the system under consideration.

We thus see that the change of free energy,

$$\Delta F(\rho) = \Delta F(\mathbf{p}) + \frac{1}{\beta} S(\rho||\mathcal{D}(\rho)), \quad (3.4)$$

has two components: a classical one (equal to the free energy change of a dephased version of a state), and a quantum one (linked to the amount of coherence present in a state). The fact that  $S(\rho||\mathcal{D}(\rho))$  quantifies the coherence of  $\rho$  and that the relative entropy can be split into classical and quantum parts will be also discussed in Chapter 5.

### 3.1.1 Work-locking

Let us recall the resource-theoretic approach to the work extraction process that we described in Sec. 1.2.4 of Chapter 1. We begin with a system initially in a state with coherence  $\rho$ , and finish with a thermal state  $\gamma$ , while optimally increasing the free energy of an ancillary work system  $W$  (also referred to as battery). The initial and final battery states,  $\rho_W$  and  $\rho'_W$ , should be incoherent, so that using Assumption 3.1 we can achieve the coherence-to-work conversion that we are looking for. Schematically:

$$\rho \otimes \rho_W \rightarrow \gamma \otimes \rho'_W. \quad (3.5)$$

Without the use of an ancillary resource state carrying coherence the above transformation is given by a thermal operation. Note that thermal operations commute with the dephasing channel  $\mathcal{D}$ . Hence, we get that if the transformation described by Eq. (3.5) is possible, the following one also is:

$$\mathcal{D}(\rho) \otimes \rho_W \rightarrow \gamma \otimes \rho'_W. \quad (3.6)$$

This implies  $\Delta F(\mathcal{D}(\rho)) \geq (F(\rho'_W) - F(\rho_W))$ , because  $F$  (being a thermodynamic monotone) is nonincreasing under thermal operations. From Assumption 3.1 we then have  $\langle W \rangle(\rho) \leq \langle W \rangle(\mathcal{D}(\rho))$ , so that the amount of extractable work is upper bounded by a classical component of free energy change. A similar argument gives also  $W_{\text{ss}}^\varepsilon(\rho) \leq W_{\text{ss}}^\varepsilon(\mathcal{D}(\rho))$ . Note that in both cases the bound is achievable because dephasing is a thermal operation. This phenomenon, first observed in Refs. [10, 13], is known under the name *work-locking* [9]. Work-locking highlights that, despite contributing to the free energy of the state, quantum coherence does not contribute to work extraction: it is “locked”. It also shows, in agreement with Ref. [14], that the standard formula  $\langle W \rangle(\rho) = \Delta F(\rho)$  applied to every state (also the ones with coherence), implicitly assumes the access to an external source of coherence.

### 3.1.2 Different thermodynamic regimes

Within the thermodynamics of non-equilibrium quantum systems, a very important distinction to make is between “single-shot” statements, which are valid for every run of the protocol, and “many-runs” statements, valid in the case of a large number  $M$  of runs. In the *asymptotic regime*  $M \rightarrow \infty$  one is focused on studying average quantities (like average extracted work), which is justified by the fact that the fluctuations around the average can be made negligible in the limit of a large number of runs of the protocol (which is often the situation of interest in the study of heat engines). On the other hand, although the expected amount of extracted work can be studied in a single-shot regime [67], it potentially carries little information about the system at hand due to the large fluctuations of non-equilibrium thermodynamics [11]. Instead, the focus in the single-shot regime is typically on probabilistic work extraction protocols that guarantee precise and sharp amount of work with a finite probability of success, or some minimum amount of guaranteed work [10, 11, 71, 72]. On top of this classification we can also differentiate between *individual processing* scenarios, in which a single (possibly nanoscale) system undergoes a thermodynamic process on its own; and *collective* scenarios, in which  $N > 1$  copies of a state are processed together.

This classification of thermodynamic regimes in which work extraction can be analysed is presented in Fig. 3.1. Work-locking described in the previous section is a feature appearing in the regime of individually processed quantum systems. On the other hand, allowing for collective processing of the systems one can “unlock” work from quantum coherence. As briefly discussed in Sec. 1.2.4 of Chapter 1 (see Example 1.5), this can be achieved by processing many copies of a system state  $\rho$  collectively and extracting work from relational degrees of freedom that live in decoherence-free subspaces [6, 9, 13, 52]. The intuitive explanation is that one copy of a state  $\rho$  with coherence can act as a reference for the other one, and we have  $\mathcal{D}(\rho^{\otimes 2}) \neq \mathcal{D}(\rho)^{\otimes 2}$ . In the case of finite number of copies  $\rho^{\otimes N}$  a non-zero amount of work is unlocked from coherences, and in the limit of processing collectively infinitely many independent and identically distributed (i.i.d.) copies, the amount of work per copy that can be extracted deterministically equals  $\Delta F(\rho)$ . In other words, in this limit coherence plays no role, and hence our focus will be on the thermodynamics of individual quantum systems ( $N = 1$ , the upper half of Fig. 3.1).

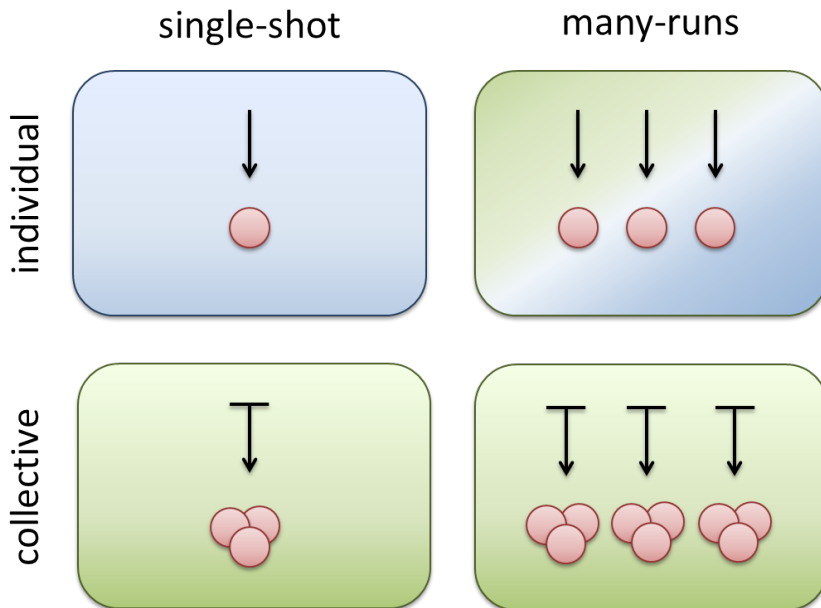


Figure 3.1: *Thermodynamic regimes*. Work extraction protocols can be investigated in different thermodynamic regimes. These can be classified by the number of systems that are processed at each run of the protocol (individual vs collective) and the number of times the protocol is repeated (single-shot vs many-runs). The green background indicates that in a given regime the maximal amount of work that can be extracted is consistent with traditional thermodynamics.

### 3.1.3 Individual processing regime

#### Average energy conservation

Before we describe how to appropriately account for coherence resources involved in a process of work extraction by introducing an ancillary reference system, let us first discuss problematic aspects of an alternative approach based on average energy conservation. Within this approach one introduces coherence implicitly by using operations that conserve the energy only on average [67]. More precisely, one replaces sharp energy conservation, as expressed by the unitary dynamics commuting with the total Hamiltonian  $H_T$  (describing the system, bath and any ancillary systems used), with the condition that such dynamics only keeps the first moment  $\langle H_T \rangle$  constant. Under this weaker condition it can be shown that an amount of energy equal to the free energy difference  $\Delta F(\rho)$  can be extracted on average from a system in an arbitrary quantum state  $\rho$  [67]. However, several problematic issues can then be raised.

Firstly, if one is interested in analysing the class of allowed quantum operations, then in the average-energy scenario this set depends on the particular state one is processing, which is conceptually less appealing and technically problematic from a resource-theoretic perspective. Secondly, restricting energy considerations to the first moment analysis can hide arbitrarily large energy fluctuations described by higher moments, that are not explicitly modelled, but may be highly relevant. To see this consider a system described by a Hamiltonian  $H$  and initially prepared in a state with coherence  $\rho$ . Now, take a unitary transformation  $V$  that conserves energy on average and maps  $\rho$  into an incoherent state  $\rho' = V\rho V^\dagger$ . Since microscopically all processes are ultimately energy-conserving,  $V$  must be realised through a joint energy-preserving unitary  $U$  involving  $\rho$  and some ancillary state  $\rho_A$ ,

$$U(\rho \otimes \rho_A)U^\dagger = \rho' \otimes \rho'_A, \quad (3.7)$$

with  $[U, H + H_A] = 0$ , and  $H_A$  denoting the Hamiltonian of the ancillary system. The uncertainty of an energy measurement on  $\rho'_A$ , quantified by the Shannon entropy  $H(\mathbf{p}^{H_A}(\rho'_A))$  (note that  $H$  as a function of probability distribution denotes the Shannon entropy, whereas  $H$  as an operator denotes the Hamiltonian), can be decomposed as<sup>1</sup> [73]:

$$H(\mathbf{p}^{H_A}(\rho'_A)) = S(\rho'_A) + A(\rho'_A), \quad (3.8)$$

where  $A(\sigma) = S(\sigma || \mathcal{D}(\sigma))$  is the relative entropy between a state and its decohered version. Because  $U$  commutes with the total Hamiltonian we have  $A(\rho \otimes \rho_A) = A(\rho' \otimes \rho'_A)$ , and as the final state of the system  $\rho'$  has no coherence we also have  $A(\rho' \otimes \rho'_A) = A(\rho'_A)$ . Using  $A(\rho \otimes \rho_A) > A(\rho_A)$ , one gets that  $A(\rho'_A) > A(\rho_A)$ . Moreover, from the invariance of the von Neumann entropy under unitary transformations we know that  $S(\rho'_A) = S(\rho_A)$ , so using Eq. (3.8) we can conclude that

$$H(\mathbf{p}^{H_A}(\rho'_A)) > S(\rho_A) + A(\rho_A) = H(\mathbf{p}^{H_A}(\rho_A)), \quad (3.9)$$

which suggests that a more detailed analysis is required to account for all the fluctuations created. Let us illustrate the above considerations with the following example:

---

<sup>1</sup>We will discuss this decomposition in more detail in Chapter 5.

**Example 3.1: Hidden energy fluctuations**

Consider a system with an equally spaced energy spectrum,  $\epsilon_n = n\epsilon$ , and a unitary  $V$  mapping a state  $|\psi_{02}\rangle := (|\epsilon_0\rangle + |\epsilon_2\rangle)/\sqrt{2}$  to  $|\epsilon_1\rangle$ , which preserves energy on average. This unitary has to be realised through a joint energy-preserving unitary  $U$  involving  $|\psi_{02}\rangle$  and some ancillary state  $\rho_A$ ,

$$U(|\psi_{02}\rangle\langle\psi_{02}| \otimes \rho_A)U^\dagger = |\epsilon_1\rangle\langle\epsilon_1| \otimes \rho'_A.$$

Let us assume that the ancillary system was initially prepared in the energy eigenstate  $|E\rangle$ , so that an energy measurement would give a sharp outcome. Then, while the system is transformed from  $|\psi_{02}\rangle$  into  $|\epsilon_1\rangle$ , the ancilla must be transformed into a superposition of energy eigenstates  $(|E + \epsilon\rangle + |E - \epsilon\rangle)/\sqrt{2}$ . Hence an energy measurement would show fluctuations in the final state of the ancilla that were not present initially.

It is important to note that the protocol that extracts work from coherence within the framework of average energy conservation [67], necessarily creates such extra fluctuations, however these are not explicitly modelled within the formalism used. As we will see it is exactly due to these fluctuations that our protocols require work to be invested in restoring the ancillary state. Finally, as the fluctuations created by operations that conserve energy only on average remain outside the formalism, one cannot properly account for the fluctuations in the extracted work outside the asymptotic regime.

**Repeatable use of coherence resources**

In the presence of energy conservation and without additional coherence resources, work-locking prevents us from extracting work from the coherence of individual quantum systems. Our approach to overcome this problem is to stay within the framework of strict energy conservation, but allow for the use of an extra source of coherence. We will refer to this extra system as a *reference* or a coherence reservoir. At one extreme one could allow for the use of an infinite source of coherence (an *unbounded reference* [52, 74])<sup>2</sup>, that entirely negates the constraints and experiences no back-reaction from its use on the

<sup>2</sup>Not to be confused with a reference described by Hamiltonian unbounded from below, which is unphysical.

quantum system. As suggested in Ref. [9], in such case we should be able to extract all the work from coherence. However, one might worry that this involves the accounting “ $\infty - c = \infty$ ”, with  $c$  being some finite resource consumed from an infinitely large coherence reservoir. Indeed, the use of such an unbounded reference allows one to simulate the operations that conserve energy only on average [52], and hide the arising extra fluctuations in the infinitely big reference system. This semiclassical treatment is typical for many standard approaches that assume the existence of a classical field experiencing no back-reaction from the system [67, 68], and works well in many circumstances. However, we are interested in the regimes in which the systems involved in the thermodynamic process (not only the system under study, but also the reference, battery and any other ancillas) may be microscopic quantum systems. Hence, it seems more reasonable to firstly consider the reference containing finite coherence resources – a *bounded reference* – and only then study the limit of an unbounded coherence reservoir [69, 75, 76]. Although a reference system may be modelled by any general quantum system, for the purpose of studying work extraction from coherence of a qubit system we will define it here as follows:

### Definition 3.1: Reference

A reference (or coherence reservoir) is given by an infinite-dimensional ladder system described by a Hamiltonian  $H_R = \sum_{j=0}^{\infty} j |\epsilon_j^R\rangle\langle\epsilon_j^R|$  (we choose just a single energy scale  $\omega = 1$ ). We characterise the quality of the reference through two numbers,  $(\langle\bar{\Delta}\rangle, M)$ . The first parameter,  $\langle\bar{\Delta}\rangle$ , measures the coherence properties of the reference and is given by

$$\langle\bar{\Delta}\rangle = \text{Tr}(\rho_R \bar{\Delta}), \quad \bar{\Delta} = \frac{\Delta + \Delta^\dagger}{2}, \quad (3.10)$$

where  $\Delta$  is the shift operator defined by

$$\Delta = \sum_{j=0}^{\infty} |\epsilon_{j+1}^R\rangle\langle\epsilon_j^R|. \quad (3.11)$$

We have that  $\langle\bar{\Delta}\rangle < 1$  and the limit case  $\langle\bar{\Delta}\rangle = 1$  is called *unbounded* or classical reference. The second parameter,  $M$ , describes the lowest occupied energy state,  $M = \min\{j : \langle\epsilon_j^R|\rho_R|\epsilon_j^R\rangle > 0\}$ .

A sequence of references that come arbitrarily close to a classical one is given by, e.g.,

uniform superpositions of  $L$  energy states when  $L \rightarrow \infty$  and coherent states with arbitrarily large amplitude. The use of  $\langle \bar{\Delta} \rangle$  and  $M$  as relevant quality parameters will soon become clear.

Results from the field of quantum reference frames [52, 74, 77–79] suggest that the back-reaction experienced by the reference will necessarily deteriorate it and consume the resources. However, if the usefulness of the reference or field is continually degraded during the work extraction process, we cannot claim that we are presenting a protocol performing work extraction from the state alone, as extra resources are consumed. Similar issues arise if free energy is continually taken away from the reference. To avoid such problems we take the following approach. We allow for the use of additional coherence resources, but demand that they are used *repeatably* in the following sense: the performance of the reference-assisted work extraction protocol, while operating individually on the  $n$ -th copy of the system, must be the same as while operating on the  $(n + 1)$ -th copy, for all  $n \in \mathbb{N}$ . In other words, repeatability requires that the reference’s ability to perform the protocol never degrades, but crucially its state is allowed to change. Essentially this means that despite the possibility that the free energy of the reference may fluctuate and its coherence properties might change, it can be used indefinitely to repeat the same protocol. To design such a protocol we will employ the surprising result of Ref. [14] that shows how a coherence resource can be used to lift the symmetry constraints imposed by energy conservation.<sup>3</sup>

## 3.2 Protocol

We are now ready to present a general protocol that processes quantum systems individually and allows us to extract work from their coherence. Even though we will focus only on qubit systems, this simplest case is already enough to illustrate the main concept behind repeatable “unlocking” of work from coherence and, in principle, it could be generalised to higher dimensional systems. We will then analyse two variations of our protocol. The first one can come arbitrarily close to extracting all the coherence (all free energy of

---

<sup>3</sup>The author of Ref. [14] actually uses the word “catalysis”, but we prefer to use the word repeatability/repeatable to avoid suggesting that there is no change in the state of the reference. Recall that traditionally a catalyst is a system in a state  $\chi$  that enables  $\rho \otimes \chi \rightarrow \sigma \otimes \chi$ , despite  $\rho \rightarrow \sigma$  being impossible (see, e.g., Refs. [1, 8, 58]). Repeatability, on the other hand, only requires the ancillary system to be as useful at the end as it was at the beginning, while its state may change.

the system) as average work with arbitrarily small failure probability, provided we make the coherence resources of the reference system large enough. However, if one does not have access to arbitrarily large coherence resources, this variation of the protocol does not guarantee perfect repeatability. Therefore, we will examine a second variation that is perfectly repeatable even for bounded references. We will show then that the performance of work extraction in both the single-shot and asymptotic regimes is enhanced only if the quality of the reference is above a certain threshold.

Without loss of generality we can describe our qubit system by a Hamiltonian  $H = |\epsilon_1\rangle\langle\epsilon_1|$ , i.e., set  $\epsilon_0 = 0$  and  $\epsilon_1 = 1$  (note that  $\epsilon_1 - \epsilon_0$  is tuned to the smallest energy gap of the reference, see Definition 3.1). A general pure qubit state with coherence,  $|\psi\rangle\langle\psi|$ , is given by

$$|\psi\rangle = \sqrt{p} |\epsilon_0\rangle + \sqrt{1-p} e^{-i\varphi} |\epsilon_1\rangle, \quad p \in (0, 1). \quad (3.12)$$

We can simplify things by setting  $\varphi = 0$ , because rotations about the  $z$  axis of the Bloch sphere conserve energy, and are hence free operations. Our aim is to unlock work from coherence through the repeatable use of a reference, while processing each copy of  $|\psi\rangle\langle\psi|$  individually. This task requires possessing a reference state  $\rho_R$  and implementing an energy-conserving unitary  $V$ :

$$\rho'_{SR} = V(|\psi\rangle\langle\psi| \otimes \rho_R) V^\dagger, \quad (3.13)$$

satisfying the following:

1. The system is pre-processed to a new state  $\rho' = \text{Tr}_R(\rho'_{SR})$  that allows for better work extraction than from the initial state  $|\psi\rangle\langle\psi|$ .
2. The final reference state  $\rho'_R = \text{Tr}_S(\rho'_{SR})$  can be processed into a state  $\rho''_R$  (perhaps using some of the extracted work) in such a way that the repeatability requirement is satisfied.
3. No collective operations, at any stage of the protocol, are allowed on multiple copies of  $|\psi\rangle\langle\psi|$ .

In Table 3.1, anticipating the conclusions of the following sections, we collect the results obtained in different frameworks and in both regimes (average and single-shot) for the paradigmatic example of a qubit in a “coherent Gibbs state”  $|\gamma\rangle\langle\gamma|$  given by:

$$|\gamma\rangle = \sqrt{\gamma_0} |\epsilon_0\rangle + \sqrt{1-\gamma_0} |\epsilon_1\rangle, \quad (3.14)$$

	Single-shot	Asymptotic
Average energy conservation	$\langle W \rangle = \Delta F$ [67] Large fluctuations [11]	$\langle W \rangle = \Delta F$ [67]
Strict energy conservation	$\varepsilon( \gamma\rangle) = \varepsilon(\gamma)$ , $\langle W \rangle = 0$ [9, 10, 13]	
Strict energy conservation with resource used repeatably	$\varepsilon( \gamma\rangle) = 0$ for unbounded reference $\varepsilon( \gamma\rangle) < \varepsilon(\gamma)$ for bounded reference	$\langle W \rangle = \Delta F$ for unbounded reference $\langle W \rangle < \Delta F$ for bounded reference

Table 3.1: **Individual processing protocols extracting work from  $|\gamma\rangle$ .**  $\langle W \rangle$  denotes the average work that can be extracted from the coherent thermal state  $|\gamma\rangle$  and  $\varepsilon$  denotes the error probability of a single-shot work extraction from a given state. A thermal state of the system is denoted by  $\gamma$ . Note that under operations strictly conserving energy, no work extraction protocol on  $|\gamma\rangle$  can outperform a work extraction protocol on  $\gamma$ , as the two states are indistinguishable.

with  $(\gamma_0, 1 - \gamma_0)$  being the thermal distribution for the system, so that  $\mathcal{D}(|\gamma\rangle\langle\gamma|) = \gamma$ .

### 3.2.1 Explicit steps of the work-extraction protocol

A protocol satisfying the introduced requirements consists of the following steps (see Fig. 3.2):

1. **Pre-processing.** The system  $|\psi\rangle\langle\psi|$  interacts through an energy-preserving unitary  $V(U)$  with the reference  $\rho_R$ . The unitary acting on the joint system  $SR$  is chosen as in Ref. [14] to be:

$$V(U) = |\epsilon_0\rangle\langle\epsilon_0| \otimes |\epsilon_0^R\rangle\langle\epsilon_0^R| + \sum_{l=1}^{\infty} V_l(U), \quad (3.15)$$

with

$$V_l(U) = \sum_{i,j=0}^1 \langle\epsilon_i|U|\epsilon_j\rangle |\epsilon_i\rangle\langle\epsilon_j| \otimes |\epsilon_{l-i}^R\rangle\langle\epsilon_{l-j}^R|. \quad (3.16)$$

We choose

$$U = \begin{pmatrix} \sqrt{1-p} & -\sqrt{p} \\ \sqrt{p} & \sqrt{1-p} \end{pmatrix}, \quad (3.17)$$

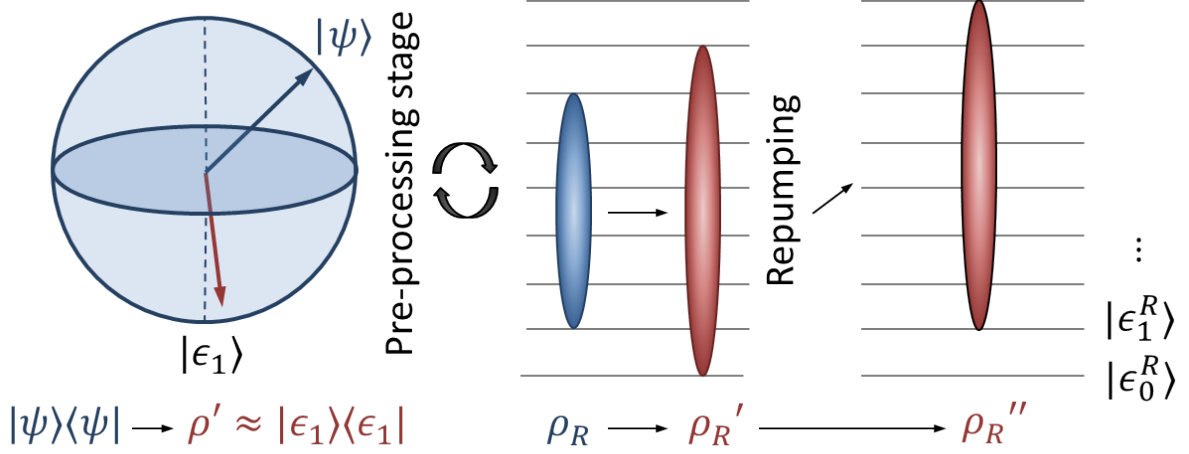


Figure 3.2: **Basic work-extraction protocol.** The evolution of the system from the initial state  $|\psi\rangle\langle\psi|$  to the final state  $\rho'$  is depicted on the Bloch ball, in blue and red respectively. The evolution of the reference from the initial state (smaller blue blob) throughout the protocol (red blobs) is depicted on the energy level ladder.

so that  $U$  rotates the qubit system from  $|\psi\rangle\langle\psi|$  to  $|\epsilon_1\rangle\langle\epsilon_1|$ .<sup>4</sup>

2. **Work extraction.** The system is now in a state  $\rho'$  and, due to work-locking (see Sec. 3.1.1), is indistinguishable from its dephased version in any work extraction protocol. So without loss of generality we can use the dephased version

$$\mathcal{D}(\rho') = q |\epsilon_0\rangle\langle\epsilon_0| + (1 - q) |\epsilon_1\rangle\langle\epsilon_1|. \quad (3.18)$$

Now, depending on the considered regime, single-shot or average (asymptotic) work can be extracted from the dephased state and stored in the battery.

3. **Repumping.** The back-reaction changes the state of the reference into  $\rho'_R$ . Using part of the extracted work (stored in the battery during the previous step) we can repump the reference to shift it up:

$$\rho'_R \rightarrow \rho''_R := \Delta \rho'_R \Delta^\dagger. \quad (3.19)$$

Although one could question such repumping, given that  $\Delta$  is not a unitary, we note that this is actually not a problem. This is because such operation can be realised

<sup>4</sup>This interaction corresponds to a modified Jaynes-Cummings model (with excitation-dependent coupling strengths). However, as shown in Ref. [14], it can also be approximately realised within the standard Jaynes-Cummings model using a reference in a coherent state  $|\alpha\rangle$ , with  $|\alpha|$  large enough.

through a joint energy-conserving unitary between the reference in a state  $\rho'_R$  and a qubit battery system in an excited state  $|\epsilon_1^W\rangle$  with  $\epsilon_1^W = 1$ . This unitary is given by  $V(U)$  from Eq. (3.15), where we take  $U = \sigma_x$ , the Pauli  $x$  operator. Then

$$V(\sigma_x) = \sigma_- \otimes \Delta + \sigma_+ \otimes \Delta^\dagger + |\epsilon_0^W\rangle\langle\epsilon_0^W| \otimes |\epsilon_0^R\rangle\langle\epsilon_0^R|, \quad (3.20)$$

where  $\sigma_+ = |\epsilon_1^W\rangle\langle\epsilon_0^W|$  and  $\sigma_- = |\epsilon_0^W\rangle\langle\epsilon_1^W|$ . As a result the final state of the weight system is  $|\epsilon_0^W\rangle$  and the final state of the reference is given by Eq. (3.19). We will describe how often we perform the repumping while analysing different variations of the protocol.

4. **Repeat.** We can repeat the protocol using  $\rho''_R$  and a fresh copy of  $|\psi\rangle\langle\psi|$ .

### 3.2.2 Performance

During the pre-processing stage the joint unitary  $V(U)$  approximately induces  $U$  on the system:

$$\rho' \approx U(|\psi\rangle\langle\psi|)U^\dagger = |\epsilon_1\rangle\langle\epsilon_1|. \quad (3.21)$$

The degree to which the above equation holds depends on the *quality* of the reference as defined in Ref. [14]. In particular, the system's final occupation in the excited state  $(1 - q) = \langle\epsilon_1|\rho'|\epsilon_1\rangle$  is given by

$$1 - q = 1 - 2p(1 - p)(1 - \langle\bar{\Delta}\rangle) - p^2 R_{00}, \quad (3.22)$$

where  $R_{00} = \langle\epsilon_0^R|\rho_R|\epsilon_0^R\rangle$  and  $\langle\bar{\Delta}\rangle$  is the quality parameter from Definition 3.1. From Eq. (3.22) it is easy to see that  $(1 - q) \rightarrow 1$  when  $R_{00} \rightarrow 0$  (i.e., when the reference quality parameter  $M > 0$ ) and  $\langle\bar{\Delta}\rangle \rightarrow 1$ . Therefore,  $R_{00}$  and  $\langle\bar{\Delta}\rangle$  are operationally well-defined quality parameters of the reference, because they directly measure the ability of the reference to induce the unitary  $U$  that we want to perform on the qubit. At the same time the reference undergoes a back-reaction induced by the joint unitary  $V(U)$ . This is described by the following Kraus operators:

$$A_0 = (1 - \sqrt{1 - p})\sqrt{p}|\epsilon_0^R\rangle\langle\epsilon_0^R| + \sqrt{p(1 - p)}(\mathbb{I} - \Delta), \quad A_1 = (1 - p)\mathbb{I} + p\Delta^\dagger, \quad (3.23)$$

so that the reference state after performing the pre-processing stage is given by:

$$\rho'_R = A_0\rho_R A_0^\dagger + A_1\rho_R A_1^\dagger. \quad (3.24)$$

From Eqs. (3.18) and (3.22) the only two parameters relevant for work extraction are the reference population in the ground-state,  $R_{00}$ , and the parameter  $\langle \bar{\Delta} \rangle$ . Using the Kraus operators specified by Eq. (3.23), the change in  $\langle \bar{\Delta} \rangle$  during the pre-processing stage (i.e., the difference between the final and initial value of  $\langle \bar{\Delta} \rangle$ ) can be computed and is found to be:

$$\delta \langle \bar{\Delta} \rangle = p \left[ (\sqrt{1-p} - 1) \text{Re}(R_{01}) - \sqrt{1-p} R_{00} \right], \quad (3.25)$$

where as before  $R_{ij} = \langle \epsilon_i^R | \rho_R | \epsilon_j^R \rangle$ . A sufficient condition for  $\langle \bar{\Delta} \rangle$  to stay constant is  $R_{00} = 0$ , i.e.,  $M > 0$ . Therefore, if the initial state satisfies  $R_{00} = 0$  exactly, performing the pre-processing stage does not change  $\langle \bar{\Delta} \rangle$ , as noted in Ref [14]. We require Step 3 of the protocol to ensure that  $\langle \epsilon_0^R | \rho_R'' | \epsilon_0^R \rangle = 0$ . If this is the case, at the end of the protocol we are left with the reference described by the same quality parameters  $R_{00}$  and  $\langle \bar{\Delta} \rangle$  as at the beginning, and the reference  $\rho_R''$  is as good as  $\rho_R$  within the protocol.

Finally, because the state of the reference changes, its free energy can fluctuate. However, notice firstly that the reference has a Hamiltonian bounded from below, so for fixed average energy it has a finite amount of free energy. Secondly, repeatability requires that the reference can be used an arbitrary number of times and the performance of the protocol never changes. It is then easy to see that on average the free energy of the reference cannot be extracted as work, as this would be incompatible with repeatability. In the worst-case scenario the free energy change in the reference fluctuates around zero. To show this let us denote by  $\Delta F_{R,n}$  the change in the free energy of the reference at the  $n$ -th repetition of the protocol. The total free energy change of the reference after  $M$  repetitions of the protocol satisfies

$$\sum_{n=1}^M -\Delta F_{R,n} \leq F(\rho_R) - F(\gamma_R), \quad (3.26)$$

where  $\gamma_R$  is the thermal state of the reference. Hence, the average change in the free energy of the reference as  $M \rightarrow \infty$  is

$$-\overline{\Delta F_R} := \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{n=1}^M -\Delta F_{R,n} = 0. \quad (3.27)$$

Moreover, these fluctuations vanish in the limit of an unbounded reference  $\langle \bar{\Delta} \rangle \rightarrow 1$ , as then the entropy of the reference stays constant, while its average energy increases. Therefore its free energy must increase at every step of the protocol.

### 3.3 Fundamental limitations of coherence-to-work conversion

We will now analyse how well the presented protocol does in terms of work extraction from coherence. First, we will emphasise the limitations that arise due to the reference being bounded, i.e., when we have access to limited coherence resources. More precisely, we will explain why the use of a bounded reference does not allow one to extract (from a state with coherence) an average amount of work equal to the free energy difference. However, we will then also show that one can construct a series of bounded reference states that come arbitrarily close to extracting the free energy  $\Delta F(|\psi\rangle\langle\psi|)$ , with protocols arbitrarily close to perfect repeatability. Thus, we will prove that in the limit of an unbounded reference all coherence can be converted into work in a repeatable way. The limiting case does not generate any entropy in the reference system and, being a reversible transformation, is optimal.

#### 3.3.1 Limitations of bounded references

In order to illustrate the limitations arising from using a bounded reference we will consider a particular model of work extraction from coherence described in Refs. [67, 68]. It was proven there that the free energy difference  $\Delta F(\rho)$  can be extracted from a system  $\rho$  as work if one allows for the use of operations that conserve energy only on average. Let us briefly recall the protocol used to achieve this. Without loss of generality we can write any state  $\rho$  as  $\sum_i p_i |\psi_i\rangle\langle\psi_i|$  with  $p_{i+1} \leq p_i$ . Let us also denote the Hamiltonian of the system by  $H = \sum_i \epsilon_i |\epsilon_i\rangle\langle\epsilon_i|$  with  $\epsilon_{i+1} \geq \epsilon_i$ . The protocol is composed of two stages. First, the system  $\rho$  interacts with a weight system in a gravitational field via the unitary

$$U_{\text{av}} = \sum_i |\epsilon_i\rangle\langle\psi_i| \otimes \Gamma_{E_i}, \quad (3.28)$$

where  $\Gamma_{E_i}$  is the shift operator on the weight system that shifts it up in energy by  $E_i = \langle\psi_i| H |\psi_i\rangle - \epsilon_i$ . Since  $U_{\text{av}}$  is a unitary, the decrease in the average energy of the system,  $\sum_i p_i E_i$ , is interpreted as work performed on the weight system. The resultant state  $\rho'$  is incoherent and can be described by a probability vector  $\mathbf{p}$ .

In the second step, work is extracted from the classical state  $\mathbf{p}$ . In accordance with Assumption 3.1, this process extracts  $\Delta F(\mathbf{p})$ . Hence, the extraction of the full free energy

$\Delta F(\rho)$  from a state  $\rho$  is equivalent to the possibility of extracting during the first step the amount of work equal to

$$\Delta F(\rho) - \Delta F(\mathbf{p}) = \frac{1}{\beta} S(\rho || \mathcal{D}(\rho)) + \Delta F(\mathbf{q}) - \Delta F(\mathbf{p}), \quad (3.29)$$

where  $\mathbf{q}$  is the probability vector describing the incoherent state  $\mathcal{D}(\rho)$ . This is equivalent to extracting  $S(\rho || \mathcal{D}(\rho))/\beta$  from the coherence of  $\rho$  (recall that this quantity is a known measure of quantum coherence [80] that we will also study in Chapter 5). Hence, the amount of work that needs to be extracted on average from the coherence of a quantum state to achieve the free energy extraction limit for arbitrary quantum states is

$$W_{\text{coh}}(\rho) = \frac{1}{\beta} S(\rho || \mathcal{D}(\rho)). \quad (3.30)$$

However, as  $U_{\text{av}}$  does not strictly conserve energy, it is not a free thermodynamic operation. One can instead ask if it can be achieved by an energy-preserving unitary  $V(U_{\text{av}})$  on a larger system that exploits some ancillary system  $\rho_A$ . In other words, we are looking for the energy-preserving unitary  $V(U_{\text{av}})$  such that

$$\mathcal{E}(\rho) := \text{Tr}_A (V(U_{\text{av}})(\rho \otimes \rho_A)V(U_{\text{av}})^\dagger) = U_{\text{av}} \rho U_{\text{av}}^\dagger. \quad (3.31)$$

Due to the imposed constraints, the ancillary system  $\rho_A$  must carry quantum coherence. In fact, if  $\rho_A$  were incoherent, then the left-hand side of Eq. (3.31) would be a time-translation covariant quantum map (meaning that  $[\mathcal{E}, \mathcal{D}] = 0$ ) [61], whereas the right-hand side is not.

Now the crucial point is that Eq. (3.31) cannot hold exactly unless  $\rho_A$  contains an unbounded reference. If  $\rho_A$  is bounded, then the reduced evolution of  $\rho$  is not exactly unitary and not all the energy change can be identified with work. To see this, first assume that  $\rho$  is a pure state. This means that the mutual information of the final state vanishes,  $I(\rho_{SA}) = 0$ . However, the only way this is possible is if  $V(U_{\text{av}}) = V_1 \otimes V_2$ , and further we would need  $V_1 = U_{\text{av}}$ . But this is not possible, because from the fact that  $V(U_{\text{av}})$  is strictly energy-preserving we can prove that  $V_1$  and  $V_2$  must both be as well. Similar reasoning can be applied for a mixed state  $\rho$  if we demand that the entropy of the ancillary system should not increase. Hence, the right-hand side of Eq. (3.31) cannot be a unitary if  $\rho_A$  contains only a bounded reference frame. In fact, Eq. (3.31) can only hold as a limiting case of using a larger and larger coherence resource. In summary, the framework of thermal operations and Assumption 3.1, together with identification of work

with energy change during unitary processes, imply that without an unbounded reference the work extraction protocol from Refs. [67, 68] cannot extract an amount of work equal to  $\Delta F(\rho)$  from a state with coherence.

The strength of this point is that we do not even need to require the repeatability of the protocol using the same ancillary system. In fact, the argument is rather general. In order to extract all the free energy from a state  $\rho$  one needs to transform it into a thermal state. This cannot be achieved by only changing the energy spectrum of  $H$ , but also requires the rotation of the energy eigenbasis, so that the system is incoherent at the end of the transformation. This can be performed perfectly only with the aid of an unbounded reference frame, because it involves unitaries that do not strictly conserve energy.<sup>5</sup>

### 3.3.2 Approaching free energy limit of work extraction

A key fact about the Carnot efficiency is that, despite being achieved only by ideal heat engines that do not actually exist in nature, we can get arbitrarily close to it through a sequence of real engines. In a similar spirit, we will now explain how to construct a sequence of bounded references that can get arbitrarily close to the coherence-to-work conversion limit set by Eq. (3.30). In other words, we will show that there exists a sequence of bounded references that allow one to approach the ideal conversion with arbitrarily high probability of success and with an arbitrarily small change in the reference quality parameters. Moreover, the limiting case is reversible and thus Eq. (3.30) provides the ultimate limit of converting coherence to average work.

After this non-technical summary of the main result of this section, let us now proceed to a detailed and rigorous statement. We want to show how one can perform the protocol described in Sec. 3.2 to extract an amount of work arbitrarily close to the free energy difference,  $\Delta F(|\psi\rangle\langle\psi|)$ , while succeeding with arbitrarily high probability and changing the quality of the reference by only a negligible amount. For simplicity, define  $f(x) = -x - h_2(x)/\beta$ , where  $h_2(x) = -x \ln x - (1-x) \ln(1-x)$  denotes the binary entropy. Our claim can be now made technically precise with the following theorem:

---

<sup>5</sup>A useful point of view is also given by the theory of quantum reference frames and recovery maps [52, 81].

**Theorem 3.1: Approaching the ideal coherence to work conversion**

Let  $\rho_R$  be an arbitrary reference state described by  $(\langle\bar{\Delta}\rangle, M)$ . In the presence of a thermal bath at inverse temperature  $\beta$  and if  $M$  is large enough, there exists a protocol individually extracting from  $M$  copies of  $|\psi\rangle\langle\psi|$  [given by Eq. (3.12)] an average amount of work  $M\langle W\rangle$ , with

$$\langle W\rangle \geq \Delta F(|\psi\rangle\langle\psi|) - f(2p(1-p)(1 - \langle\bar{\Delta}\rangle)) - O(M^{-\frac{1}{3}}). \quad (3.32)$$

The probability of success  $p_{\text{succ}}$  of the protocol is

$$p_{\text{succ}} \gtrsim [1 - 2p(1-p)(1 - \langle\bar{\Delta}\rangle)]^M, \quad (3.33)$$

i.e., it approaches unity when  $M(1 - \langle\bar{\Delta}\rangle) \rightarrow 0$ . The protocol changes the quality parameters of the reference as follows:

$$\delta M = 0, \quad \delta\langle\bar{\Delta}\rangle \leq 2\sqrt{1 - p_{\text{succ}}}. \quad (3.34)$$

Before presenting the main steps of the proof of this theorem (the details of calculations can be found in Appendix B), let us first comment on its scope. Note that the same result holds when a reference  $(\langle\bar{\Delta}\rangle, M)$  is used a number of times  $M' < M$ , as long as  $M' \gg 1$  (this will be clear from the proof). In the case in which  $M' > M$ , we can apply the Theorem every  $M$  uses of the reference. The changes in the quality parameters will eventually sum up, but the theorem gives a bound on them. Also, it will be clear from the proof that failure of the protocol implies a destruction of the coherence properties of the reference.

The theorem is proven by constructing an explicit variation of the protocol introduced in Sec. 3.2 and showing that it performs as claimed. Recall that by an energy conserving unitary we can rotate  $|\psi\rangle\langle\psi|$  around the  $z$  axis of the Bloch sphere. Hence, without loss of generality, we can set  $\varphi = 0$  in Eq. (3.12). We then perform Steps 1-2 of the protocol described in Sec. 3.2  $M$  times, i.e., individually processing each of  $M$  copies of  $|\psi\rangle\langle\psi|$  using a reference  $\rho_R$  described by  $(\langle\bar{\Delta}\rangle, M)$ . The choice of  $\rho_R$  ensures that during this process the reference state will have no population in the ground state, and so  $\langle\bar{\Delta}\rangle$  will stay constant. Then, the final state of the reference is described, with probability  $p_1^M$ , by  $\rho_{R,1} = A_1^M \rho_R A_1^{\dagger M} / p_1^M$ , where  $p_1 = (1 - 2p(1-p)(1 - \langle\bar{\Delta}\rangle))$ . Notice that having access to

a reference described by parameters  $\langle \bar{\Delta} \rangle$  and  $M$  such that  $M(1 - \langle \bar{\Delta} \rangle) \rightarrow 0$ , the probability  $p_1^M$  can be made arbitrarily close to 1. This happens because by taking  $\langle \bar{\Delta} \rangle$  close enough to 1, we can get arbitrarily close to unitary evolution of a system state  $|\psi\rangle\langle\psi|$  into a pure, incoherent state  $|\epsilon_1\rangle\langle\epsilon_1|$ . As the final state of the system is almost pure, the final joint state of the system and the reference factorises and an arbitrarily small amount of entropy is generated (the back-reaction on the reference is given by the Kraus operator  $A_1$  alone).

Next, we repump the reference  $\bar{M} + s\sigma_M^{4/3}$  times, where  $\bar{M} = Mp$ ,  $\sigma_M = \sqrt{Mp(1-p)}$  and  $s > 0$ . This guarantees that the reference has arbitrarily small population in states  $\{|\epsilon_0^R\rangle \dots |\epsilon_M^R\rangle\}$ , so that by performing a measurement we can project the reference to a state  $\rho_R''$  with support on the subspace spanned by  $\{|\epsilon_i^R\rangle\}_{i>M}$  with arbitrarily high probability ( $s$  fixes the confidence level, see Appendix B for details). More precisely, after repeating Steps 1-2 of the protocol  $M$  times and repumping as explained above, the reference is described by a state  $\rho_R''$  with probability

$$p_{\text{succ}} \geq [1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle)]^M \mathbb{E}_s(M^{1/6}), \quad (3.35)$$

where  $\mathbb{E}_s(x) = \text{erf}(sx/\sqrt{2})$  and erf denotes the error function. The final state is given by  $\rho_R''$ , described by  $(\langle \bar{\Delta} \rangle'', M'')$ , where  $M'' = M$  and

$$\delta\langle \bar{\Delta} \rangle := \langle \bar{\Delta} \rangle'' - \langle \bar{\Delta} \rangle \leq 2\sqrt{1 - p_{\text{succ}}} \leq 2\sqrt{1 - p_1^M \mathbb{E}_s(M^{1/6})}. \quad (3.36)$$

Notice that by taking  $s$  large enough (but finite) we can make the factor  $\mathbb{E}_s(M^{1/6})$  in the previous two equations arbitrarily close to 1,  $\mathbb{E}_s(M^{1/6}) \approx 1$ . In the appropriately chosen limit  $\langle \bar{\Delta} \rangle \rightarrow 1$  and  $M \rightarrow \infty$  the quality parameters of the reference state are then unchanged with probability 1. Let us also note that the cost of the measurement  $W_E$  described above is bounded by  $h_2(p_{\text{succ}})/\beta$  (see Appendix B).

We have just shown that following the procedure above we can guarantee repeatability with arbitrary confidence level. Hence, we now proceed to proving that it also allows for extracting an average amount of work per system arbitrarily close to the free energy difference  $\Delta F(|\psi\rangle\langle\psi|)$ . To see this, note that after repeating the protocol on  $M$  copies of  $|\psi\rangle\langle\psi|$  we are left with  $M$  copies of a state  $\mathcal{D}(\rho')$  from Eq. (3.18) with  $(1-q)$  given by Eq. (3.22),  $(1-q) = 1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle)$ . This state is diagonal in the energy eigenbasis and the average work  $\langle \tilde{W} \rangle$  extracted from it is given by  $\Delta F(\mathcal{D}(\rho'))$ :

$$\langle \tilde{W} \rangle = 1 + \frac{1}{\beta} \log Z - 2p(1-p)(1 - \langle \bar{\Delta} \rangle) - \frac{1}{\beta} h_2(2p(1-p)(1 - \langle \bar{\Delta} \rangle)), \quad (3.37)$$

where as before  $h_2(\cdot)$  denotes the binary entropy. By choosing  $M$  large enough, we can ensure that the extracted work is arbitrarily peaked around the average given by the above equation ( $M$  can be bounded using the results of Ref. [11]). This ensures that when we need to repump the reference, we actually have enough work to invest to do it. The repumping costs  $\bar{M} + s\sigma_M^{4/3}$  units of extracted work and the cost of measurement  $W_E$  is independent from  $M$ . Hence, the net gain per processed copy of  $|\psi\rangle\langle\psi|$  is given by

$$\langle\tilde{W}\rangle - \frac{(\bar{M} + s\sigma_M^{4/3}) + W_E}{M} = \langle W \rangle - O(M^{-1/3}), \quad (3.38)$$

where

$$\langle W \rangle = \Delta F(|\psi\rangle\langle\psi|) - 2p(1-p)(1 - \langle\bar{\Delta}\rangle) - \frac{1}{\beta}h_2(2p(1-p)(1 - \langle\bar{\Delta}\rangle)). \quad (3.39)$$

Therefore, the deficit per copy scales as  $M^{-1/3}$  and by choosing  $M$  large enough it can be made arbitrarily small. Moreover, the previous equation gives us the relation between the quality of the reference and the average extracted work, showing that  $\langle W \rangle \rightarrow \Delta F(|\psi\rangle\langle\psi|)$  as  $\langle\bar{\Delta}\rangle \rightarrow 1$ ,  $M \rightarrow \infty$ ,  $M(1 - \langle\bar{\Delta}\rangle) \rightarrow 0$ .

We conclude that it is possible, with arbitrarily large success probability, to extract an amount of work arbitrarily close to the free energy change from a pure state with coherence. Moreover, it can be achieved, while processing the states individually and properly taking account of all the resources used, i.e., ensuring arbitrarily exact repeatability.

### 3.4 Extracting work with perfect repeatability and bounded references

In the previous section we have shown how to extract an amount of work equal to the free energy of a pure quantum state with coherence. However we allowed for

1. The limiting case of an unbounded reference,  $\langle\bar{\Delta}\rangle \rightarrow 1$ .
2. An asymptotic protocol individually processing a large number  $M$  of copies of the system.

These assumptions may be too strong if the reference itself is a microscopic system involved in the thermodynamic processing, and their applicability to single-shot scenarios

is excluded. What if we only want to process a small number of systems? This requires us to go beyond the results of the previous section.

Moreover, even if we only want to release the first of the two assumptions, i.e., put a bound on the coherence properties of the reference, we are still left with open questions. In this case the general result stated by Theorem 3.1 is applicable, however the presented work extraction protocol always entails a failure probability  $1 - p_{\text{succ}}$  that can lead to a complete destruction of coherent properties of the reference. Even if this probability is relatively small, we may not be willing to take this risk. Also, the reference inevitably deteriorates, even if by a small amount bounded by Eq. (3.36). A crucial question is then: are there work extraction protocols with  $\langle \bar{\Delta} \rangle < 1$  such that  $\delta \langle \bar{\Delta} \rangle = 0$  and  $p_{\text{succ}} = 1$ ? In other words, can we extract work from coherence using a protocol that *never* fails and gives back the reference with *exactly* the same quality parameters, even if the reference is bounded?

Here we will construct such protocols for both average and single-shot work extraction. These ensure perfect repeatability, but the price we pay is that the average amount of extracted work is strictly smaller than the free energy difference and it is only possible for  $\langle \bar{\Delta} \rangle$  above a certain threshold value  $\bar{\Delta}_{\text{crit}}$ . In the case of single-shot work extraction we show similarly that there exists a threshold over which the reference allows us to outperform the single-shot protocol with no coherence. For clarity of the discussion, we focus on the paradigmatic case of the class of states  $|\gamma\rangle\langle\gamma|$  introduced in Eq. (3.14).

### 3.4.1 Average work extraction

In absence of an external source of coherence no work can be extracted from the state  $|\gamma\rangle\langle\gamma|$  on average [9, 13]. However, if we allow for a repeatable use of the reference, positive work yield can be obtained. In order to achieve this, during Step 2 of the protocol we perform average work extraction from the state  $\mathcal{D}(\rho')$  specified by Eq. (3.18). As  $\mathcal{D}(\rho')$  is diagonal in the energy eigenbasis, by Assumption 3.1 we have that the average work yield is given by the free energy difference  $\Delta F(\mathcal{D}(\rho')) = \Delta F(\mathbf{q})$ . To ensure perfect repeatability we repump the reference at each run, so that if  $\langle \epsilon_0 | \rho_R | \epsilon_0 \rangle = 0$ , then  $\langle \epsilon_0 | \rho_R'' | \epsilon_0 \rangle = 0$ , and the reference quality parameters do not change. The repumping requires a unit of work,

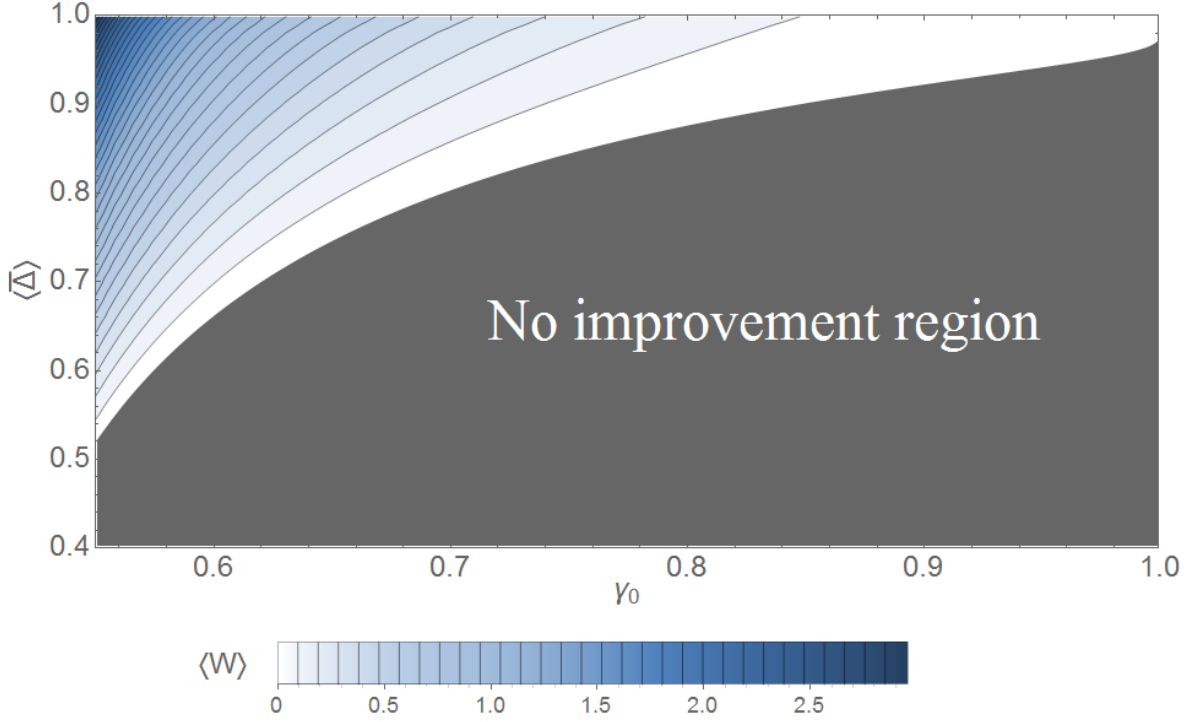


Figure 3.3: *Coherence boost to average work extraction from  $|\gamma\rangle$* . Work can be unlocked from the coherence of the system using a reference that never deteriorates. The quality of the reference, measured by  $\langle \bar{\Delta} \rangle$ , must be bigger than some threshold value (boundary of the gray region) to ensure  $\langle W \rangle > 0$  ( $\gamma_0$  denotes the thermal occupation of the ground state,  $\gamma_0 = 1/Z$ ). Over the threshold, the higher the quality the greater is the average work yield from quantum coherence. Graph plotted for  $\gamma_0 \in [0.55, 1]$  as  $\langle W \rangle$  diverges when  $\gamma_0 \rightarrow 0.5$ , which corresponds to a heat bath at infinite temperature.

so that the work extracted on average during one run of the protocol is given by:

$$\langle W \rangle = \Delta F(\mathbf{q}) - 1 = \frac{1}{\beta}(\log Z - h_2(q)) - q, \quad (3.40)$$

where  $Z = 1 + e^{-\beta}$  is the partition function (recall that we have chosen  $\epsilon_0 = 0$  and  $\epsilon_1 = 1$ , so that  $\gamma_0 = 1/Z$ ). The connection between the properties of the reference and the work yield, described above by  $q$ , is given by Eq. (3.22) with  $R_{00} = 0$  and  $p = \gamma_0$ , i.e.,

$$q = 2\gamma_0(1 - \gamma_0)(1 - \langle \bar{\Delta} \rangle). \quad (3.41)$$

In Fig. 3.3 we show how much work on average,  $\langle W \rangle$ , can be unlocked through our protocol as a function of the quality of the reference  $\langle \bar{\Delta} \rangle$  and the thermal occupation  $\gamma_0$  of the ground state. The graph shows that the quality of the reference needs to be above a certain threshold in order to get positive average work yield. As expected, the advantage

is the most significant for low  $\gamma_0$ , because the states  $|\gamma\rangle\langle\gamma|$  and  $\gamma$  differ most in this case (the amount of coherence to be unlocked is bigger), and because the heat bath has higher temperature .

As already mentioned in Sec. 3.1, in the asymptotic regime of individually processing large number of copies of  $|\gamma\rangle\langle\gamma|$ , the fluctuations in the work yield, Eq. (3.40), become negligible. Notice, however, that even if  $\langle W \rangle > 0$  we may not be able to perform Step 3 every time, as the fluctuations around the average mean that we will not always have enough work to invest in the repumping. To resolve this problem we can follow a strategy analogous to the case of unbounded reference. That is, we repump after having extracted work  $M$  times, where  $M$  is sufficiently large to neglect the fluctuations around  $\langle W \rangle$ . The protocol will be repeatable up to an arbitrarily small probability of failure, if the support of the reference initially starts high enough in the energy ladder. It is important to stress, however, that the “failure” in this case does not entail a destruction of the coherence properties of the reference, as in Sec. 3.3.2. It only requires the investment of extra work in order to ensure perfect repeatability.

### 3.4.2 Single-shot work extraction

Finally, we proceed to a single-shot protocol for an individual quantum state. This version of the protocol does not assume possessing an unbounded reference nor it requires asymptotic number of runs. In absence of an external source of coherence we can perform  $\varepsilon$ -deterministic work extraction from  $|\gamma\rangle\langle\gamma|$  – as  $|\gamma\rangle\langle\gamma|$  is indistinguishable from  $\gamma$ , we can extract  $\log Z/\beta$  work with failure probability  $\varepsilon_0 := (1 - \gamma_0)$  or  $(1 + \log Z/\beta)$  with failure probability  $\gamma_0$  [see Eq. 3.2 and Eqs. (1.36)-(1.37)]. We now note that while exploiting the reference in a perfectly repeatable way, the failure probability  $\varepsilon_0$  for extracting  $\log Z/\beta$  can be decreased – and the higher the quality of the reference, the stronger the improvement. To achieve this, during Step 2 of the protocol we perform  $\varepsilon$ -deterministic work extraction from the state  $\mathcal{D}(\rho')$  specified by Eq. (3.18), in accordance with Assumption 3.1. With probability  $(1 - q)$  we extract  $1 + \log Z/\beta$  work and with probability  $\tilde{\varepsilon} := q$  our protocol fails. As we need one unit of work to repump the reference [see Eq. (3.19)], the net gain is  $\log Z/\beta$ . When the protocol fails (with probability  $q$ ), the reference is returned in the state  $\rho'_R$  and one has to invest one unit of work to ensure repeatability.

In Fig. 3.4 we present the relative decrease in the failure probability achieved by our

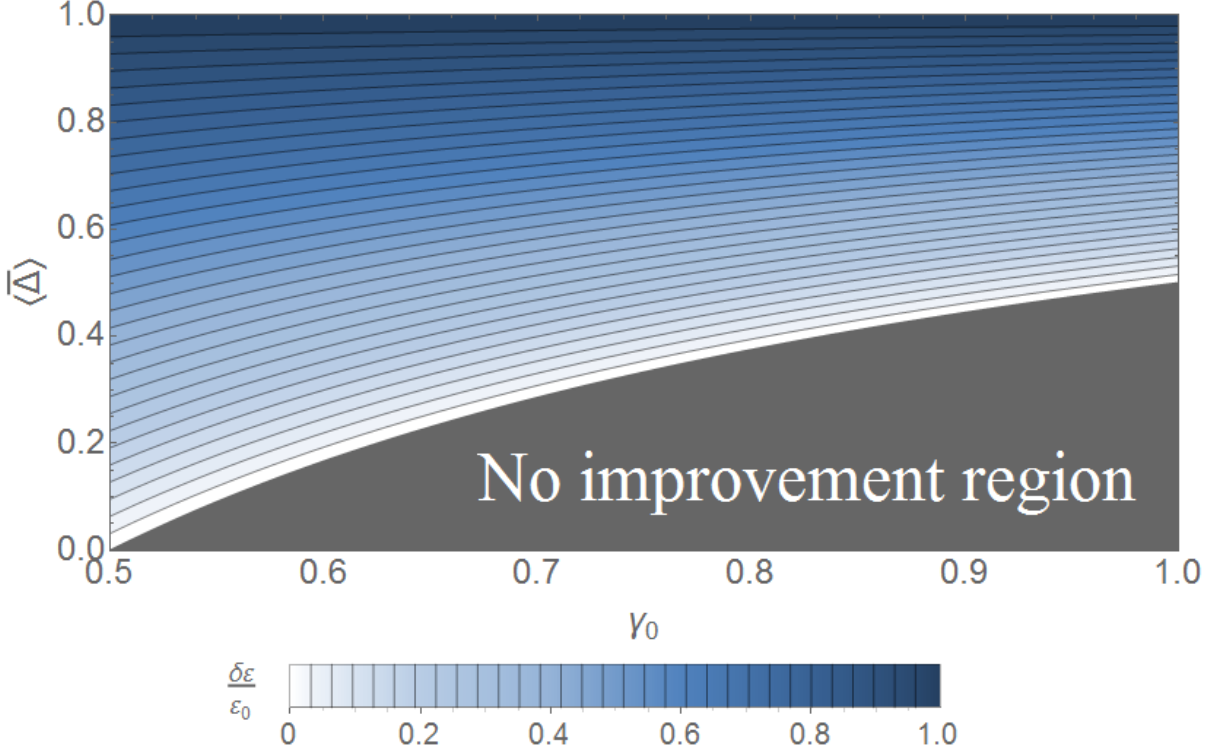


Figure 3.4: *Coherence boost to single-shot work extraction from  $|\gamma\rangle$ .* A reference used in a repeatable way can exploit the quantum coherence of the system to decrease the failure probability in single-shot work extraction. Similarly to the case of average work extraction, the quality of the reference  $\langle \bar{\Delta} \rangle$  must be over some threshold value to lead to any improvement ( $\gamma_0$  denotes the thermal occupation of the ground state). As  $\langle \bar{\Delta} \rangle$  increases, the relative change in the failure probability increases to 1, i.e., to the point when the single-shot work extraction from the pure quantum state  $|\gamma\rangle$  becomes fully deterministic.

protocol as compared to work extraction from  $\gamma$ :

$$\frac{\delta \varepsilon}{\varepsilon_0} = \frac{\varepsilon_0 - \tilde{\varepsilon}}{\varepsilon_0} = 1 - 2\gamma_0(1 - \langle \bar{\Delta} \rangle). \quad (3.42)$$

We see that if the quality of the reference is high enough (the threshold value being  $1 - (2\gamma_0)^{-1}$ ), the coherence content of  $|\gamma\rangle\langle\gamma|$  can be exploited to provide an advantage in the work extraction. In the limit of a very high quality (unbounded) reference,  $\langle \bar{\Delta} \rangle \rightarrow 1$ , the failure probability can be sent to zero, i.e., the work extraction from  $|\gamma\rangle\langle\gamma|$  becomes deterministic.

# Chapter 4

## Order theory approach

In the previous two chapters we have studied the symmetry constraints on thermodynamic processing of quantum states that arise from the fact that thermal operations are time-translation covariant. However, as already discussed in Sec. 1.1 of Chapter 1, another crucial property of thermal operations is that they preserve the thermal Gibbs state. In order to better understand the consequences of this property, in this final chapter of Part I we will focus solely on Gibbs-preserving maps. Thus, we will study only “entropic” aspects of thermal operations – recall that GP operations form the biggest set of CPTP maps that does not allow one to create a perpetual mobile, and thus, in a sense, the restriction of free operations to GP maps captures the very essence of the second law. Our main aim will be to investigate the structural properties of the *thermodynamic ordering* of states. This ordering (that we will rigorously define later) is induced by the laws of thermodynamics on the system state space: whenever a state  $\rho$  can freely, i.e., without any extra thermodynamic resources used, evolve to a state  $\sigma$ , then  $\rho$  precedes  $\sigma$ . In other words, the ordering under scrutiny, which we will also refer to as the *thermodynamic arrow of time*, is defined by thermal cones (recall Definition 1.1): states in the future thermal cone  $\mathcal{T}_+(\rho)$  succeed  $\rho$ , whereas the ones in the past thermal cone  $\mathcal{T}_-(\rho)$  precede it.

In classical equilibrium thermodynamics the ordering of states is particularly simple. A transformation from a state  $A$  to  $B$  is possible if and only if the final state has smaller free energy than the initial one (as then no work has to be invested, see Sec. 1.2.4 of Chapter 1). The ordering between equilibrium states is then fully specified by one function and, therefore, we deal with a total order: either  $A$  can be freely transformed to  $B$ , or  $B$  can be freely transformed to  $A$ . However, as we have seen in Sec. 1.2.2 of Chapter 1, the

ordering that arises in the resource theory of thermodynamics between classical states is much more complicated. Recall that to be able to transform a classical state  $\mathbf{p}$  into  $\mathbf{q}$ , it is required that a whole family of  $\alpha$ -Rényi divergences,  $S_\alpha(\mathbf{p}||\boldsymbol{\gamma})$ , has to decrease. Hence, we deal with a partial order, as not every two states are comparable, i.e., it may happen that neither  $\mathbf{p}$  can be transformed to  $\mathbf{q}$ , nor the other way round.

Here, instead of studying the ordering induced by GP maps for a particular state, we will focus on the global properties of the thermodynamic arrow of time. More precisely, a partial order is a very general structure, studied within the field of mathematics known as *order theory*, defined as follows:

**Definition 4.1: Partial order and preorder**

A partial order is a binary relation  $\succ$  over a set  $S$  that satisfies for all  $s_1, s_2, s_3 \in S$  the following three conditions:

1. Reflexivity:  $s_1 \succ s_1$ .
2. Transitivity: if  $s_1 \succ s_2$  and  $s_2 \succ s_3$  then  $s_1 \succ s_3$ .
3. Antisymmetry: if  $s_1 \succ s_2$  and  $s_2 \succ s_1$  then  $s_1 = s_2$ .

Binary relations satisfying only the first two properties are known as *preorders* and are usually denoted by  $\succsim$ .

Being such a broad and general concept, it seems natural to ask whether thermodynamic ordering has a more rigid and specific structure. Inspired by order-theoretic studies we will focus on a special kind of partial order known as a *lattice*<sup>1</sup> and interpret it from a thermodynamic perspective. We will show that in the infinite temperature limit the thermodynamic arrow of time actually reflects the structure of a lattice, but it is lost in the classical regime at finite temperatures. Surprisingly, however, we will also prove that in the simplest quantum scenario of a two-dimensional system, the lattice structure is preserved at finite temperatures. This suggests that coherence can play a role in providing structure to the thermodynamic arrow of time.

<sup>1</sup>It is interesting to note that the order-theoretic concept of a lattice has been used in physics before to study the structure of quantum logic [82].

## 4.1 Order theory

We will start by introducing the necessary concepts from the field of order theory, so that we can investigate the structure of thermodynamic ordering more rigorously. As the considerations from this section can be applied to both classical states (represented by probability distributions), as well as to general quantum states, for the clarity of discussion we will just restrict to the latter generic case. To obtain classical statements one simply needs to replace a density matrix  $\rho \in \mathcal{S}_d$  with a probability vector  $\mathbf{p} \in \mathcal{P}_d$  and a GP quantum channel  $\mathcal{E}$  with a GP stochastic matrix  $\Lambda$ .

### 4.1.1 Partial order

Since the identity operation is a Gibbs-preserving CPTP map and a composition of two GP operations is also GP, we have that  $\rho \in \mathcal{T}_\pm(\rho)$  and

$$(\rho \in \mathcal{T}_\pm(\sigma) \text{ and } \sigma \in \mathcal{T}_\pm(\tau)) \Rightarrow \rho \in \mathcal{T}_\pm(\tau).$$

Hence, the relation of one state belonging to a thermal cone of another state is reflexive and transitive. Thus, the set of all past thermal cones induces a preorder  $\succsim_-$  “oriented along” the thermodynamic arrow of time, i.e.,  $\rho$  precedes  $\sigma$ ,  $\rho \succsim_- \sigma$ , if  $\rho \in \mathcal{T}_-(\sigma)$ . On the other hand the set of all future thermal cones induces a preorder  $\succsim_+$  “oriented against” the thermodynamic arrow of time, i.e.,  $\rho$  precedes  $\sigma$ ,  $\rho \succsim_+ \sigma$ , if  $\rho \in \mathcal{T}_+(\sigma)$ . As these two preorders are dual, meaning that  $\rho \succsim_+ \sigma$  is equivalent to  $\sigma \succsim_- \rho$ , without loss of generality we will focus only on the preorder  $\succsim_-$  oriented along the thermodynamic arrow of time and simply denote it by  $\succsim$ . In other words, instead of writing  $\rho \in \mathcal{T}_-(\sigma)$  we can write  $\rho \succsim \sigma$ , meaning that  $\rho$  precedes  $\sigma$  in the thermodynamic preorder.

For two states  $\rho$  and  $\sigma$  it may happen that  $\rho \succsim \sigma$  and  $\sigma \succsim \rho$ . We then say that  $\rho$  and  $\sigma$  are *reversibly interconvertible* under GP operations and denote it by  $\rho \sim \sigma$ . Since the relation  $\sim$  is reflexive, transitive and symmetric, it is an equivalence relation and thus the set of states that are reversibly interconvertible forms a *thermodynamic equivalence class*. Using thermodynamic equivalence classes we can simplify thermodynamic preorders in the following way. Instead of considering the ordering between all states, we can restrict our study to the ordering between single representatives of each thermodynamic equivalence class. This results in promoting the thermodynamic preorder  $\succsim$  into a partial order  $\succ$  (not to be confused with majorisation, which is just a particular example of a partial

order), which has an additional property of being antisymmetric, i.e., if  $\rho \succ \sigma$  and  $\sigma \succ \rho$  then  $\rho = \sigma$ . Let us explain this construction with the use of the following example:

**Example 4.1: Thermodynamic equivalence class**

Consider a qubit system described by a Hamiltonian  $H = \sigma_z$  prepared in a general state with Bloch vector  $\mathbf{r} = (r \cos \phi, r \sin \phi, z)$ . Since a unitary  $U(t) = e^{-iHt}$ , mapping  $\mathbf{r}$  to  $\mathbf{r}' = (r \cos \phi', r \sin \phi', z)$ , is a reversible GP map, states with fixed  $r$  and  $z$  belong to the same thermodynamic equivalence class. Hence, we can focus only on one representative of each class parametrised by  $r$  and  $z$ , e.g., with  $\phi = 0$  (see

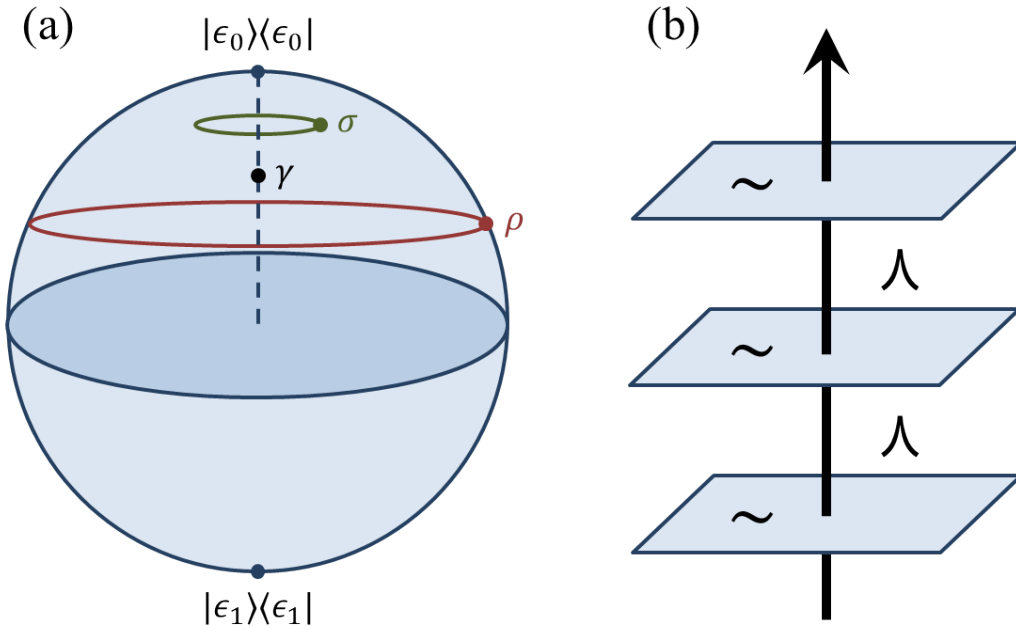


Figure 4.1: *Decomposition of preorder into partial order between equivalence classes.* (a) States on the unitary orbit generated by  $U(t) = e^{-iHt}$  are all reversibly interconvertible under GP transformations, as  $U(t)$  is GP. Hence, if a state  $\rho$  can thermodynamically evolve to  $\sigma$  then any state  $U(t)\rho U^\dagger(t)$  can evolve to any state  $U(t')\sigma U^\dagger(t')$ . Thus, without loss of generality one can only study the ordering between single representatives of each orbit. (b) States that are reversibly interconvertible by GP operations can be thought of as living in a plane “perpendicular” to the thermodynamic arrow of time, i.e., members of a thermodynamic equivalence class are at the same stage of the evolution towards equilibrium. The thermodynamic ordering of states is then given by the partial order between planes.

Fig. 4.1).

In fact we have just used a standard procedure that allows one to decompose any pre-order  $\succsim$  on a generic set  $S$  into a partial order  $\succ$  between equivalence classes (subsets of  $S$  with elements connected via an equivalence relation  $\sim$ ). Let us summarise this section by stating a formal definition of thermodynamic ordering:

**Definition 4.2: Thermodynamic ordering**

Consider a preorder  $\succsim$  on the set  $\mathcal{S}_d$  of quantum states defined by a relation of belonging to the past thermal cone, i.e.,  $\rho \succsim \sigma$  if and only if  $\rho \in \mathcal{T}_-(\sigma)$ . Identify each set of states that are reversibly interconvertible via a GP map, i.e.,  $\rho \succsim \sigma$  and  $\sigma \succsim \rho$ , with a thermodynamic equivalence class  $\sim$ . Thermodynamic ordering is a partial order  $\succ$  between those equivalence classes, i.e., a partial order on the quotient set  $\mathcal{S}_d / \sim$ .

### 4.1.2 Lattice

We will now provide a definition and interpret a special kind of partial order: a lattice. However, before we can do it we first need to introduce a few more notions, which are illustrated in Fig. 4.2. For any two states  $\rho$  and  $\sigma$  let us introduce a set of states  $\mathcal{T}_-(\rho, \sigma) = \mathcal{T}_-(\rho) \cap \mathcal{T}_-(\sigma)$ , i.e., the set of all states whose future thermal cones contain both  $\rho$  and  $\sigma$ . The thermodynamic interpretation of  $\mathcal{T}_-(\rho, \sigma)$  is that of a set of states in the past that are allowed by the thermodynamic arrow of time to evolve both into  $\rho$  and  $\sigma$  at present. Similarly, let us introduce a set of states  $\mathcal{T}_+(\rho, \sigma) = \mathcal{T}_+(\rho) \cap \mathcal{T}_+(\sigma)$ , i.e., the set of all states whose past thermal cones contain both  $\rho$  and  $\sigma$ . Thermodynamically  $\mathcal{T}_+(\rho, \sigma)$  is the set of states in the future that are allowed by the thermodynamic arrow of time to be reached from both  $\rho$  and  $\sigma$  at present.

Now, if there exists  $\tau_- \in \mathcal{T}_-(\rho, \sigma)$  such that for all  $\tau \in \mathcal{T}_-(\rho, \sigma)$  we have  $\tau \succ \tau_-$  then  $\tau_-$  is called the *join* of  $\rho$  and  $\sigma$  and is usually denoted by  $\rho \vee \sigma$ . The notation is justified by the fact that  $\mathcal{T}_+(\tau_-)$  is the smallest thermal cone that contains  $\mathcal{T}_+(\rho) \cup \mathcal{T}_+(\sigma)$ . Thermodynamically we can interpret the join of  $\rho$  and  $\sigma$  as the unique state in the past that is consistent both with  $\rho$  and  $\sigma$  at present, as well as with all possible joint pasts of  $\rho$  and  $\sigma$ . The join can also be seen as the extremal moment in the past evolution, at

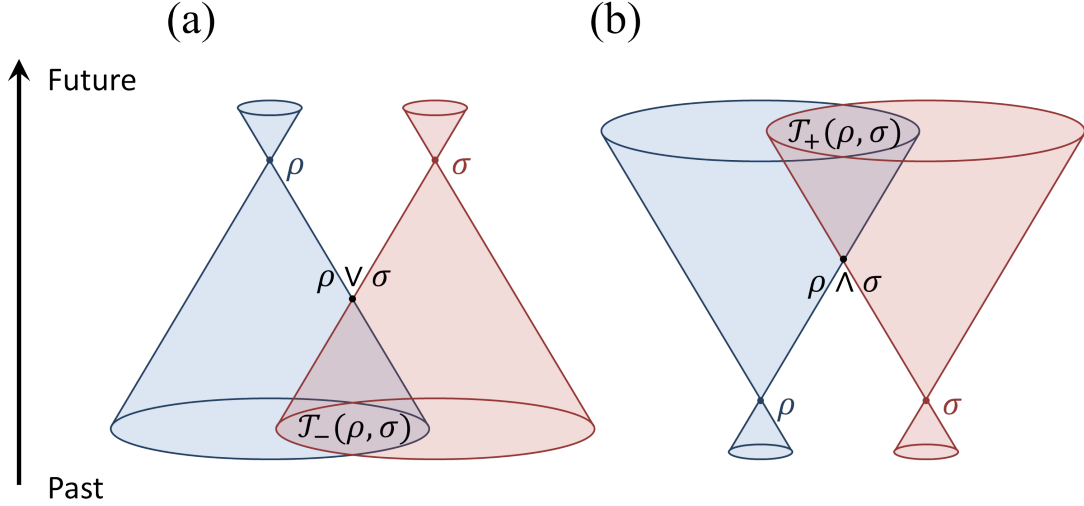


Figure 4.2: *Visualising the join and meet.* (a) The intersection of past thermal cones of  $\rho$  and  $\sigma$ , denoted by  $\mathcal{T}_-(\rho, \sigma)$ , is a set of states that can thermodynamically evolve to both  $\rho$  and  $\sigma$ . The join  $\rho \vee \sigma$  is the unique state belonging to  $\mathcal{T}_-(\rho, \sigma)$  that can be thermodynamically reached from all states in  $\mathcal{T}_-(\rho, \sigma)$ . (b) The intersection of future thermal cones of  $\rho$  and  $\sigma$ , denoted by  $\mathcal{T}_+(\rho, \sigma)$ , is a set of states that can be thermodynamically reached from both  $\rho$  and  $\sigma$ . The meet  $\rho \wedge \sigma$  is the unique state belonging to  $\mathcal{T}_+(\rho, \sigma)$  that can thermodynamically evolve to all states in  $\mathcal{T}_+(\rho, \sigma)$ .

which the system has to “decide” whether to evolve into  $\rho$  or  $\sigma$ .

Similarly, if there exists  $\tau_+ \in \mathcal{T}_+(\rho, \sigma)$  such that for all  $\tau \in \mathcal{T}_+(\rho, \sigma)$  we have  $\tau_+ \succ \tau$  then  $\tau_+$  is called the *meet* of  $\rho$  and  $\sigma$  and is usually denoted by  $\rho \wedge \sigma$ . Again, the notation is justified by the fact that  $\mathcal{T}_+(\tau_+)$  is the biggest thermal cone that is contained in  $\mathcal{T}_+(\rho) \cap \mathcal{T}_+(\sigma)$ . Thermodynamically the meet of  $\rho$  and  $\sigma$  is the unique state in the future that is consistent both with  $\rho$  and  $\sigma$  at present, as well as with all possible joint futures of  $\rho$  and  $\sigma$ . The meet can also be seen as the extremal moment in the future evolution, after which the system “forgets” whether it evolved from  $\rho$  or  $\sigma$ , as its state is consistent with both pasts.

#### Definition 4.3: Thermodynamic lattice

The thermodynamically ordered set of quantum states  $(\mathcal{S}_d, \succ)$  forms a thermodynamic lattice if for every pair of states  $\rho, \sigma \in \mathcal{S}_d$  there exists a join and meet.

The existence of a thermodynamic lattice would not only bring a new understanding of

the thermodynamic arrow of time (with a unique consistent future and past for each subset of states), but could also allow us to use new algebraic tools to study thermodynamics. Namely, if  $(\mathcal{S}_d, \succ)$  forms a thermodynamic lattice, then it can be fully described as an algebraic structure  $(\mathcal{S}_d, \vee, \wedge)$  consisting of a set of quantum states  $\mathcal{S}_d$  and two binary operations  $\vee$  and  $\wedge$  satisfying the following axioms for all  $\rho, \sigma, \tau \in \mathcal{S}_d$ :

$$\rho \vee \sigma = \sigma \vee \rho, \quad (4.1a)$$

$$\rho \vee (\sigma \vee \tau) = (\rho \vee \sigma) \vee \tau \quad (4.1b)$$

$$\rho \vee (\rho \wedge \sigma) = \rho, \quad (4.1c)$$

and another three obtained from the above by exchanging  $\vee$  with  $\wedge$ . In Fig. 4.3 we present examples of partial orders that do and do not form a lattice.

## 4.2 Ordering of classical states

We are now ready to analyse the structure of the thermodynamic ordering of states. We will start with classical states described by probability distributions  $\mathbf{p} \in \mathcal{P}_d$ , considering separately the infinite temperature limit,  $\beta \rightarrow 0$ , and the case of finite temperatures. The former can also be thought of as an information-theoretic limit, because with energy states all being degenerate, negentropy (or information) is the only thermodynamic resource. In this case, as we will show, the thermodynamic arrow of time exhibits the structure of a lattice (known as the information lattice [83]). However, we will also prove that as soon as different energy states become distinct, the lattice structure is broken within the classical theory.

### 4.2.1 Infinite temperature and a lattice structure

As discussed in Sec. 1.2.2 of Chapter 1 in the infinite temperature limit a classical state  $\mathbf{p}$  can be transformed into  $\mathbf{q}$  if and only if  $\mathbf{p}$  majorises  $\mathbf{q}$  (recall Definition M3). Within the space of  $d$ -dimensional probability distributions  $\mathcal{P}_d$ , majorisation forms a preorder, not a partial order, because for two probability distributions  $\mathbf{p}$  and  $\mathbf{q}$  that are connected by a permutation,  $\mathbf{q} = \Pi\mathbf{p}$ , we have  $\mathbf{p}$  majorising  $\mathbf{q}$  and  $\mathbf{q}$  majorising  $\mathbf{p}$ , but  $\mathbf{p} \neq \mathbf{q}$ . However, if we identify all probability distributions that are connected by some permutation (which is a reversible bistochastic map) with an equivalence class, we arrive at a partial order

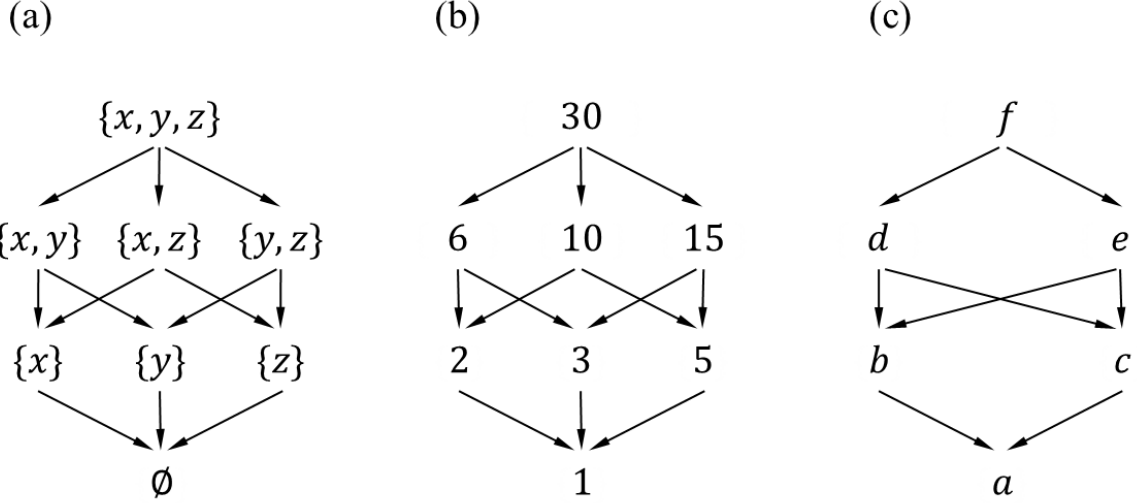


Figure 4.3: *Examples of lattice and non-lattice partial orders.* Partially ordered sets can be represented by their Hasse diagrams: an arrow from  $a$  to  $b$  denotes  $a \succ b$ . (a) The power set of any set  $A$  forms a lattice under the partial order induced by subset inclusion. The join and meet are given by set union and intersection, respectively. Here we choose a three-element set  $A = \{x, y, z\}$ . (b) Natural numbers partially ordered by divisibility, i.e.,  $a \succ b$  if and only if  $b$  is a divisor of  $a$ , form a lattice. The join and meet are given by the operations of taking the least common multiple and greatest common divisor, respectively. Here we present a sublattice of divisors of 30. (c) A set  $\{a, b, c, d, e, f\}$  with partial ordered defined by the presented Hasse diagram does not form a lattice. Although  $b$  and  $c$  have common upper bounds  $d, e$ , and  $f$ , neither of them is a join, i.e, the least upper bound.

between those equivalence classes. It is known that majorisation partial order forms a lattice: for any two probability distributions,  $\mathbf{p}$  and  $\mathbf{q}$ , there exists meet  $\mathbf{p} \wedge \mathbf{q}$  and join  $\mathbf{p} \vee \mathbf{q}$ . Here we will provide their construction as described by the authors of Ref. [83].

The meet is given by a probability vector  $\mathbf{l}$  with components given by

$$l_i = \min \left\{ \sum_{j=1}^i p_j^\downarrow, \sum_{j=1}^i q_j^\downarrow \right\} - \min \left\{ \sum_{j=1}^{i-1} p_j^\downarrow, \sum_{j=1}^{i-1} q_j^\downarrow \right\}. \quad (4.2)$$

To explain why this is the case, let us use thermo-majorisation curves in the limit of infinite temperatures (recall Definition M7), that we will simply refer to as majorisation curves. Such curves,  $f_{\mathbf{p}}$  and  $f_{\mathbf{q}}$ , consist of linear segments connecting points  $(i, \sum_{j=1}^i p_j^\downarrow)$  for  $i \in \{0, \dots, d\}$ , and analogously for  $\mathbf{q}$ . Now, the condition  $\mathbf{p} \succ \mathbf{q}$  is equivalent to  $f_{\mathbf{p}} \geq f_{\mathbf{q}}$  everywhere. Hence, for the meet,  $\mathbf{l} = \mathbf{p} \wedge \mathbf{q}$ , its majorisation curve  $f_{\mathbf{l}}$  must be the

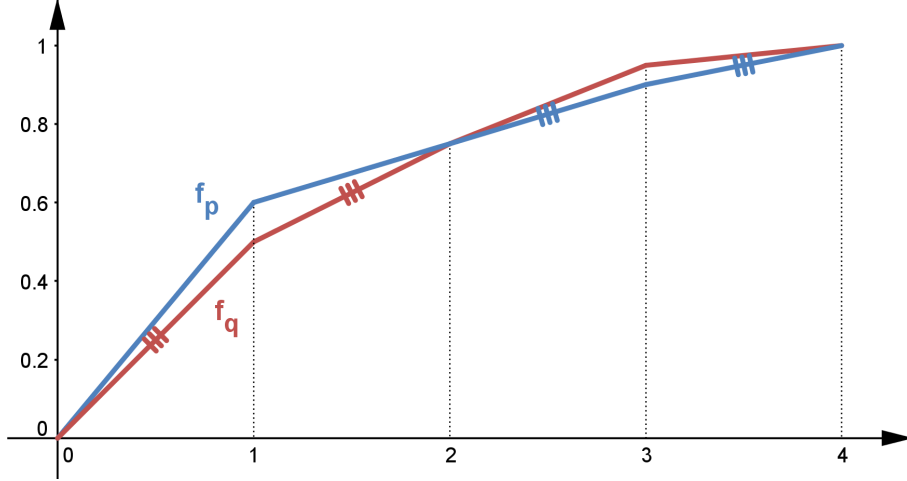


Figure 4.4: *Meet of a majorisation lattice.* Majorisation curves  $f_p$  and  $f_q$  (solid lines) for  $\mathbf{p} = \frac{1}{100}(60, 15, 15, 10)$  and  $\mathbf{q} = \frac{1}{100}(50, 25, 20, 5)$ , together with their meet  $\mathbf{p} \wedge \mathbf{q} = \frac{1}{100}(50, 25, 15, 5)$  (given by the line consisting of triple-crossed segments).

“maximal” curve lying below both  $f_p$  and  $f_q$ . The expression in Eq. (4.2) ensures that, as the resulting majorisation curve  $f_l$  is equal to  $\min\{f_p, f_q\}$  at each point. We illustrate this for exemplary probability vectors in Fig. 4.4.

The join  $\mathbf{p} \vee \mathbf{q}$  can be constructed with the use of the following algorithm. First, define a probability vector  $\mathbf{g}^{(0)}$  with components given by

$$g_i^{(0)} = \max \left\{ \sum_{j=1}^i p_j^\downarrow, \sum_{j=1}^i q_j^\downarrow \right\} - \max \left\{ \sum_{j=1}^{i-1} p_j^\downarrow, \sum_{j=1}^{i-1} q_j^\downarrow \right\}. \quad (4.3)$$

Now, the iterative application of the following transformation on  $\mathbf{g}^{(k)}$  results in  $\mathbf{p} \vee \mathbf{q}$  in no more than  $d - 1$  steps. Start with  $k = 0$ . Denote by  $n \geq 2$  the smallest integer such that  $g_n^{(k)} > g_{n-1}^{(k)}$  and by  $m \leq n - 1$  the greatest integer such that

$$g_{m-1}^{(k)} \geq \frac{\sum_{i=m}^n g_i^{(k)}}{n - m + 1} := a_k. \quad (4.4)$$

Define  $\mathbf{g}^{(k+1)}$  by setting its components  $g_i^{(k+1)} = a_k$  for  $i \in \{m \dots n\}$  and  $g_i^{(k+1)} = g_i^{(k)}$  otherwise. Repeat until for some  $k'$  the vector  $\mathbf{g}^{(k')}$  has components ordered in a decreasing order. The join  $\mathbf{p} \vee \mathbf{q}$  is then given by  $\mathbf{g}^{(k')}$ . The procedure just described starts similarly to the one used to define meet: we introduce a state  $\mathbf{g}^{(0)}$ , whose majorisation curve is the “minimal” curve lying above both  $f_p$  and  $f_q$ , i.e.,  $f_{\mathbf{g}^{(0)}}$  is equal to  $\max\{f_p, f_q\}$  at each point. The problem is that the resulting curve may not be concave, and since the majorisation curve is constructed from the components arranged in a decreasing order, each such curve must be concave. What the described algorithm does to overcome

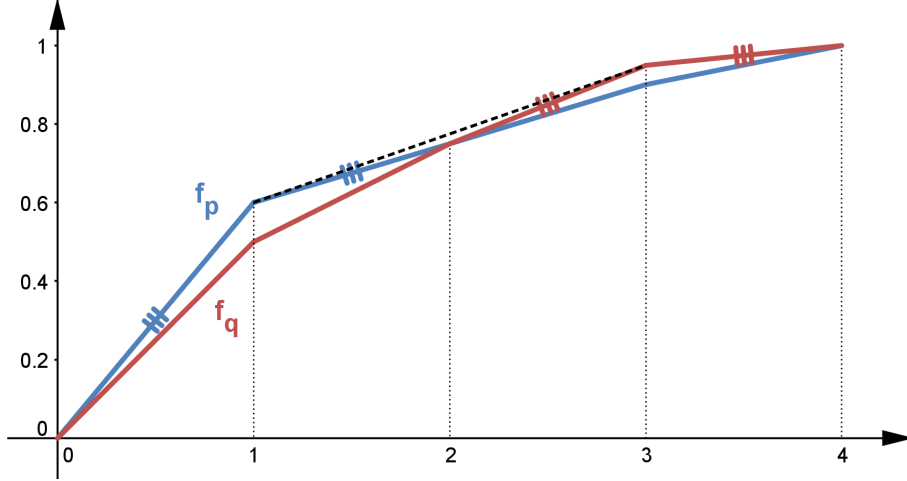


Figure 4.5: *Join of a majorisation lattice.* Majorisation curves  $f_{\mathbf{p}}$  and  $f_{\mathbf{q}}$  (solid lines) for  $\mathbf{p} = \frac{1}{100}(60, 15, 15, 10)$  and  $\mathbf{q} = \frac{1}{100}(50, 25, 20, 5)$ , together with the curve corresponding to  $\mathbf{g}^{(0)} = \frac{1}{100}(60, 15, 20, 5)$  from Eq. (4.3) (given by the line consisting of triple-crossed segments). This curve is not concave, since  $\mathbf{g}^{(0)}$  is not arranged in the decreasing order. Hence, in order to obtain the join  $\mathbf{p} \vee \mathbf{q}$ , one needs to “smooth” the curve between the points  $i = 1$  and  $i = 3$  (dashed black line).

this problem, is to identify points at which the curve breaks concavity, and “smooth” it over sufficient number of points to guarantee concavity. We illustrate this for exemplary probability vectors in Fig. 4.5.

### 4.2.2 Finite temperatures and a broken lattice structure

We will now first show that in the simplest case of a two-level system the partial order induced by thermo-majorisation on the full state space  $\mathcal{P}_2$  does not form a lattice. Then we will present how to generalise this result for  $d > 2$ . However, we will also prove that a lattice structure is preserved within subspaces of  $\mathcal{P}_d$  containing probability vectors belonging to the same  $\beta$ -ordering (recall Definition M6). In other words, for two states  $\mathbf{p}$  and  $\mathbf{q}$  belonging to the same  $\beta$ -ordering, there may be many incomparable “candidate” states for meet and join; but within the subset consisting only of probability vectors with the same  $\beta$ -ordering as  $\mathbf{p}$  and  $\mathbf{q}$ , the meet and join will be defined uniquely.

For any given two-dimensional thermal state  $\boldsymbol{\gamma} = (\gamma_0, 1 - \gamma_0)$  with  $\gamma_0 \neq 1/2$ , let us choose two states,  $\mathbf{p} = (p, 1 - p)$  and  $\mathbf{q} = (q, 1 - q)$ , with

$$p = \frac{1 + \gamma_0}{2} \geq \gamma_0, \quad q = \frac{2\gamma_0 - 1}{\gamma_0} \leq \gamma_0. \quad (4.5)$$

We will prove that there does not exist a join for these two states. A generic two level system is described by  $\mathbf{r} = (r, 1 - r)$ . It is easy to verify that for  $r \geq \gamma_0$  the only state that thermo-majorises both  $\mathbf{p}$  and  $\mathbf{q}$  is given by  $r = 1$ . On the other hand, for  $r \leq \gamma_0$  we get that  $\mathbf{r}$  thermo-majorises both  $\mathbf{p}$  and  $\mathbf{q}$  if  $r \leq \gamma_0/2$ . Among these states the ones with  $r < \gamma_0/2$  thermo-majorise the one with  $r = \gamma_0/2$ . Hence, we are left only with two candidates for the join of  $\mathbf{p}$  and  $\mathbf{q}$ , namely  $(1, 0)$  and  $(\gamma_0/2, 1 - \gamma_0/2)$ . By direct inspection we find that those states are incomparable under thermo-majorisation partial order, so no join exists and we do not have a lattice structure. In a similar fashion one can prove that there always exists states  $\mathbf{p}$  and  $\mathbf{q}$  for which no meet exists [for example by choosing  $p = (3 + \gamma_0)/4$  and  $q = (\gamma_0^2 + 2\gamma_0 - 1)/(4\gamma_0)$ ].

Let us now consider two  $d$ -dimensional probability distributions  $\mathbf{p} = (0, \dots, 0, p, 1 - p)$  and  $\mathbf{q} = (0, \dots, 0, q, 1 - q)$ , i.e., classical states with only the two highest energy levels  $\epsilon_{d-1}$  and  $\epsilon_d$  occupied. For simplicity let us assume that these two energy levels are non-degenerate. This simplifying assumption allows us to map a  $d$ -dimensional problem to a two-dimensional one, showing that thermo-majorisation does not have the structure of a lattice for  $d > 2$ . Let us now choose  $p \geq \frac{\gamma_{d-1}}{\gamma_d + \gamma_{d-1}}$  and  $q \leq \frac{\gamma_{d-1}}{\gamma_d + \gamma_{d-1}}$ . This way the corresponding thermo-majorisation curves  $f_{\mathbf{p}}$  and  $f_{\mathbf{q}}$  (see Fig. 4.6 and Definition M7) will consist of linear segments joining the following points:

$$f_{\mathbf{p}} : \{(0, 0), (\gamma_{d-1}, p), (\gamma_d + \gamma_{d-1}, 1)\}, \quad (4.6a)$$

$$f_{\mathbf{q}} : \{(0, 0), (\gamma_d, 1 - q), (\gamma_d + \gamma_{d-1}, 1)\}. \quad (4.6b)$$

We will now try again to find a candidate state for the join of  $\mathbf{p}$  and  $\mathbf{q}$ . We first note that any state  $\mathbf{r}$  that has non-zero occupation in at least two levels that are not the highest energy levels, let say  $i$  and  $j$ , does not thermo-majorise either  $\mathbf{p}$  or  $\mathbf{q}$ . This is because the thermo-majorisation curve of such a state can only reach 1 at  $\gamma_i + \gamma_j$ , which is bigger than  $\gamma_d + \gamma_{d-1}$ , and hence  $f_{\mathbf{r}}$  will not lie above either  $f_{\mathbf{p}}$  or  $f_{\mathbf{q}}$ . Let us now focus on the states  $\mathbf{r}^{(i)}$  that have a single non-zero entry at  $i$ -th position for  $i \in \{1, \dots, d - 2\}$ . The thermo-majorisation curve  $f_{\mathbf{r}^{(i)}}$  of such state at a point  $\gamma_{d-1}$  will take the value  $\gamma_{d-1}/\gamma_i < 1$ . Hence, the curve  $f_{\mathbf{r}^{(i)}}$  will not lie above  $f_{\mathbf{p}}$  as long as  $p > \max_i \gamma_{d-1}/\gamma_i = \gamma_{d-1}/\gamma_{d-2}$ . The choice

$$p = \frac{1 + \max\left(\frac{\gamma_{d-1}}{\gamma_{d-2}}, \frac{\gamma_{d-1}}{\gamma_d + \gamma_{d-1}}\right)}{2} \quad (4.7)$$

guarantees this, as well as the consistency with the initial assumption  $p \geq \frac{\gamma_{d-1}}{\gamma_d + \gamma_{d-1}}$ . Thus

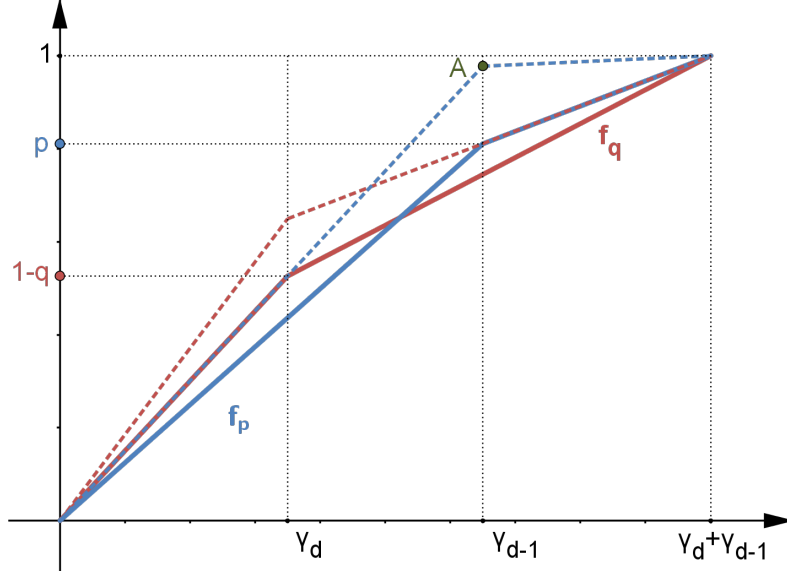


Figure 4.6: *Thermo-majorisation order is not a lattice.* Thermo-majorisation curves  $f_{\mathbf{p}}$  and  $f_{\mathbf{q}}$  (solid lines), together with the candidates for the join  $\mathbf{p} \vee \mathbf{q}$ , i.e., optimal curves (plotted with dashed lines) thermo-majorising both  $\mathbf{p}$  and  $\mathbf{q}$ . Satisfying the inequalities given by Eq. (4.8) guarantees that: 1.  $f_{\mathbf{q}} > f_{\mathbf{p}}$  at  $\gamma_d$ ; 2.  $f_{\mathbf{q}} < f_{\mathbf{p}}$  at  $\gamma_{d-1}$ ; 3. Point  $A$  lies below 1, which results in the existence of two incomparable candidates for  $\mathbf{p} \vee \mathbf{q}$ .

the only candidate states for the join of  $\mathbf{p}$  and  $\mathbf{q}$  are of the form  $\mathbf{r} = (0, 0, \dots, 0, r, 1 - r)$ . But this is exactly a two-level case discussed before and, using the same reasoning, one can show that for any  $\mathbf{p}$  and  $\mathbf{q}$ , consistent with our initial assumptions, the choice of  $q$  satisfying

$$\frac{\gamma_d}{\gamma_{d-1}}p < (1 - q) < \min \left( \frac{\gamma_d}{\gamma_{d-1}}, 1 - \frac{\gamma_{d-1}}{\gamma_d}(1 - p) \right), \quad (4.8)$$

guarantees that no join for  $\mathbf{p}$  and  $\mathbf{q}$  exists. The first inequality on the left guarantees that  $f_{\mathbf{q}}$  will be above  $f_{\mathbf{p}}$  at the point  $\gamma_d$  (refer to Fig. 4.6). For  $(1 - q)$  being smaller than the second argument of the minimum guarantees that  $f_{\mathbf{q}}$  will be below  $f_{\mathbf{p}}$  at the point  $\gamma_{d-1}$ . Hence,  $\mathbf{p}$  and  $\mathbf{q}$  are incomparable. Finally, ensuring  $(1 - q)$  to be smaller than the first argument of the minimum guarantees that there exist exactly two candidates for a join of  $\mathbf{p}$  and  $\mathbf{q}$  and that these are incomparable.

A careful reader would have noticed that in our two-dimensional example we have  $\gamma_0 = 1/2$  not only for infinite temperatures, but also at finite temperatures if the two energy levels are degenerate. More generally, the transformations within any degenerate subspace are governed by the same rules as the infinite temperature limit from the previous

section.<sup>2</sup> Hence, the lattice structure arises within subspaces of states whose energies cannot be distinguished. But degenerate energy subspaces are not the only ones in which the lattice structure can be preserved, as there exist subspaces of  $\mathcal{P}_d$  in which thermo-majorisation is effectively described by majorisation. These consist of probability vectors that belong to the same  $\beta$ -ordering, i.e., the same permutation matrix rearranges their Gibbs-rescaled versions into a decreasing order (recall Definition M6). To see this, consider two probability vectors  $\mathbf{p}$  and  $\mathbf{q}$  with the same  $\beta$ -ordering that we will denote by  $\beta_1$ . Let us also denote by  $\gamma^{\beta_1}$  the  $\beta_1$ -ordered version of the thermal Gibbs state  $\gamma$ . Now, the extremal points of the segments that constitute thermo-majorisation curves  $f_{\mathbf{p}}$  and  $f_{\mathbf{q}}$  will have the same  $x$  coordinates:  $x_i = \sum_{j=1}^i \gamma_j^{\beta_1}$ . Hence, to verify if one curve lies above the other, one only needs to compare their  $y$  coordinates. This means that  $\mathbf{p}$  thermo-majorises  $\mathbf{q}$  if and only if  $\mathbf{p}^{\beta_1} \succ \mathbf{q}^{\beta_1}$ , where  $\succ$  denotes standard majorisation. This allows us to use a slightly modified version of the construction presented in Sec. 4.2.1 to find the meet and join. Meet is given by Eq. (4.2) simply by replacing the entries of  $\mathbf{p}$  arranged in a decreasing order with the  $\beta_1$ -ordered entries (and similarly for  $\mathbf{q}$ ). To verify that the resulting probability vector is  $\beta_1$ -ordered, note that its thermo-majorisation curve is concave. To find the join we also replace decreasing order with  $\beta_1$ -ordering in Eq. (4.3), and modify the described iterative procedure in the following way. We define  $n$  as the smallest integer such that  $g_n^{(k)}/\gamma_n^{\beta_1} > g_{n-1}^{(k)}/\gamma_{n-1}^{\beta_1}$ , and by  $m$  the greatest integer satisfying

$$\frac{g_{m-1}^{(k)}}{\gamma_{m-1}^{\beta_1}} \geq \frac{\sum_{i=m}^n g_i^{(k)}}{\sum_{i=m}^n \gamma_i^{\beta_1}} := b_k. \quad (4.9)$$

Finally, we define  $\mathbf{g}^{(k+1)}$  by setting its components  $g_i^{(k+1)} = b_k \gamma_i^{\beta_1}$  for  $i \in \{m \dots n\}$  and  $g_i^{(k+1)} = g_i^{(k)}$  otherwise. The role of this modified procedure is the same as in Sec. 4.2.1: to ensure that the thermo-majorisation curve of the resulting join  $\mathbf{p} \vee \mathbf{q}$  is concave, which also guarantees that  $\mathbf{p} \vee \mathbf{q}$  is  $\beta_1$ -ordered.

As a final remark, let us notice that in the infinite temperature limit states connected by a permutation are reversibly interconvertible, and so they belong to the same thermal equivalence class. Hence, when we speak of a join  $\mathbf{r} = \mathbf{p} \vee \mathbf{q}$ , it is unique only because states  $\mathbf{r}$  and  $\Pi \mathbf{r}$ , where  $\Pi$  denotes arbitrary permutation, are equivalent. At finite temperature, however, this permutation invariance is broken and so is the uniqueness of join and

---

<sup>2</sup>Also, an important point to note is that if some energy states are degenerate, then thermo-majorisation actually forms a preorder. It can be replaced with a partial order only once we identify all states connected via a permutation between degenerate states with corresponding equivalence classes.

meet. It is preserved only if we restrict our considerations to a particular class of states described by the same  $\beta$ -ordering.

### 4.3 Ordering of quantum states

Let us now proceed to analysing the structure of the thermodynamic ordering of quantum states. In the infinite temperature limit we will show that, similarly to the classical case, we are dealing with a lattice structure. This could be expected as in this limit the unitary operations are GP and, since unitaries are reversible, for every state with coherence there exists a diagonal (classical) state belonging to the same thermodynamic equivalence class. Therefore, the quantum and classical states share the same structure of thermodynamic ordering.

The situation becomes much more complicated at finite temperatures. In fact, the set of states that a given state  $\rho \in \mathcal{S}_d$  can be mapped to via a GP map has not been explicitly found for any dimension  $d$ . Therefore, we solve this problem in the simplest case of  $d = 2$  and provide future thermal cones  $\mathcal{T}_+(\rho)$  for all states  $\rho \in \mathcal{S}_d$ . This will allow us to prove that the thermodynamic arrow of time for qubit systems exhibits a lattice structure. This contrast with a classical two-level system provides evidence that coherence may play an important role in thermodynamics by adding structure to the thermodynamic ordering of states. However, whether the lattice structure persists beyond the qubit case for  $d \geq 3$  remains an open question for future investigation.

#### 4.3.1 Infinite temperature and a lattice structure

In the infinite temperature limit the Gibbs state is the maximally mixed state  $\gamma = \mathbb{I}/d$ . Hence, GP maps are replaced by unital maps  $\mathcal{E}(\mathbb{I}) = \mathbb{I}$ . The existence of a unital map  $\mathcal{E}$  satisfying  $\mathcal{E}(\rho) = \sigma$  is equivalent to the spectrum of  $\rho$  majorising the spectrum of  $\sigma$ . This is due to the fact that the set of unital maps contains all unitaries, so that  $\rho$  and  $\sigma$  can be brought to a diagonal form in the same basis, and then the transformation between two diagonal states via a unital CPTP map is described by a bistochastic matrix. Thus, the problem can be mapped to the one discussed in the previous section. The slight difference only lies in thermodynamic equivalence classes. Namely, in the classical case these were composed of probability vectors connected via a permutation, whereas in the quantum

case these are composed of density matrices connected via a unitary. Regardless of this difference, the partial order between density matrices modulo unitaries forms a lattice.

### 4.3.2 Finite temperatures: qubit evidence for a lattice structure

We will consider a generic qubit system described by a Hamiltonian  $H = \epsilon_1 |\epsilon_1\rangle\langle\epsilon_1|$ . Let us denote a thermal state of such a system with respect to inverse temperature  $\beta$  by  $\gamma = e^{-\beta H}/Z$ , where as usual  $Z = \text{Tr}(e^{-\beta H})$ . In what follows we will describe qubit states  $\rho$ ,  $\rho'$  and a thermal state  $\gamma$  using the Bloch sphere representation,

$$\rho = \frac{\mathbb{I} + \mathbf{r}_\rho \cdot \boldsymbol{\sigma}}{2}, \quad (4.10)$$

where  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  denotes the vector of Pauli matrices. The Bloch vectors will be parametrised in the following way:

$$\mathbf{r}_\rho = (x, y, z), \quad \mathbf{r}_{\rho'} = (x', y', z'), \quad \mathbf{r}_\gamma = (0, 0, \zeta), \quad (4.11)$$

where  $\zeta = 2Z^{-1} - 1 \geq 0$ .

The following theorem, which may be of independent interest (as it provides geometric characterisation of the family of qubit CPTP maps with a given fixed point), specifies the necessary and sufficient conditions for the existence of a GP quantum channel between generic qubit states  $\rho$  and  $\rho'$ :

#### Theorem 4.1: Existence of a GP transformation between qubit states

Given two qubit states  $\rho$  and  $\rho'$  there exists a GP quantum channel  $\mathcal{E}$  such that  $\mathcal{E}(\rho) = \rho'$  if and only if  $R_\pm(\rho) \geq R_\pm(\rho')$  where  $R_\pm(\rho) = \delta(\rho) \pm \zeta z$  and

$$\delta(\rho) := \sqrt{(z - \zeta)^2 + (x^2 + y^2)(1 - \zeta^2)}, \quad (4.12)$$

with analogous (primed) definitions for  $\rho'$ .

The future thermal cone  $\mathcal{T}_+(\rho)$  of any qubit state  $\rho$  can be found as a corollary of the above theorem:

#### Corollary 4.2: Future thermal cone of a qubit system

Consider a generic qubit state  $\rho$  and orient the Bloch sphere so that its  $xz$  plane coincides with the plane containing  $\rho$  and a thermal state  $\gamma$ , i.e.,  $\mathbf{r}_\rho = (x, 0, z)$ .

Define two disks,  $D_1(\rho)$  and  $D_2(\rho)$  with corresponding circles  $C_1(\rho)$  and  $C_2(\rho)$ , of radii

$$R_1(\rho) = \frac{R_-(\rho) + \zeta^2}{1 - \zeta^2}, \quad R_2(\rho) = \frac{R_+(\rho) - \zeta^2}{1 - \zeta^2}, \quad (4.13)$$

centred at

$$\mathbf{z}_1(\rho) = [0, 0, \zeta(1 + R_1(\rho))], \quad \mathbf{z}_2(\rho) = [0, 0, \zeta(1 - R_2(\rho))]. \quad (4.14)$$

Then the set of qubit states that a state  $\rho$  can be mapped to by a GP quantum channel is given, in the Bloch sphere, by the region obtained from revolving the intersection  $D_1(\rho) \cap D_2(\rho)$  around the  $z$  axis. In other words, it is given by the intersection of two balls of radii  $R_1(\rho)$  and  $R_2(\rho)$  centred at  $\mathbf{z}_1(\rho)$  and  $\mathbf{z}_2(\rho)$ .

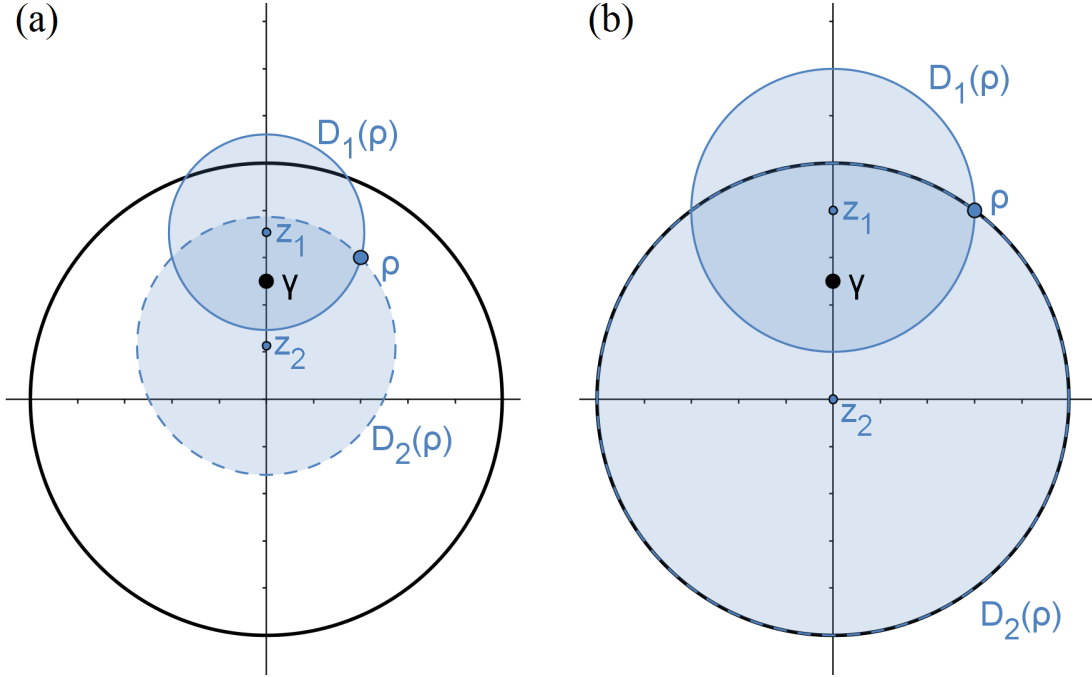


Figure 4.7: *Future thermal cones for qubits (GP operations)*. A general qubit state  $\rho$  and a thermal state  $\gamma$  with  $\mathbf{r}_\gamma = (0, 0, 0.5)$  presented in the Bloch sphere. The disk  $D_1(\rho)$  corresponds to a set of states  $\{\sigma\}$  with  $R_-(\sigma) \leq R_-(\rho)$ , whereas the disk  $D_2(\rho)$  correspond to a set  $\{\sigma\}$  with  $R_+(\sigma) \leq R_+(\rho)$ . The equalities are obtained at the edges of the disks, i.e., on circles  $C_1(\rho)$  and  $C_2(\rho)$ . The set of states  $\rho$  can be mapped to via a GP quantum channel is given by the intersection  $D_1(\rho) \cap D_2(\rho)$  (which can also be freely revolved around the  $z$  axis). (a) A mixed state with  $\mathbf{r}_\rho = (0.4, 0, 0.6)$ . (b) A pure state with  $\mathbf{r}_\rho = (0.6, 0, 0.8)$ .

The proof of the above results is based on the Alberti-Uhlmann theorem [84] and can be found in Appendix C. We illustrate the statement of Corollary 4.2 in Fig. 4.7. Let us also briefly discuss a few particular cases. For every pure state  $\rho$  we have  $x^2 + y^2 + z^2 = 1$ , which results in  $R_+(\rho) = 1$  and  $R_-(\rho) = 1 - 2z\zeta$ . Hence, pure qubit states are totally ordered by the value of  $z$ : the top state is described by a Bloch vector  $(0, 0, -1)$ , and the bottom one by  $(0, 0, 1)$ . For every incoherent state  $\rho$  one of the disks,  $D_1(\rho)$  or  $D_2(\rho)$ , is always contained within the other. Hence, for an incoherent state  $\rho$  with  $z \geq \zeta$  the future thermal cone is given by  $D_1(\rho)$ , whereas when  $z \leq \zeta$  it is given by  $D_2(\rho)$ . Finally, note that when  $\zeta = 0$ , i.e., we consider the infinite temperature limit, both disks are centred at the origin and have the same radius equal to the length of  $\mathbf{r}_\rho$ . We thus recover the majorisation result, as the spectrum of the state  $\rho$  majorises that of  $\rho'$  if and only if  $\mathbf{r}_\rho \succeq \mathbf{r}_{\rho'}$ .

We are now ready to state the anticipated result that the thermodynamic ordering of qubit states at finite temperatures, unlike the ordering of classical states of a two-level system, forms a lattice. Fig. 4.8 serves as an illustration of the following theorem:

### Theorem 4.3: Thermodynamic lattice for qubits

The thermodynamically ordered set of qubit states forms a lattice. The partially ordered thermodynamic equivalence classes consist of states connected via a unitary conjugation with  $U = \exp -iHt$ . For two distinct equivalence classes consider their representatives,  $\rho$  and  $\rho'$ , lying in the  $xz$  plane of the Bloch sphere with  $x \geq 0$ . The join and meet of  $\rho$  and  $\rho'$  are defined as follows. Introduce

$$\rho_m^{\max} = \arg \max\{R_m(\rho), R_m(\rho')\}, \quad (4.15a)$$

$$\rho_m^{\min} = \arg \min\{R_m(\rho), R_m(\rho')\}, \quad (4.15b)$$

for  $m \in \{1, 2\}$ . The join is then given by a state lying in the Bloch sphere at the intersection of two circles  $C_1(\rho_1^{\max})$  and  $C_2(\rho_2^{\max})$ , and the meet is defined analogously by replacing max with min.

*Proof.* First of all, we can restrict our considerations to states lying in the  $xz$  plane of the Bloch sphere with Bloch vectors  $(x, 0, z)$  and  $(-x, 0, z)$  being equivalent. Now, assume that two states,  $\rho$  and  $\rho'$ , are comparable,  $\rho \succ \rho'$ . Then, by Theorem 4.1 and Corollary 4.2, the two disks  $D_1(\rho')$  and  $D_2(\rho')$  are fully contained inside  $D_1(\rho)$  and  $D_2(\rho)$ . As a result

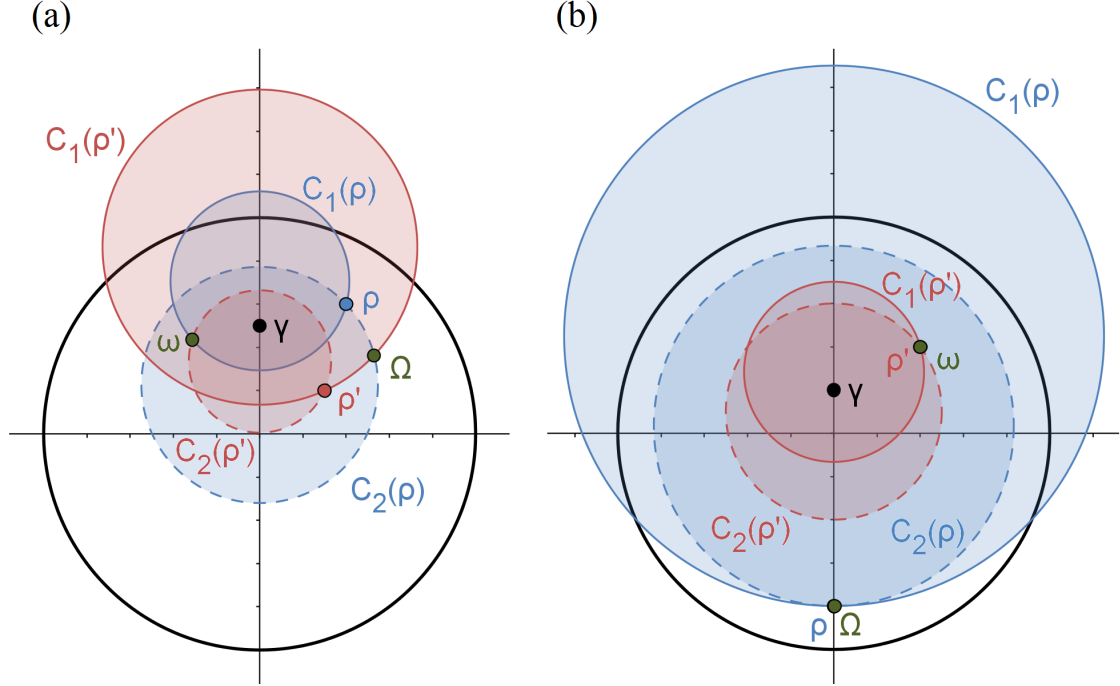


Figure 4.8: *Thermodynamic lattice for qubits.* A thermal state  $\gamma$  and two states  $\rho$  and  $\rho'$  presented in the Bloch sphere. (a) States described by  $\mathbf{r}_\rho = (0.4, 0, 0.6)$ ,  $\mathbf{r}_{\rho'} = (0.3, 0, 0.2)$ ,  $\mathbf{r}_\gamma = (0.0, 0, 0.5)$ . The join  $\Omega$  of  $\rho$  and  $\rho'$  lies at the intersection of  $C_1(\rho')$  and  $C_2(\rho)$ , whereas their meet  $\omega$  lies at the intersection of  $C_1(\rho)$  and  $C_2(\rho')$ . (b) States described by  $\mathbf{r}_\rho = (0, 0, -0.8)$ ,  $\mathbf{r}_{\rho'} = (0.4, 0, 0.4)$ ,  $\mathbf{r}_\gamma = (0.0, 0, 0.2)$ . The join  $\Omega$  of  $\rho$  and  $\rho'$  lies at the intersection of  $C_1(\rho)$  and  $C_2(\rho)$  (which coincides with  $\rho$ ), whereas their meet  $\omega$  lies at the intersection of  $C_1(\rho')$  and  $C_2(\rho')$  (which coincides with  $\rho'$ ).

$\rho_1^{\max} = \rho$ ,  $\rho_2^{\max} = \rho$ ,  $\rho_1^{\min} = \rho'$  and  $\rho_2^{\min} = \rho'$ . Hence, the join is given by  $\rho$  and the meet by  $\rho'$ , consistently with the fact that for every lattice if  $\rho \succ \rho'$  then  $\rho \vee \rho' = \rho$  and  $\rho \wedge \rho' = \rho'$ .

Now, let us consider the case when  $\rho$  and  $\rho'$  are incomparable, i.e., neither  $\rho \succ \rho'$  nor  $\rho' \succ \rho$ . Then without the loss of generality we have  $R_+(\rho) > R_+(\rho')$  and  $R_-(\rho) < R_-(\rho')$ . Consider a set states  $\mathcal{T}_-(\rho, \rho')$  whose future thermal cones contain both  $\rho$  and  $\rho'$ . According to Theorem 4.1,  $\tau \in \mathcal{T}_-(\rho, \rho')$  if and only if  $R_+(\tau) \geq R_+(\rho)$  and  $R_-(\tau) \geq R_-(\rho')$ . Now, if there existed a state  $\Omega$  such that  $R_+(\Omega) = R_+(\rho)$  and  $R_-(\Omega) = R_-(\rho')$  it would clearly be a join  $\rho \vee \rho'$ . This is because one could reach both  $\rho$  and  $\rho'$  from  $\Omega$  and also  $\Omega$  itself could be reached from any  $\tau \in \mathcal{T}_-(\rho, \rho')$ . We will now prove that such a state  $\Omega$  exists for any choice of incomparable states  $\rho$  and  $\rho'$ . The condition  $R_+(\Omega) = R_+(\rho)$  means that  $\Omega \in C_2(\rho)$ , whereas the condition  $R_-(\Omega) = R_-(\rho')$  means that  $\Omega \in C_1(\rho')$ . Hence, such a

state exists if and only if the circles  $C_2(\rho)$  and  $C_1(\rho')$  intersect (see Fig. 4.8). To prove this first note that a thermal state  $\gamma$  is contained inside both circles [straightforward from Eqs.(4.13)-(4.14)]. This means that either the circles intersect or one is fully contained inside the other. However, the latter is not possible, because a circle  $C_2(\rho)$  contains a point  $\rho$  that is inside  $C_1(\rho')$ ; and a circle  $C_1(\rho')$  contains a point  $\rho'$  that is inside  $C_2(\rho)$ . We thus conclude that the circles  $C_2(\rho)$  and  $C_1(\rho')$  do intersect and that a state  $\Omega$  lying at their intersection is a join  $\rho \vee \rho'$ .

Analogously, if a state  $\omega$  exists such that  $R_+(\omega) = R_+(\rho')$  and  $R_-(\omega) = R_-(\rho)$ , then it would be a meet  $\rho \wedge \rho'$ . The existence of such state is equivalent the circles  $C_1(\rho)$  and  $C_2(\rho')$  intersecting. Again, the state  $\gamma$  is contained in both circles, so that they either intersect or one is contained inside the other. The latter is impossible, because a circle  $C_1(\rho)$  contains a point  $\rho$  that is outside  $C_2(\rho')$ ; and a circle  $C_2(\rho')$  contains a point  $\rho'$  that is outside  $C_1(\rho)$ . Hence, the circles  $C_1(\rho)$  and  $C_2(\rho')$  do intersect and that a state  $\omega$  lying at their intersection is a meet  $\rho \wedge \rho'$ .

□

## Part II

# Uncertainty relations

# Summary of results

In Chapter 5 we study uncertainty relations for mixed states, when some of our inability to make perfect predictions is actually classical. We are interested in separating the uncertainty that arises from classical ignorance and the “genuinely quantum” uncertainty coming from coherence in a particular basis. In other words, we want to ask: how much of the uncertainty in predicting measurement outcomes for noncommuting quantum observables is *de facto* quantum mechanical? Within the framework of entropic uncertainty relations we provide a natural decomposition of the uncertainty (as measured by the Shannon entropy of the outcome statistics) into a classical component, and an intrinsically quantum mechanical component (which is also a proper measure of coherence). We show that the sum of quantum components for two noncommuting observables is never lower or upper bounded by any state-independent quantities. Instead, this total quantum uncertainty admits *fixed-entropy* lower bounds that generalise entropic formulations of uncertainty relations such as the one provided by Maassen and Uffink [34]. These relations reveal a non-trivial interplay between quantum and classical randomness in any finite-dimensional state. We further develop certain fixed-entropy bounds, compare them with other recently developed entropic bounds [38, 85, 86], and discuss their nonlinear character. Finally, we point out that the constraint on the von Neumann entropy of a state affects the structure of states minimising the total uncertainty, which leads us to the following chapter.

In Chapter 6 we analyse how classical lack of knowledge affects quantum states that minimise the unavoidable uncertainty arising from the noncommutativity of two observables. We first note that, due to the plethora of uncertainty measures, there are many minimum uncertainty states, each optimal with respect to a different measure. Instead of restricting our study to a particular measure, as in Chapter 5 where we focus on the Shannon entropy, we follow a “universal” approach introduced in Refs. [85, 87, 88]. We

thus present plausible axioms for the set  $\mathcal{F}$  of bona-fide information-theoretic uncertainty functions, and then discuss the existence of states minimising uncertainty with respect to all members of  $\mathcal{F}$ , i.e., *universal minimum uncertainty states* (universal MUSs). We prove that such states do not exist within the full state space, i.e., when there are no constraints on purity of the system under study. However, we present an explicit example of a universal MUS for a qubit that arises when purity is constrained by introducing a threshold amount of noise. For higher dimensional systems we derive several no-go results limiting the existence of noisy universal MUSs, but also conjecture that universality may emerge in an approximate sense. We conclude by discussing connections with thermodynamics, and highlight the privileged role that one particular non-equilibrium free energy functional plays close to equilibrium.

Finally, in Chapter 7, we focus on the difference between quantum and classical information reflected by the fact that the former can be destroyed solely by observation. Our investigations are thus concerned with error-disturbance trade-off relations, i.e., how the information gained about one observable disturbs the information about another observable that does not commute with the first one. More precisely, we want to clarify some confusing aspects of these trade-off relations for sequential measurements by questioning the commonly used state-dependent measures of error and disturbance. Inspired by the operational approach of quantum information theory, we argue for an *operational requirement* (OR) that all state-dependent measures of disturbance should satisfy. Motivated by this natural criterion, we prove that in any  $d$ -dimensional Hilbert space and for any pair of noncommuting operators,  $A$  and  $B$ , there exists a set of at least  $2^{d-1}$  *zero-noise, zero-disturbance states* (ZNZD states), for which the first observable can be measured without noise and the second will not be disturbed. Moreover, we show that it is possible to construct such ZNZD states for which the expectation value of the commutator  $[A, B]$  does not vanish. Therefore, any state-dependent error-disturbance relation, based on the expectation value of the commutator as a lower bound, must violate the OR. We also discuss Ozawa's state-dependent error-disturbance relation in light of our results and show that the disturbance measure used in this relation exhibits unphysical properties. We conclude that the trade-off is inevitable only between state-independent measures of error and disturbance.

# Chapter 5

## Quantum and classical sources of uncertainty

The Heisenberg-Robertson uncertainty relation [28] is indisputably the most celebrated constraint on our ability to make predictions about measurement outcomes, sometimes simply referred to as “the” uncertainty relation (UR). For measurement outcomes of two observables,  $A$  and  $B$ , it bounds the product of their variances in a given quantum state  $\rho$  in terms of the expectation of the commutator  $[A, B]$ ,

$$V(A, \rho)V(B, \rho) \geq \frac{1}{4}|\mathrm{Tr}(\rho[A, B])|^2, \quad (5.1)$$

where  $V(A, \rho) = \mathrm{Tr}(\rho A^2) - \mathrm{Tr}(\rho A)^2$ . A weakness of this formulation, however, lies in its dependence on the particular quantum state of the system to be measured, as such a bound becomes trivial for states lying in the kernel of the commutator. A different approach, pioneered by Deutsch [33], is to lower bound the total uncertainty about both measurement outcomes in a state-independent manner. Here, the most famous result is due to Maassen and Uffink [34], who bounded the sum of the Shannon entropies  $H(\cdot)$  (recall Definition M1) of the measurement outcome probability distributions,

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB}, \quad (5.2)$$

where  $\mathbf{p}^A(\rho)$  and  $\mathbf{p}^B(\rho)$  denote outcome probability distributions for measurements of  $A$  and  $B$  in a state  $\rho$ , and  $c_{AB} = \max_{ij} |\langle a_i | b_j \rangle|$  yields the state-independent lower bound.

This *entropic uncertainty relation* (EUR) has been improved in many respects. A tighter relation for observables fulfilling  $c_{AB} > 1/\sqrt{2}$  has been found using the Landau-Pollak uncertainty relation [89]. The study of EURs in the presence of classical or quantum

side information brought other improvements [38, 90]. In particular, the case of a trivial memory gives a bound on  $H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho))$  in terms of the von Neumann entropy of  $\rho$  [38]. Recently a majorisation approach [85, 88] led to improvements based on a more fine-grained measure of overlap between observables, where one looks at all  $|\langle a_i | b_j \rangle|$  [86]. In this chapter we want to broaden our understanding of entropic uncertainty relations by investigating the ways classical lack of knowledge affects them. Hence, we focus on EURs for mixed states, which arise not only due to imperfect experimental preparations, but also in scenarios of pure entangled states when we are interested in the uncertainties of two observables  $A$  and  $B$  on a particular subsystem. We will answer how one can cleanly separate the uncertainty that arises due to the intrinsic noncommutativity of the observables (i.e., due to the eigenstates of one observable being a coherent superposition of the eigenstates of the other observable) from that which does not. Moreover, using this splitting, we will show how the quantum component of uncertainty is unconstrained, but instead can be bounded relative to the amount of classical uncertainty present in the state. Finally, we will present a general result about the nonlinearity of all fixed-entropy bounds and introduce the concept of minimum uncertainty states in the presence of classical noise (studied in more detail in Chapter 6).

## 5.1 Splitting total uncertainty into quantum and classical parts

While intuitive and operationally meaningful, the entropic measure  $H(\mathbf{p}^A(\rho))$  quantifies total uncertainty and, as such, does not yield information about its origin: whether it is classical or quantum. To illustrate the difference between these two kinds of uncertainty, let us use the following example:

### Example 5.1: Classical vs quantum uncertainty for qubits

Consider a qubit system and the Pauli observable  $\sigma_z$ . Then, the Shannon entropy of the measurement outcome distribution is maximal and equal to  $\ln 2$  for both states  $|+\rangle\langle+|$  and  $\mathbb{I}/2$ . However, it is evident that in the former case the uncertainty is entirely quantum mechanical (as it arises from the fact the  $|+\rangle$  is a coherent superposition of  $|0\rangle$  and  $|1\rangle$ ), whereas in the latter case it is entirely classical.

Although there is no single correct way to decompose the total measurement uncertainty of an observable  $A$  in a state  $\rho$  into quantum and classical components, it is quite straightforward to establish natural criteria that such a decomposition should obey. Such a set of conditions was formulated by Luo [91], and represents the minimal requirements that any *quantum uncertainty measure*  $Q(\mathbf{p}^A(\rho))$  and any *classical uncertainty measure*  $C(\mathbf{p}^A(\rho))$  should satisfy. To these criteria we add additional two conditions specific to an entropic scenario, resulting in the following requirements:

#### Criteria for quantum-classical splitting of uncertainty

1. If a state  $\rho$  is pure, then  $C(\mathbf{p}^A(\rho))$  should vanish.
2. If  $[\rho, A] = 0$ , then the state is diagonal in the eigenbasis of the observable  $A$  and so  $Q(\mathbf{p}^A(\rho))$  should vanish.
3. Classical mixing increases the classical, but not the quantum, uncertainty, and so  $Q(\mathbf{p}^A(\rho))$  should be convex and  $C(\mathbf{p}^A(\rho))$  should be concave.
4.  $0 \leq Q(\mathbf{p}^A(\rho)), C(\mathbf{p}^A(\rho)) \leq H(\mathbf{p}^A(\rho))$ .
5.  $Q(\mathbf{p}^A(\rho))$  and  $C(\mathbf{p}^A(\rho))$  are functions of the probability distribution over the measurement outcomes of observable  $A$  and not of its eigenvalues.

The relative merits or weaknesses of these conditions can certainly be debated, but in what follows we simply use them as a guide for our entropic decomposition. We first focus on non-degenerate observables and then show how to extend our results also to the degenerate ones.

In light of the presented criteria, we observe that the central measurement entropy for non-degenerate observables can be expressed as  $H(\mathbf{p}^A(\rho)) = S(\mathcal{D}_A(\rho))$ , where  $S(\rho)$  is the von Neumann entropy of a quantum state  $\rho$ , and  $\mathcal{D}_A(\cdot)$  is the dephasing map in the basis of the eigenstates of  $A$ . The projective measurement of  $A$  is repeatable, and so it is reasonable to demand that a second measurement of  $A$  (in which the prior classical outcome of the first measurement is discarded) should not reveal any quantum uncertainties in the state, and be entirely classical. Therefore, we propose the relative entropy  $S(\rho||\mathcal{D}_A(\rho))$  as the appropriate measure of quantum uncertainty for the measurement of  $A$  in the state  $\rho$ . The following geometrical characterisation [80] of the relative entropy as a “distance”

from the manifold of classical states, further justifies our definition:

$$Q(\mathbf{p}^A(\rho)) := S(\rho||\mathcal{D}_A(\rho)) = \min_{\sigma \in \mathcal{I}_A} S(\rho||\sigma), \quad (5.3)$$

where  $\mathcal{I}_A$  denotes the set of states diagonal in the eigenbasis of  $A$ . Moreover, if we define

$$C(\mathbf{p}^A(\rho)) := S(\rho) \quad (5.4)$$

as our measure of classical uncertainty, we obtain an *additive* decomposition of the total entropic uncertainty,

$$H(\mathbf{p}^A(\rho)) = Q(\mathbf{p}^A(\rho)) + C(\mathbf{p}^A(\rho)). \quad (5.5)$$

To prove the equality in Eq. (5.5) we note that

$$\mathrm{Tr}(\rho \ln(\mathcal{D}_A(\rho))) = \mathrm{Tr}(\rho \mathcal{D}_A[\ln(\mathcal{D}_A(\rho))]) = \mathrm{Tr}(\mathcal{D}_A(\rho) \ln(\mathcal{D}_A(\rho))) = -S(\mathcal{D}_A(\rho)),$$

where we have used the fact that the dephasing operator is idempotent,  $\mathcal{D}_A(\mathcal{D}_A(\rho)) = \mathcal{D}_A(\rho)$ , and that trace is cyclic. By inserting the above result into the definition of  $Q(\mathbf{p}^A(\rho))$  we thus obtain

$$Q(\mathbf{p}^A(\rho)) = S(\rho||\mathcal{D}_A(\rho)) = -S(\rho) - \mathrm{Tr}(\rho \ln(\mathcal{D}_A(\rho))) = -S(\rho) + S(\mathcal{D}_A(\rho)).$$

To verify that the introduced measures conform to the presented criteria note that the von Neumann entropy is concave and vanishes for pure states, whilst the relative entropy  $S(\rho||\mathcal{D}_A(\rho))$  vanishes if and only if  $\rho = \mathcal{D}_A(\rho)$ , i.e., when  $\rho$  commutes with  $A$ . Finally, convexity of  $Q$  follows from the joint convexity of relative entropy.

In addition to providing an additive decomposition, let us note that our choice of quantum uncertainty  $Q(\mathbf{p}^A(\rho))$  has a natural interpretation as a measure of the superpositions present in  $\rho$  with respect to the eigenbasis of  $A$  [92], and as a resource within the resource theories of coherence [93] and  $U(1)$ -asymmetry [54]. Indeed, for the Hamiltonian observable  $H$ , the quantity  $Q(\mathbf{p}^H(\rho))$  coincides with the coherence measure that we used to study the resource theory of thermodynamics within Part I (see Eq. (3.4), where  $S(\rho||\mathcal{D}_H(\rho))$  is the quantum component of the free energy of  $\rho$ ). Moreover, within the quantum memory approach our choice of  $Q(\mathbf{p}^A(\rho))$  corresponds to quantum side information introduced by the system  $E$  purifying  $\rho$ , i.e., it is equivalent to conditional entropy  $H(A|E)$ , as discussed in Ref. [94]. These connections make the measure  $Q$  additionally attractive, and facilitate interpretation within a broader framework.

Note that our measure of classical uncertainty does not depend on the choice of the observable. This is due to the fact that we are considering here non-degenerate observables, projective measurements of which are perfectly sharp. In this situation any classical uncertainty can only be due to the fact that we are sampling a mixed state. A similar situation would be true for perfectly sharp measurements in classical physics. However, similarly to coarse-grained measurements in classical physics, the projective measurements of degenerate observables in quantum physics can be the source of additional classical uncertainty dependent on the degeneracy. We will now proceed to analyse this case in more detail.

### 5.1.1 Extension to general projective measurements

In the case of general degenerate observables the corresponding projective measurement is no longer sharp (rank-1), and so it is natural to demand that the classical uncertainty should reflect this degeneracy. Intuitively, the more a measurement coarse-grains the Hilbert space, the smaller the classical uncertainty will be. Consider a state  $\rho$  and a projective measurement  $\{\Pi_i\}$ . Define

$$\rho_i = \frac{\Pi_i \rho \Pi_i}{p_i}, \quad (5.6)$$

where  $p_i = \text{Tr}(\rho \Pi_i)$ . Let us also define the measurement map

$$\rho \longmapsto D_\Pi(\rho) = \sum_i p_i \rho_i = \sum_i \Pi_i \rho \Pi_i, \quad (5.7)$$

which associates to each state the post-measurement state (without post-selection). The relative entropy between the initial state and the post-measurement state is given by

$$S(\rho || D_\Pi(\rho)) = -S(\rho) - \text{Tr}(D_\Pi(\rho) \ln D_\Pi(\rho)) = -S(\rho) + S(D_\Pi(\rho)).$$

Given that  $\{\rho_i\}$  have orthogonal supports, one has

$$S(D_\Pi(\rho)) = S\left(\sum_i p_i \rho_i\right) = H(\mathbf{p}^\Pi(\rho)) + \sum_i p_i S(\rho_i),$$

so that

$$S(\rho || D_\Pi(\rho)) = -S(\rho) + H(\mathbf{p}^\Pi(\rho)) + \sum_i p_i S(\rho_i), \quad (5.8)$$

which yields the final splitting

$$H(\mathbf{p}^\Pi(\rho)) = Q(\mathbf{p}^\Pi(\rho)) + C(\mathbf{p}^\Pi(\rho)), \quad (5.9)$$

with

$$Q(\mathbf{p}^\Pi(\rho)) = S(\rho || D_\Pi(\rho)), \quad (5.10a)$$

$$C(\mathbf{p}^\Pi(\rho)) = S(\rho) - \sum_i p_i S(\rho_i). \quad (5.10b)$$

Notice that the classical uncertainty is now a function of the measurement  $\{\Pi_i\}$ . This is because the uncertainty depends on degeneracy of the measured observable, so on the coarse-graining (sharpness) of the corresponding measurement, and in general it will be lower than the von Neumann entropy, which is the uncertainty for a perfectly discriminating measurement. Let us illustrate this with the following example:

### Example 5.2: Unsharp measurement of a qutrit

Consider a qutrit state

$$\xi = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|),$$

and a projective measurement  $\{\Pi_i\}$  with

$$\Pi_1 = |0\rangle\langle 0| + |1\rangle\langle 1|, \quad \Pi_2 = |2\rangle\langle 2|.$$

Even though the von Neumann entropy of  $\xi$  is non-zero, the classical uncertainty of such  $\{\Pi_i\}$  measurement on  $\xi$  should vanish (as the outcome associated with  $\Pi_1$  has probability 1) and, as can be easily checked with the definition given by Eq. (5.10b), it does vanish.

Let us now verify that the introduced quantities satisfy the basic criteria we imposed. As the Criteria 1 and 5 are clearly satisfied, we will only comment on the remaining ones:

2.  $Q(\mathbf{p}^\Pi(\rho)) = 0$  if and only if the measurement is classical, i.e.,  $[\rho, \Pi_i] = 0 \quad \forall i$ .
3.  $Q(\mathbf{p}^\Pi(\cdot))$  is a convex function as it is defined in terms of relative entropy. To show that  $C(\mathbf{p}^\Pi(\cdot))$  is a concave function simply note that  $H(\mathbf{p}^\Pi(\cdot))$  is a concave function and taking into account the additive splitting, Eq. (5.9), together with convexity of  $Q(\mathbf{p}^\Pi(\cdot))$ , it is straightforward to show that  $C(\mathbf{p}^\Pi(\cdot))$  must be concave.
4.  $Q(\mathbf{p}^\Pi(\rho)) \geq 0$  as it is defined in terms of relative entropy. To show that  $C(\mathbf{p}^\Pi(\rho)) \geq 0$  it suffices to observe that it is equal to the quantum-classical mutual information introduced by Sagawa [95] (and independently by Groenewold [96] and Ozawa [97]), which is known to be positive. The upper bound for both classical and quantum

uncertainty clearly holds, as both quantities are positive and their sum equals the upper bound  $H(\mathbf{p}^\Pi(\rho))$ .

Finally, let us note one more property, which is not among the introduced criteria, but that supports our interpretation of  $C(\mathbf{p}^\Pi(\rho))$  as classical uncertainty. Namely, if a measurement  $\Pi'$  is a refinement of a measurement  $\Pi$ ,

$$\Pi_i = \sum_j \Pi'_{ij}, \quad p'_{ij} = \text{Tr}(\rho \Pi'_{ij}), \quad \rho'_{ij} = \frac{\Pi'_{ij} \rho \Pi'_{ij}}{p'_{ij}} \quad (5.11)$$

then

$$\forall \rho \quad C(\mathbf{p}^\Pi(\rho)) \leq C(\mathbf{p}^{\Pi'}(\rho)). \quad (5.12)$$

In other words,  $C(\mathbf{p}^\Pi(\rho))$  decreases under coarse-graining, as expected. To prove this, we simply need to show that

$$\sum_i p_i S(\rho_i) \geq \sum_{ij} p'_{ij} S(\rho'_{ij}).$$

Since all the terms are positive it is sufficient to prove that for every  $i$  we have

$$S(\rho_i) \geq \sum_j \frac{p'_{ij}}{p_i} S(\rho'_{ij}), \quad (5.13)$$

which can be done in the following way [98]. We first note that for any linear operator  $L$  we have  $S(LL^\dagger) = S(L^\dagger L)$ , as the eigenvalues of  $LL^\dagger$  and  $L^\dagger L$  are the same (which can be seen with the use of singular value decomposition). Using this fact and the concavity of the von Neumann entropy we get:

$$\sum_j \frac{p'_{ij}}{p_i} S(\rho'_{ij}) = \sum_j \frac{p'_{ij}}{p_i} S\left(\frac{\sqrt{\rho} \Pi'_{ij} \sqrt{\rho}}{p'_{ij}}\right) \leq S\left(\sum_j \frac{\sqrt{\rho} \Pi'_{ij} \sqrt{\rho}}{p_i}\right) = S\left(\frac{\sqrt{\rho} \Pi_i \sqrt{\rho}}{p_i}\right) = S(\rho_i).$$

Therefore, the inequality stated in Eq. (5.13) holds, so that the measure of classical uncertainty introduced in Eq. (5.10b) decreases under coarse-graining.

## 5.2 Non-existence of state-independent bounds

Whereas the Maassen-Uffink relation, Eq. (5.2), bounds the total entropic uncertainty  $H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho))$ , we would like to establish a finer set of conditions on the quantum component of the total uncertainty. However, as we will now show, a direct approach to bound  $Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho))$  fails, as for uncertainty measures satisfying the introduced

criteria this expression is entirely unconstrained over the set of all states, i.e., no state-independent lower (or upper) bound is possible for the total quantum uncertainty of  $A$  and  $B$  in the state  $\rho$ . Thus, in the next section we will present an indirect approach that bounds  $Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho))$  conditioned on the amount of classical uncertainty in the state.

While we are interested in a particular quantum-classical splitting, the following argument applies more generally. As we will discuss in detail in Chapter 6, one can define general families of functions satisfying some basic axioms and use any member of the family as a measure of uncertainty. The most general such family, introduced in Ref. [88], contains all functions  $u : \mathbf{p} \mapsto \mathbb{R}^+$  invariant under relabelling of the probability vector  $\mathbf{p}$  and monotonically increasing under random relabelling.<sup>1</sup> The results presented below apply to any such function  $u$  that is Schur-concave, i.e., whenever  $\mathbf{p}$  majorises  $\mathbf{q}$  we have  $u(\mathbf{p}) \leq u(\mathbf{q})$  (see Definition M3 for details on majorisation).

We start with the following general result concerning the total uncertainty of the measurement outcomes of two observables.

**Lemma 5.1: Existence of maximally uncertain pure state**

For every pair of observables  $A$  and  $B$ , and any general uncertainty function  $u$ , there exists a pure state  $|\psi_{AB}^*\rangle$  that simultaneously maximises the total uncertainty of both observables,

$$u(\mathbf{p}^A(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = u(\mathbf{p}^B(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = \max_{\mathbf{p}} u(\mathbf{p}). \quad (5.14)$$

*Proof.* For every pair of observables  $A, B$  there exists a pure state  $|\psi_{AB}^*\rangle$  that is unbiased in the eigenbases of both observables<sup>2</sup> [99]:

$$\mathbf{p}^A(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|) = \mathbf{p}^B(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|) = (1/d, \dots, 1/d) = \boldsymbol{\eta}.$$

The uniform distribution  $\boldsymbol{\eta}$  is majorised by all other distributions. Hence, as  $u$  is Schur-concave, one gets

$$\forall \mathbf{q} : \quad u(\mathbf{q}) \leq u(\mathbf{p}^A(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = u(\mathbf{p}^B(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)). \quad (5.15)$$

□

<sup>1</sup>For every  $q \in [0, 1]$  and every permutation  $\pi$ ,  $u(\mathbf{p}) \leq u(q\mathbf{p} + (1-q)\pi\mathbf{p})$ .

<sup>2</sup>This result will be discussed in more detail in Sec. 7.2.1 of Chapter 7.

Suppose we now want to split the general measure of total uncertainty into the sum of classical and quantum uncertainty components,

$$u(\mathbf{p}^A(\rho)) + u(\mathbf{p}^B(\rho)) = [u_Q(\mathbf{p}^A(\rho)) + u_Q(\mathbf{p}^B(\rho))] + [u_C(\mathbf{p}^A(\rho)) + u_C(\mathbf{p}^B(\rho))], \quad (5.16)$$

where we have two non-negative real-valued functions  $u_Q : (A, \rho) \mapsto \mathbb{R}^+$ , the *quantum uncertainty*, and  $u_C : (A, \rho) \mapsto \mathbb{R}^+$ , the *classical uncertainty*. Given the additive splitting, defined by Eq. (5.16), we could wonder if we can find a state-independent upper or lower bound on  $u_Q(\mathbf{p}^A(\cdot)) + u_Q(\mathbf{p}^B(\cdot))$  only. From the previous lemma we immediately infer this is impossible, if one demands that quantum and classical uncertainties satisfy the introduced criteria. Specifically, we have the following,

### Theorem 5.2: No state-independent bounds

No additive splitting admits a non-trivial state-independent bound for  $u_Q(\mathbf{p}^A(\cdot)) + u_Q(\mathbf{p}^B(\cdot))$  if the Criteria 1 and 2 are satisfied. In other words, there are no  $A, B, c(A, B) > 0$  and  $d(A, B) < 2 \max_{\mathbf{p}} u(\mathbf{p})$  such that:

$$\forall \rho : \quad c(A, B) < u_Q(\mathbf{p}^A(\rho)) + u_Q(\mathbf{p}^B(\rho)) < d(A, B). \quad (5.17)$$

*Proof.* Let us fix general observables  $A$  and  $B$ . From Lemma 5.1 there is always a pure state  $|\psi_{AB}^*\rangle$  achieving

$$u(\mathbf{p}^A(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = u(\mathbf{p}^B(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = \max_{\mathbf{p}} u(\mathbf{p}).$$

But from Criterion 1,  $u_C$  vanishes on pure states, so

$$u_Q(\mathbf{p}^A(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) + u_Q(\mathbf{p}^B(|\psi_{AB}^*\rangle\langle\psi_{AB}^*|)) = 2 \max_{\mathbf{p}} u(\mathbf{p}).$$

The maximally mixed state is diagonal in any basis, therefore from Criterion 2

$$u_Q(\mathbf{p}^A(\mathbb{I}/d)) + u_Q(\mathbf{p}^B(\mathbb{I}/d)) = 0.$$

□

### Remark 5.3

Given the additive splitting and the fact that the sum of  $u_Q$ 's is unconstrained, we can also deduce that there is no constraint on the sum of classical components.

The above theorem implies that for the particular quantum-classical splitting with  $u_Q = Q$

and  $u_C = C$ , as defined in Eqs. (5.3)-(5.4), we have

$$0 \leq Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \leq 2 \ln d, \quad (5.18)$$

and only the total uncertainty has a state-independent lower bound.

## 5.3 Quantum uncertainty relations and fixed-entropy lower bounds

Since there are no non-trivial state-independent bounds for the sum of quantum uncertainties, we will seek bounds that are conditional on the degree of classical uncertainty in the state. For our choice of classical and quantum uncertainties, this will lead us to a refinement of the Maassen-Uffink relation in terms of *fixed-entropy* lower bounds. Schematically, we would like to obtain entropic relations for the total quantum uncertainty of the form

$$u_Q(\mathbf{p}^A(\rho)) + u_Q(\mathbf{p}^B(\rho)) \geq f(A, B, \text{“mixedness of } \rho\text{”}). \quad (5.19)$$

Specialising to the particular case where  $u_Q(\mathbf{p}^A(\rho))$  and  $u_Q(\mathbf{p}^B(\rho))$  are the entropic quantum uncertainties of Eq. (5.3), and the mixedness is measured by the von Neumann entropy, we are looking for

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq f(A, B, S(\rho)). \quad (5.20)$$

Among all fixed-entropy bounds we will be particularly interested in these that are tighter than the known results, and thus we define the following:

### Definition 5.1: Strong fixed-entropy bound (SFEB)

We say that a fixed-entropy bound is *strong* if it satisfies the following two conditions:

1. Being at least as strong as the Maassen-Uffink bound:

$$\forall \rho, \forall A, B, \text{ Eq.(5.20)} \Rightarrow H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq -2 \log c_{AB}.$$

2. For  $d$ -dimensional systems  $f(A, B, \ln d) = 0$ .

The second requirement captures the classical feature of the maximally mixed state,

namely that it should not exhibit any quantum uncertainty, consistently with the vanishing of all coherences.

As the quantum uncertainty measure is the relative entropy between the state  $\rho$  and its dephased version following the projective measurement, we can make use of certain well-known entropic properties to develop meaningful lower bounds on the total quantum uncertainty in a state. Let us define the states  $\rho_A = \mathcal{D}_A(\rho)$  and  $\rho_B = \mathcal{D}_B(\rho)$  to shorten the notation. We then have

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) = S(\rho||\rho_A) + S(\rho||\rho_B) \geq S(\rho_B||\mathcal{D}_B(\rho_A)) + S(\rho||\rho_B),$$

where the inequality follows from the fact that the relative entropy is contractive under CPTP maps, so in particular under a dephasing map. We further have

$$S(\rho_B||\mathcal{D}_B(\rho_A)) = -S(\rho_B) - \text{Tr}(\rho_B \ln(\mathcal{D}_B(\rho_A))) = -S(\rho_B) - \text{Tr}(\rho \ln(\mathcal{D}_B(\rho_A))),$$

which implies that

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq -S(\rho) - \text{Tr}(\rho \ln(\mathcal{D}_B(\rho_A))), \quad (5.21a)$$

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq -S(\rho) - \text{Tr}(\rho \ln(\mathcal{D}_A(\rho_B))), \quad (5.21b)$$

where the second inequality is obtained simply by inverting the roles of  $A$  and  $B$ .

For the special case of the observables  $A$  and  $B$  being mutually unbiased, we have that  $\mathcal{D}_A(\rho_B) = \mathcal{D}_B(\rho_A) = \mathbb{I}/d$ , which implies

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq \ln d \left(1 - \frac{S(\rho)}{\ln d}\right). \quad (5.22)$$

We see that in that case one can factor out a “mixedness” factor of  $(1 - S(\rho)/\ln d)$  that accounts for the contribution from the classical uncertainty in the state. This bound turns out to be an optimal one, as it is achieved for states diagonal in the eigenbasis of either  $A$  or  $B$ . Let us also note that Eq. (5.22) implies a refinement of the Maassen-Uffink relation for mutually unbiased observables:

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq \ln d + S(\rho), \quad (5.23)$$

which agrees with the result found in Ref. [38] for the case of a trivial quantum memory.

Beyond the case of mutually unbiased observables we see that the right-hand sides of

inequalities (5.21a)-(5.21b) may be written as

$$-S(\rho) - \sum_{i=1}^d p_i^B(\rho) \ln \left( \sum_{j=1}^d |c_{ij}|^2 p_j^A(\rho) \right), \quad (5.24a)$$

$$-S(\rho) - \sum_{i=1}^d p_i^A(\rho) \ln \left( \sum_{j=1}^d |c_{ij}|^2 p_j^B(\rho) \right), \quad (5.24b)$$

where  $c_{ij} = \langle a_i | b_j \rangle$ . One can obtain a convenient lower bound by replacing the terms  $|c_{ij}|$  with their maximum value  $c_{AB}$ , which yield the following relation

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} \left( 1 + \frac{S(\rho)}{2 \ln c_{AB}} \right). \quad (5.25)$$

Let us again note that Eq. (5.25) implies a refinement of Maassen-Uffink uncertainty relation, specified in Eq. (5.2),

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} + S(\rho). \quad (5.26)$$

We note that the above refinement was also obtained in Refs. [38] and [100] using different techniques.

### 5.3.1 Strong fixed-entropy lower bounds are nonlinear for $d > 2$

In contrast to the MUB case, the bound given by Eq. (5.25) has the disadvantage of having a mixedness factor which is not independent of the observables  $A$  and  $B$ . We might conjecture that a stronger bound holds, where the mixedness factor is independent of the observables, just like for the mutually unbiased observables [Eq. (5.22)],

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} \left( 1 - \frac{S(\rho)}{\ln d} \right). \quad (5.27)$$

We note that this bound would be an SFEB that is linear in von Neumann entropy  $S(\rho)$ . By the following theorem, it cannot hold for any dimension  $d > 2$ .

**Theorem 5.3: No linear SFEB for  $d \geq 3$**

For all  $d \geq 3$  all strong fixed-entropy bounds are nonlinear in  $S(\rho)$ .

Before presenting the proof let us note that since  $H(\mathbf{p}^A(\rho)) = Q(\mathbf{p}^A(\rho)) + S(\rho)$ , the same conclusion applies to the usual entropic uncertainty relations. Hence, all bounds proposed so far in the literature that are linear in  $S(\rho)$ , are inevitably either weaker than Maassen-Uffink for at least some states and observables, or they are not tight for the maximally

mixed state. The case of  $d = 2$  is special and will be discussed separately in the next section.

*Proof.* The proof proceeds in three steps: first, we construct the weakest linear bound satisfying SFEB conditions 1 and 2. A violation of this bound will imply the violation of any other bound with the same properties. Second, we produce counterexamples for dimension 3, 4, 5, i.e., construct observables  $A$  and  $B$  as well as states  $\rho_3$ ,  $\rho_4$  and  $\rho_5$  for which the weakest bound is violated. Third, we show that these already imply that a counterexample exists in any dimension  $d > 5$ .

**Step 1.** All bounds linear in  $S(\rho)$  have the form

$$f(A, B, S(\rho)) = a + bS(\rho), \quad (5.28)$$

for some  $a, b \in \mathbb{R}$ . Condition 2 implies  $b = -a/\ln d$ , hence

$$f(A, B, S(\rho)) = a \left( 1 - \frac{S(\rho)}{\ln d} \right).$$

Condition 1 implies  $a \geq -2 \ln c_{AB}$ , so

$$f(A, B, S(\rho)) \geq -2 \ln c_{AB} \left( 1 - \frac{S(\rho)}{\ln d} \right) := f_w(A, B, S(\rho)), \quad (5.29)$$

where  $f_w$  denotes the weakest linear bound satisfying the requirements 1 and 2. Notice that for any  $d > 2$ ,

$$f_w(A, B, S(\rho)) \geq -2 \ln c_{AB} \left( 1 - \frac{S(\rho)}{\ln 3} \right). \quad (5.30)$$

**Step 2.** Let  $|a_i\rangle$  and  $|b_i\rangle$  denote the eigenstates of observables  $A$  and  $B$  acting on a  $d$ -dimensional Hilbert space. Assume that these two bases are linked by a rotation  $R_d = \exp(\theta_d S_d)$ , where  $S_d$  is the  $d$ -dimensional skew-symmetric matrix such that  $|S_{ij}| = 1 - \delta_{ij}$  and  $\theta_d$  are real numbers. We choose

$$\theta_3 = \frac{4\pi}{7}, \quad \theta_4 = \frac{\pi}{2}, \quad \theta_5 = \pi.$$

The following states, written in the  $\{|a_i\rangle\}$  basis, violate the bound of Eq. (5.29) in dimensions  $d = 3, 4, 5$ , respectively:

$$\rho_3 = \frac{1}{100} \begin{bmatrix} 61 & -15 & 0 \\ -15 & 26 & 8 \\ 0 & 8 & 13 \end{bmatrix},$$

$$\rho_4 = \frac{1}{100} \begin{bmatrix} 8 & 3 & 4 & -3 \\ 3 & 6 & -3 & 0 \\ 4 & -3 & 8 & -3 \\ -3 & 0 & -3 & 78 \end{bmatrix},$$

$$\rho_5 = \frac{1}{100} \begin{bmatrix} 19 & 5 & -6 & -2 & -5 \\ 5 & 49 & -11 & 0 & 0 \\ -5 & -11 & 22 & -3 & 2 \\ -2 & 0 & -3 & 5 & 0 \\ -5 & 0 & 2 & 0 & 5 \end{bmatrix}.$$

**Step 3.** We will now show that the existence of the counterexample in dimension 3 immediately implies the existence of counterexamples in any dimension  $d > 5$ . Hence our counterexamples for  $d = 3, 4, 5$  immediately imply the result for all  $d > 2$ . Fix a  $d$ -dimensional Hilbert space  $\mathcal{H}_d$  and consider the 3-dimensional subspace  $\mathcal{H}_3$  spanned by  $\{|a_i\rangle\}$  and  $\{|b_i\rangle\}$  given by the counterexample of Step 2. We can complete  $\{|a_i\rangle\}$  and  $\{|b_i\rangle\}$  to a basis in  $\mathcal{H}_d = \mathcal{H}_3 \oplus \mathcal{H}_3^\perp$  by choosing

$$\langle a_i | b_j \rangle = \frac{1}{\sqrt{d-3}}, \quad \forall i, j = 4, \dots, d. \quad (5.31)$$

This can be done as  $\mathcal{H}_3^\perp$  is a  $d-3$  dimensional Hilbert space and one can always find two mutually unbiased bases as long as  $d > 4$ . Let us call  $c_{AB}^{(3)} = \max_{i,j=1}^3 |\langle a_i | b_j \rangle| = 0.6851$ . Notice that

$$c_{AB}^{(d)} := \max_{i,j=1}^d |\langle a_i | b_j \rangle| = \max \left\{ c_{AB}^{(3)}, \frac{1}{\sqrt{d-3}} \right\} = c_{AB}^{(3)}, \quad (5.32)$$

for all  $d \geq 6$ . Exploiting this construction and Eq. (5.30) it is easy to see that  $\rho_3$  (seen now as a quantum state in  $\mathcal{H}_d$ ) violates the bound of Eq. (5.29) in any dimension  $d > 5$ .  $\square$

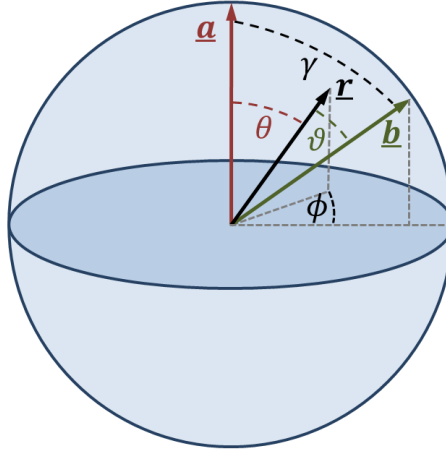
### 5.3.2 Linear strong fixed-entropy lower bound for qubits

As is a common situation in quantum information, the case of a qubit system is special and the bound conjectured in Eq. (5.27) holds. Thus, we have the following theorem:

#### Theorem 5.4: SFEB for qubit systems

Given arbitrary qubit observables  $A$  and  $B$ , for all states  $\rho$  we have

$$Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} \left( 1 - \frac{S(\rho)}{\ln 2} \right). \quad (5.33)$$


 Figure 5.1: *Parametrisation of the Bloch sphere.*

*Proof.* A general qubit observable has the form  $A = \alpha_1 \mathbb{I} + \alpha_2 \mathbf{a} \cdot \boldsymbol{\sigma}$ . However, as entropic uncertainty measures depend only on the eigenstates of observables and not on their eigenvalues, we can restrict our considerations to observables  $A = \mathbf{a} \cdot \boldsymbol{\sigma}$  and  $B = \mathbf{b} \cdot \boldsymbol{\sigma}$ . Without loss of generality one can choose

$$\mathbf{a} = (0, 0, 1), \quad \mathbf{b} = (\sin \gamma, 0, \cos \gamma), \quad (5.34)$$

and it is enough to restrict to  $\gamma \in [0, \pi/2]$ , as for entropic quantities  $\mathbf{a}$  and  $-\mathbf{a}$  are indistinguishable. In this setting one has  $c_{AB} = \cos \frac{\gamma}{2}$ . A general qubit state can now be written as

$$\rho = \frac{\mathbb{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}, \quad \mathbf{r} = r(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (5.35)$$

with  $r \in [0, 1]$ ,  $\theta \in [0, \pi]$  and  $\phi \in [0, 2\pi]$ . All three Bloch vectors  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{r}$  are depicted in the Bloch sphere in Fig. 5.1. The probability distributions of the outcomes of  $A$  and  $B$  in a state  $\rho$  are given by

$$\mathbf{p}^A(\rho) = \left( \frac{1 + r \cos \theta}{2}, \frac{1 - r \cos \theta}{2} \right), \quad \mathbf{p}^B(\rho) = \left( \frac{1 + r \cos \vartheta}{2}, \frac{1 - r \cos \vartheta}{2} \right), \quad (5.36)$$

where  $\cos \vartheta = \cos \phi \sin \theta \sin \gamma + \cos \theta \cos \gamma$ .

Introducing the binary entropy to shorten the notation,

$$h_2(p) = -p \ln p - (1 - p) \ln(1 - p), \quad (5.37)$$

the conjectured entropic uncertainty relation for qubits, Eq. (5.33), takes the following form

$$h_2\left(\frac{1 + r \cos \theta}{2}\right) + h_2\left(\frac{1 + r \cos \vartheta}{2}\right) + 2 \ln\left(\cos \frac{\gamma}{2}\right) - 2h_2\left(\frac{1 + r}{2}\right) \left(1 + \frac{\ln\left(\cos \frac{\gamma}{2}\right)}{\ln 2}\right) \geq 0.$$

In order to prove that the above inequality always holds let us first denote its left hand side by  $F(\theta, \vartheta, \gamma, r)$ . We will find its minimum and show that it is greater or equal to zero. Computing the derivative with respect to  $r$ ,

$$\begin{aligned} \frac{\partial}{\partial r} F(\theta, \vartheta, \gamma, r) &= -\tanh^{-1}(r \cos \theta) \cos \theta - \tanh^{-1}(r \cos \vartheta) \cos \vartheta \\ &\quad + 2 \tanh^{-1}(r) \left( 1 + \frac{\ln \left( \cos \frac{\gamma}{2} \right)}{\ln 2} \right), \end{aligned}$$

we will show it is always non-negative. Using the Taylor series representation of  $\tanh^{-1}$  one finds that the above derivative is given by

$$\sum_{n=0}^{\infty} \frac{r^{2n+1}}{2n+1} \left( 2 - \cos^{2n+2} \theta - \cos^{2n+2} \vartheta + \frac{2}{\ln 2} \ln \left( \cos \frac{\gamma}{2} \right) \right).$$

Now, let us denote the coefficient in parentheses standing by the  $n$ -th term by  $a_n$  and note that for all  $n$  we have  $a_0 \leq a_n$ . Therefore, if we can show that  $a_0 \geq 0$  then all the coefficients are positive and, taking into account the positivity of  $r$ , the considered derivative is positive for all  $\theta$ ,  $\vartheta$  and  $\gamma$ . Using the explicit dependence of  $\cos \vartheta$  on  $\theta$ ,  $\gamma$  and  $\phi$  we see that

$$a_0 \geq \sin^2 \theta + \frac{2}{\ln 2} \ln \left( \cos \frac{\gamma}{2} \right) + \min_{+,-} \sin^2(\theta \pm \gamma).$$

Introducing  $\delta = \theta \pm \gamma/2$  one gets that for  $a_0 \geq 0$  to be true one has to prove that

$$1 - \cos 2\delta \cos \gamma + \frac{2}{\ln 2} \ln \left( \cos \frac{\gamma}{2} \right) \geq 0.$$

The minimum of the left hand side of above inequality is achieved for  $\delta = 0$ , hence it is enough to prove that

$$2 \sin^2 \frac{\gamma}{2} + \frac{2}{\ln 2} \ln \left( \cos \frac{\gamma}{2} \right) \geq 0,$$

which requires just a straightforward calculation. Therefore, we have proven that

$$\forall \theta, \vartheta, \gamma \quad \frac{\partial}{\partial r} F(\theta, \vartheta, \gamma, r) \geq 0, \quad (5.38)$$

and, since  $F(\theta, \vartheta, \gamma, 0) = 0$ , we get  $F(\theta, \vartheta, \gamma, r) \geq 0$ .  $\square$

Let us now compare this qubit-specific uncertainty relation with other known bounds: the Maassen-Uffink bound [34] given by Eq. (5.2), the bound coming from uncertainty

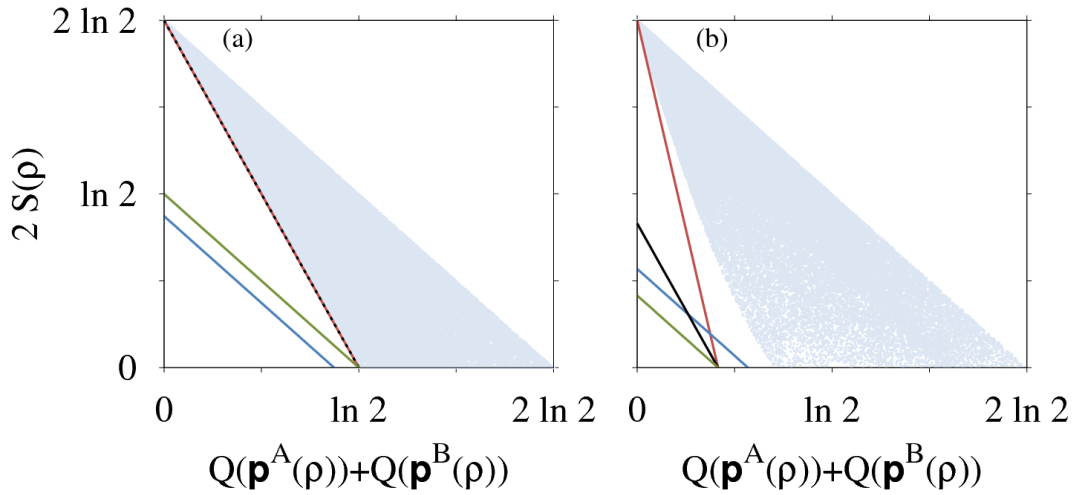


Figure 5.2: *QC uncertainty plots for qubits.* Sum of classical vs sum of quantum uncertainties plotted for  $10^5$  random qubit states. The red line is the strong fixed-entropy bound, the green line is the Maassen-Uffink bound, the black line is the bound coming from uncertainty relations with trivial quantum memory, and the blue line is the strong majorisation bound. The bounds are shown for (a)  $A$  and  $B$  mutually unbiased; (b) Eigenstates of  $A$  and  $B$  separated by angle  $\gamma = \pi/3$  on the Bloch sphere, corresponding to  $c_{AB} \approx 0.8660$ .

relations with trivial quantum memory [38] and the strong majorisation bound of Ref. [86]. The latter two for qubit systems read:

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} + S(\rho), \quad (5.39a)$$

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq h_2(c_{AB}), \quad (5.39b)$$

where  $h_2$  is the binary entropy defined in Eq. (5.37). In order to provide the comparison we plot the sum of classical uncertainties,  $C(\mathbf{p}^A(\rho)) + C(\mathbf{p}^B(\rho)) = 2S(\rho)$ , versus the sum of quantum uncertainties,  $Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho))$ , for random quantum states  $\rho$  (we will refer to such plots as quantum-classical, or QC, uncertainty plots). In Fig. 5.2 this data is presented for qubit systems, together with all four discussed bounds. As can be seen, for states with low von Neumann entropy the majorisation bound outperforms Eq. (5.33), whereas for states with high von Neumann entropy the discussed bound outperforms the majorisation bound. None of the bounds are however optimal as they are linear in the von Neumann entropy, whereas numerical investigations show that the minimum uncertainty curve in the  $Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho))$  versus  $2S(\cdot)$  plane is nonlinear (apart from the case of

$A$  and  $B$  being mutually unbiased). Hence, none of these bounds can be optimal for all values of von Neumann entropy. Finally, let us note that the discussed relation for qubits implies that

$$H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) \geq -2 \ln c_{AB} + 2S(\rho) \left(1 + \frac{\ln c_{AB}}{\ln 2}\right), \quad (5.40)$$

which yields an even stronger refinement of the Maassen-Uffink bound for qubits than the one presented in Eq. (5.26).

## 5.4 Minimum uncertainty states with fixed von Neumann entropy

As demonstrated, the fixed-entropy bounds on the sum of quantum uncertainties are generally nonlinear in the von Neumann entropy  $S$ . Therefore, in order to get more insight into their form, we will now focus directly on the states with fixed  $S$  that minimise  $Q(\mathbf{p}^A(\rho)) + Q(\mathbf{p}^B(\rho))$ . We will refer to them as *minimum uncertainty states* (MUS) with fixed von Neumann entropy (these are the states that form the optimal curve in the QC uncertainty plots). In this chapter we will only consider the case of  $d = 2$  in order to point out that even in this simplest scenario mixed MUSs have non-trivial structure that is independent of the pure MUS with  $S(\rho) = 0$ . Hence, we will show that finding the optimal pure state is not enough to construct an optimal state with fixed  $S(\rho) > 0$ . We will then investigate this issue more deeply in Chapter 6.

The form of minimum uncertainty pure states for qubits has been studied previously in Refs. [101, 102] and it has been shown that they exhibit the following dependence on  $\gamma$ . For  $\gamma < \gamma_c$  (where  $\gamma_c \approx 1.17056$  was found numerically [102]) the optimal state is represented by the Bloch vector

$$\mathbf{r}_< = \left(\sin \frac{\gamma}{2}, 0, \cos \frac{\gamma}{2}\right), \quad (5.41)$$

i.e., it lies in the middle between the eigenstates of  $A$  and  $B$  on the Bloch sphere. For  $\gamma > \gamma_c$  a parametric bifurcation occurs – the number of optimal states doubles and they are represented by the Bloch vectors

$$\mathbf{r}_> = \left(\sin \left(\frac{\gamma}{2} \pm \beta\right), 0, \cos \left(\frac{\gamma}{2} \pm \beta\right)\right), \quad (5.42)$$

where  $\beta$  is a non-elementary function of  $\gamma$  that increases from  $\beta = 0$  for  $\gamma = \gamma_c$  to  $\beta = \gamma/2$  for  $\gamma = \pi/2$ . In the Bloch sphere picture when  $\gamma$  exceeds  $\gamma_c$  the two optimal states start to move away from the vector lying symmetrically between the eigenstates of  $A$  and  $B$ , and move towards these eigenstates, eventually overlapping with them for  $\gamma = \pi/2$ , i.e., for  $A$  and  $B$  being mutually unbiased.

One might suspect that a general MUS with fixed von Neumann entropy can be

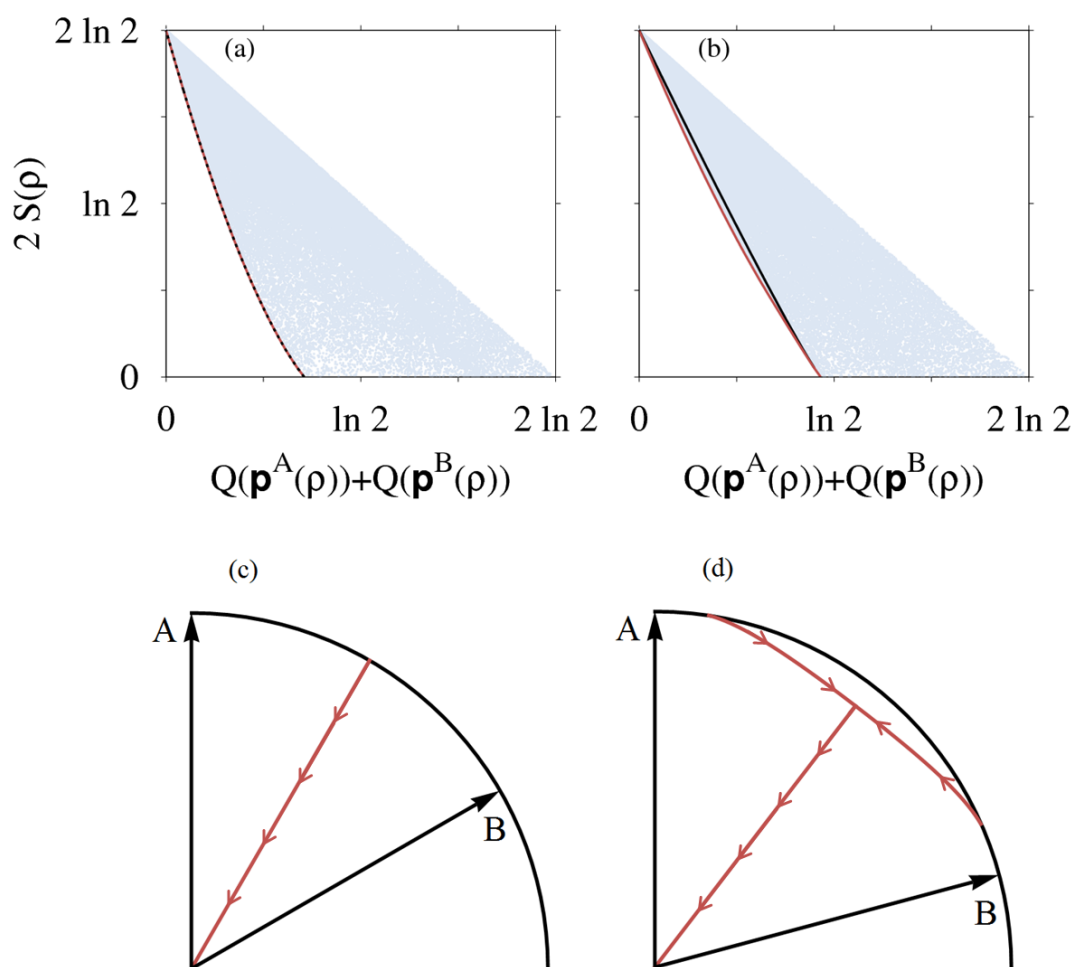


Figure 5.3: *MUSs with fixed von Neumann entropy for qubits.* (a-b) Sum of classical vs sum of quantum uncertainties plotted for  $10^5$  random qubit states. Red line represents numerically optimised MUSs with fixed von Neumann entropy, whereas the black line represents mixtures of a pure MUS with the maximally mixed state. In (a) the eigenstates of  $A$  and  $B$  are separated by  $\gamma = 60^\circ < \gamma_c$ , in (b) they are separated by  $\gamma = 75^\circ > \gamma_c$ . (c-d) The trajectory of MUSs with fixed von Neumann entropy in the first quadrant of the plane spanned by  $\mathbf{a}$  and  $\mathbf{b}$ . In (c)  $A$  and  $B$  are given as in (a), in (d) as in (b).

obtained just by mixing the pure MUS with the maximally mixed state. Interestingly, however, we will show that this is not the case, which supports the claim that MUSs with fixed von Neumann entropy are not just a trivial extension of pure MUSs. The behaviour of such a qubit MUS is shown in Fig. 5.3. It is straightforward to prove that a qubit MUS, independently of its von Neumann entropy, must lie in the plane spanned by  $\mathbf{a}$  and  $\mathbf{b}$ , thus having the form

$$\mathbf{r} = r(\sin \theta, 0, \cos \theta), \quad (5.43)$$

where  $r \in [0, 1]$ . One can numerically verify that for  $\gamma < \gamma_c$ , when a pure MUS is represented by the Bloch vector  $\mathbf{r}_<$ , a general MUS with fixed von Neumann entropy is given by  $r\mathbf{r}_<$ , i.e., the same Bloch vector, just shorter [see Fig. 5.3(a-c)]. Hence in this regime a general MUS is obtained by mixing the pure MUS with the maximally mixed state. However, for  $\gamma > \gamma_c$  this is no longer the case, as for a given  $\mathbf{r}_>$  increasing the mixedness decreases  $\beta$ , and MUSs flow towards the  $r\mathbf{r}_<$  solution [see Fig. 5.3(b-d)]. Numerical investigations performed for qutrits (see Fig. 5.4) suggest that this non-trivial

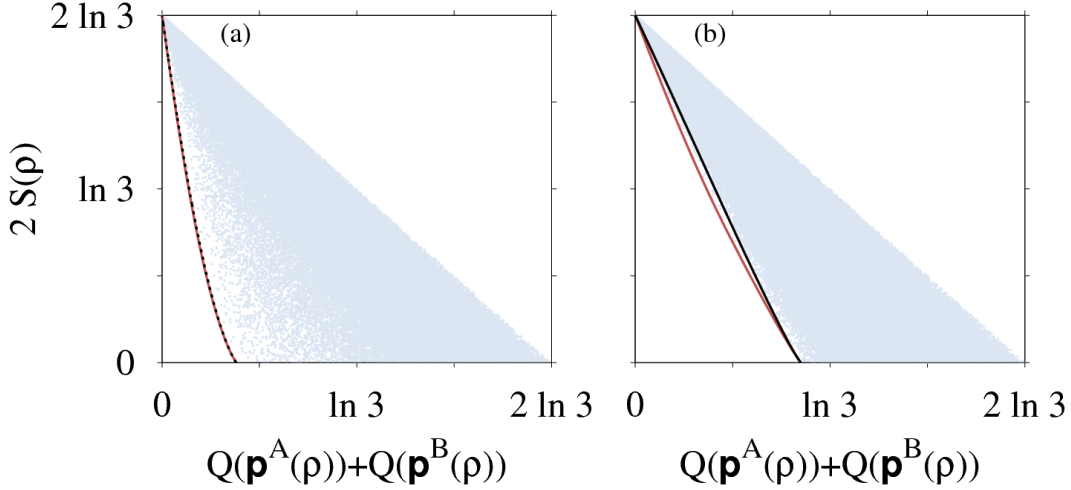


Figure 5.4: *QC uncertainty plots for qutrits.* Sum of classical vs sum of quantum uncertainties plotted for  $10^5$  random qutrit states. Red line represents numerically optimised MUSs with fixed von Neumann entropy, whereas the black line represents mixtures of a pure MUS with the maximally mixed state. Observables  $A$  and  $B$  chosen so that the eigenstates of  $B$  are connected with the eigenstates of  $A$  by a rotation around  $(1,1,1)$  axis by  $\alpha$ . (a)  $\alpha = \pi/6$ , corresponding to  $c_{AB} \approx 0.9107$  and (b)  $\alpha = \pi/3$ , corresponding to  $c_{AB} \approx 0.6667$ .

structure of MUSs is a general feature, not only limited to qubits. We devote the following chapter to further investigate this problem.

Finally, we would also like to make a short comment on the states that are the opposite of MUS with fixed von Neumann entropy - the maximal uncertainty states with fixed von Neumann entropy. Such states are important from the point of view of *certainty relations* [103], where one is interested in the upper (instead of lower) bound on the total uncertainty. As can be seen in Figs. 5.2-5.4, these states form a straight line in QC uncertainty plots, connecting maximally mixed state and the pure state  $|\psi_{AB}^*\rangle$ , which is unbiased in eigenbases of both observables (see Lemma 5.1). Thus, the states of fixed von Neumann entropy that maximise the sum of quantum uncertainties have a particularly simple form:  $p\mathbb{I}/d + (1-p)|\psi_{AB}^*\rangle\langle\psi_{AB}^*|$ . Note that this means that for every fixed von Neumann entropy there exists a state  $\rho$  that maximises the sum of total uncertainties, i.e., for which  $H(\mathbf{p}^A(\rho)) + H(\mathbf{p}^B(\rho)) = 2\log d$ .

# Chapter 6

## Minimum uncertainty states in the presence of classical noise

Entropic uncertainty relations quantify the impossibility of preparing a quantum state with the statistics of two noncommuting observables being simultaneously sharp. As such, they describe fundamental features of the quantum world. On the other hand, whenever a given uncertainty relation is based on a specific choice of uncertainty measure, it is biased. Although this choice may be justified by other assumptions, it unavoidably limits the universality of the results obtained. The particular example of this problem that we will focus on in this chapter is the form of *minimum uncertainty states* (MUSs). Already in the simplest case of two qubit observables,  $A = \mathbf{a} \cdot \boldsymbol{\sigma}$  and  $B = \mathbf{b} \cdot \boldsymbol{\sigma}$ , one easily finds that a MUS is not unique and depends on the chosen measure. Indeed, if as a measure of uncertainty we choose the Shannon entropy  $H_1$  of the outcome probabilities (as we did in Chapter 5), a pure MUS may be given by eigenvectors of either  $A$  or  $B$  [101, 102]; if we choose the min-entropy  $H_\infty$  instead, pure MUSs are always described by the Bloch vector lying in the middle between the two closest eigenvectors of  $A$  and  $B$ . Moreover, as we have shown in the previous chapter, the structure of a state minimising a particular uncertainty measure can be affected by the introduction of classical noise (mixedness constraint). Hence, the problem of finding an optimal MUS being constrained by some level of classical ignorance, has many solutions corresponding to different uncertainty measures. Although each such solution may shed some light on new aspects of uncertainty relations, the general effect of classical noise may be hard to grasp.

In this chapter we follow a recent “universal” approach [85–88] that overcomes this

problem by making statements that are independent of the particular measure of uncertainty being used. It is based on a minimal requirement that all valid uncertainty functions should satisfy: the act of forgetting information about a random variable cannot decrease its uncertainty. This approach restricts uncertainty functions to the family of Schur-concave functions, so that any relation valid for all such functions provides a general statement about uncertainty. Here, we also analyse the effect of two additional requirements: the additivity of uncertainty for independent random variables and continuity. Adding these further restrictions one after the other gives two more frameworks for studying uncertainty. These allow us to restrict the set of all uncertainty measures to the one-parameter family of Rényi entropies  $H_\alpha$  with parameter  $\alpha \in \mathbb{R}$  or  $\alpha \in \mathbb{R}_+$ .

Within these three frameworks we will then study the structure of quantum states that minimise uncertainty. In the spirit of the universal approach we ask: what quantum states - if any - simultaneously minimise all possible uncertainty measures? In other words we will be looking for a *universal* minimum uncertainty state that minimises uncertainty of the measurement outcome distributions of two noncommuting observables with respect to all information-theoretic uncertainty measures belonging to a chosen family. As we will prove, such state does not exist within the full unconstrained state space for any choice of the family of uncertainty functions. However, we will show that the introduction of classical noise simplifies the structure of MUSs and can lead to the emergence of a universal MUS or its approximate notion. We will also identify the crucial role played by Rényi entropy of order  $\alpha = 2$  for MUSs in the presence of classical noise. Finally, exploiting relations between the measures of uncertainty and the measures of departure from thermodynamic equilibrium, we will link the results on uncertainty presented in this chapter with the thermodynamic considerations from Part I of this thesis. In particular, we will point out how a so-far neglected measure of non-equilibrium, the  $\alpha = 2$  Rényi divergence between a given state and a thermal Gibbs state, plays a crucial role in near-equilibrium thermodynamic transformations.

## 6.1 General families of uncertainty measures

Intuitively, an uncertainty measure  $u$  is a function that assigns a real positive number to every probability distribution  $\mathbf{p}$ , reflecting the “spread” of  $\mathbf{p}$ . However, there is no unique

way of measuring the uncertainty of a probability distribution; quite the contrary, there exists a plethora of different information-theoretic functions [42]. This is linked to the fact that there are different ways of assessing uncertainty and making bets, depending on the rules of the probabilistic game being played. For example, making a bet on a single event is very different from making bets on many, independent and identically distributed ones. In the former case one would look at a single-shot entropy, whereas in the latter one may choose the Shannon entropy. Also, depending on the stake, one may want to follow a very risk-averse strategy (and, e.g., look at the Hartley entropy  $H_0$ ) or, on the contrary, be risk-prone (and, e.g., look at the min-entropy  $H_\infty$ ). This is reflected by different choices of the relevant uncertainty functions, as each of them captures a different aspect of the “spread” of  $\mathbf{p}$ . However, one can ask: what is the set of all possible uncertainty functions?

The basic idea is that all uncertainty functions must satisfy some elementary requirements; e.g., all of them should assign zero uncertainty to the sharp probability distribution  $\mathbf{p} = (1, 0, \dots, 0)$ . In what follows we will describe and motivate conditions defining general families of bona-fide uncertainty functions. We will call a probability distribution  $\mathbf{p}$  universally less uncertain than  $\mathbf{q}$  – according to some chosen set of uncertainty measures  $\mathcal{F}$  – if  $u(\mathbf{p}) \leq u(\mathbf{q})$  for all  $u \in \mathcal{F}$ .

### 6.1.1 Minimal requirement of Schur-concavity

Recently, a general condition has been proposed that a function  $u$  should satisfy in order to measure uncertainty [88]. It is given by

$$u(\mathbf{p}) \leq u(\lambda\mathbf{p} + (1 - \lambda)\Pi\mathbf{p}) \quad \text{for } \lambda \in [0, 1], \quad (6.1)$$

where  $\Pi$  is any permutation of the probability vector. In other words, a random relabelling of a probability distribution cannot decrease the uncertainty. Notice that since permutations are reversible, this immediately implies that any  $u$  must be a function of the probability vector only and not of the way we label events, i.e.,  $u(\mathbf{p}) = u(\Pi\mathbf{p})$ . This is in accordance with a much older concept, introduced by Deutsch [33], that an information-theoretic measure of uncertainty for a given observable should not depend on its eigenvalues.

As Birkhoff’s theorem states that the convex hull of permutation matrices is given by

the set of bistochastic matrices  $\{\Lambda\}$  [44], the above axiom is equivalent to

$$u(\mathbf{p}) \leq u(\Lambda\mathbf{p}) \text{ for all bistochastic matrices } \Lambda. \quad (6.2)$$

Notice that  $\mathbf{q} = \Lambda\mathbf{p}$  if and only if  $\mathbf{p}$  majorises  $\mathbf{q}$ ,  $\mathbf{p} \succ \mathbf{q}$  [44] (recall Definition M3 and Remark M1). Therefore, functions satisfying Eq. (6.2), the Shannon entropy being the best known example, are Schur-concave. We shall denote this set by  $\mathcal{U}_0$ . Hence, the condition given by Eq. (6.1) specifies that a minimal requirement for  $u$  to be a bona-fide uncertainty function is to be Schur-concave. In Ref. [88] no further properties are imposed, i.e, it is assumed that actually any  $u \in \mathcal{U}_0$  can be considered as a meaningful uncertainty function. Thus, within this approach a probability distribution  $\mathbf{p}$  is universally less uncertain than  $\mathbf{q}$  if and only if  $\mathbf{p} \succ \mathbf{q}$ .

### 6.1.2 Requirement of context-independence

We note that not all Schur-concave functions may be appropriate uncertainty measures, as some of them possess potentially undesired properties. In particular, one can show that there exist probability distributions  $\mathbf{p}$ ,  $\mathbf{q}$  and  $\mathbf{r}$  such that

1.  $\exists u \in \mathcal{U}_0 : u(\mathbf{p}) > u(\mathbf{q})$ ,
2.  $\forall u \in \mathcal{U}_0 : u(\mathbf{p} \otimes \mathbf{r}) \leq u(\mathbf{q} \otimes \mathbf{r})$ .

Notice, that this is linked with the phenomenon of catalysis studied within entanglement theory [58]. Therefore, allowing any Schur-concave function to measure uncertainty leads to the existence of a measure  $u$  according to which  $\mathbf{p}$  is more uncertain than  $\mathbf{q}$ , but  $\mathbf{p} \otimes \mathbf{r}$  is less uncertain than  $\mathbf{q} \otimes \mathbf{r}$ .<sup>1</sup> As a result, uncertainty functions are allowed to be *context-dependent*, i.e., an independent random variable  $\mathbf{r}$  can change our assessment of which of two probability distributions,  $\mathbf{p}$  or  $\mathbf{q}$ , is more uncertain. Here we will be interested in uncertainty functions that are *context-independent*, in the sense that independent events do not affect the uncertainty ordering between probability distributions.

In order to remove context-dependence we propose a single and natural additional assumption restricting the set of allowed measures of uncertainty. We require that all

---

<sup>1</sup>As a particular example one can take:  $\mathbf{p} = (0.5, 0.25, 0.25, 0)$ ,  $\mathbf{q} = (0.4, 0.4, 0.1, 0.1)$ ,  $\mathbf{r} = (0.6, 0.4)$  and  $u$  to be the sum of two smallest elements of a probability vector [58].

bona-fide measures of uncertainty should not only be Schur-concave, but also additive:

$$u(\mathbf{p} \otimes \mathbf{q}) = u(\mathbf{p}) + u(\mathbf{q}). \quad (6.3)$$

The above condition reflects the extensiveness of uncertainty for independent events, a standard assumption for information and uncertainty measures [42]. Thus, we define the general family of uncertainty functions by the set of additive Schur-concave functions and denote it by  $\mathcal{U}$ . It is straightforward to check that by getting rid of non-additive functions the problem of context-dependence is solved. Indeed, due to additivity, for any  $u \in \mathcal{U}$  we have

$$u(\mathbf{p}) > u(\mathbf{q}) \Leftrightarrow u(\mathbf{p} \otimes \mathbf{r}) > u(\mathbf{q} \otimes \mathbf{r}) \quad \forall \mathbf{r}.$$

As before we can ask when one random variable is universally less uncertain than another. The answer is that if  $\mathbf{p}$  is not simply a permutation of  $\mathbf{q}$  this is the case if and only if

$$H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q}) \quad \forall \alpha \in \mathbb{R}, \quad (6.4)$$

where  $H_\alpha$  are the Rényi entropies (see Definition M1). The fact that Eq. (6.4) implies the same inequality for all  $u \in \mathcal{U}$  is non-trivial and it is a consequence of the results of Refs. [46, 47] that show the equivalence between Eq. (6.4) and the *trumping* relation  $\succ_T$  (see Definition M4 and Remark M2). A probability distribution  $\mathbf{p}$  is said to trump  $\mathbf{q}$  if there exists a context in which  $\mathbf{p}$  majorises  $\mathbf{q}$ . Therefore, choosing  $\mathcal{U}$  as the set of uncertainty functions, we can alternatively say that  $\mathbf{p}$  is universally less uncertain than  $\mathbf{q}$  if and only if  $\mathbf{p}$  trumps  $\mathbf{q}$ ,  $\mathbf{p} \succ_T \mathbf{q}$ .

### 6.1.3 Requirement of decidability

We will now show that allowing Rényi entropies of non-positive order  $\alpha \leq 0$  to measure uncertainty leaves us with an important problem of undecidability; more precisely, arbitrarily small changes in the probability of events can switch our assessment of which between two probability distributions is more uncertain. Note that this is a physically significant issue, as any physical experiment allows us to determine the probability of events only up to an arbitrarily small, but non-zero error. Hence, using such non-continuous uncertainty functions may lead to a situation in which we need to change our assessment of which of two probability distributions is more uncertain according to an unobservable event.

To illustrate this problem, let us consider the example of two distributions  $\mathbf{p} \otimes \mathbf{r}$  and  $\mathbf{q} \otimes \mathbf{r}$  (both with full support) and fix  $\alpha < 0$ . Then if  $H_\alpha(\mathbf{p}) > H_\alpha(\mathbf{q})$  we have that  $\mathbf{p} \otimes \mathbf{r}$  is *more* uncertain than  $\mathbf{q} \otimes \mathbf{r}$  according to the chosen measure  $H_\alpha$ . However, as we can only know the probabilities of events up to an arbitrarily good approximation, the probability distribution  $\mathbf{r}$  on the right hand side may actually be  $\mathbf{r}^\epsilon$  with  $\|\mathbf{r} - \mathbf{r}^\epsilon\|_1 \leq \epsilon$  and some arbitrarily small  $\epsilon$  (here  $\|\cdot\|_1$  denotes the  $\ell_1$  norm). Then by choosing  $\mathbf{r} = (1, 0)$  and  $\mathbf{r}^\epsilon = (\epsilon, 1 - \epsilon)$  we get that  $\mathbf{p} \otimes \mathbf{r}$  is *less* uncertain than  $\mathbf{q} \otimes \mathbf{r}^\epsilon$ , according to  $H_\alpha$ , for any non-zero  $\epsilon$ , whereas it is more uncertain if  $\epsilon = 0$  exactly. Hence, our assessment of which probability distribution is more uncertain is reversed by an undecidable fact (i.e., if  $\epsilon$  is exactly zero or not).

To overcome the problem of undecidability one can simply require the continuity of uncertainty functions: given any  $\mathbf{p}$ , for all  $\delta$  there should exist  $\epsilon$  such that

$$\|\mathbf{p} - \mathbf{p}^\epsilon\|_1 \leq \epsilon \implies |u(\mathbf{p}) - u(\mathbf{p}^\epsilon)| \leq \delta. \quad (6.5)$$

It is then clear that given  $\mathbf{p}$  and  $\mathbf{q}$ , with  $u(\mathbf{p}) > u(\mathbf{q})$ , also all elements of an  $\epsilon$ -ball around  $\mathbf{p}$  are more uncertain than all elements of an  $\epsilon$ -ball around  $\mathbf{q}$ , for  $\epsilon > 0$  small enough. Note that Rényi entropies of order  $\alpha \leq 0$  and the Burges entropy are not continuous for distributions without full support. Hence, if we decide to exclude measures affected by this problem, we further restrict the set of uncertainty functions to Schur-concave, continuous and additive functions, denoted by  $\mathcal{U}_+$ . As before, a probability distribution  $\mathbf{p}$  is universally less uncertain than  $\mathbf{q}$  if and only if  $H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q})$  for all  $\alpha > 0$ .

Let us now summarise the main message of this section. We have defined three families of bona-fide uncertainty functions (see Fig. 6.1) by means of three natural axioms:<sup>2</sup>

#### Axioms for measures of uncertainty

1. Non-increasing under random relabelling, Eq. (6.1), as introduced in Ref. [88].
2. Additivity for independent random variables.
3. Continuity.

Within  $\mathcal{U}_0$ ,  $\mathbf{p}$  is universally more uncertain than  $\mathbf{q}$  if and only if  $p \succ q$ ; within  $\mathcal{U}$ , if and

<sup>2</sup>It may be also worth exploring the set of Schur-concave and continuous functions. In fact, if any of the three presented conditions may be dropped and still give a physically reasonable framework, this seems to be Axiom 2.

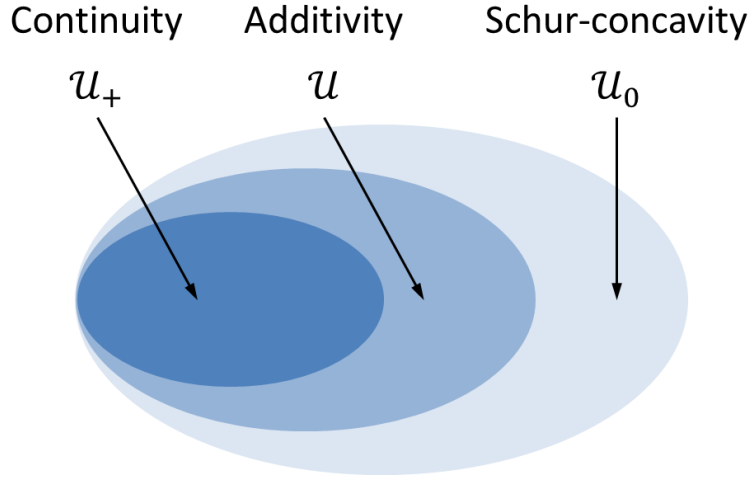


Figure 6.1: *Three families of uncertainty functions.* Schur-concave functions  $\mathcal{U}_0$  satisfy Axiom 1, Rényi entropies of any real order  $\mathcal{U}$  satisfy Axioms 1-2 and, finally, Rényi entropies of positive order  $\mathcal{U}_+$  satisfy Axioms 1-3.

only if  $H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q})$  for all  $\alpha \in [-\infty, \infty]$ ; and within  $\mathcal{U}_+$ , if and only if  $H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q})$  for all  $\alpha \in (0, \infty]$ .

## 6.2 Minimum uncertainty states

Having identified the set of conditions that characterise when one probability distribution is universally more uncertain than another, we now have a general framework to study the uncertainty relations. We can investigate the unavoidable uncertainty of the outcome statistics for two noncommuting observables independently of the chosen uncertainty measure. In particular, we will be interested in answering the following question: are there quantum states that simultaneously minimise uncertainty with respect to all uncertainty measures?

In fact, any uncertainty function  $u$  defines an uncertainty relation for given observables  $A$  and  $B$ :

$$u(\mathbf{p}^A(\rho) \otimes \mathbf{p}^B(\rho)) \geq c_{AB}^u \quad \forall \rho \in \mathcal{S}'_d. \quad (6.6)$$

Here  $\mathcal{S}'_d$  is some subset of the set of  $d$ -dimensional quantum states  $\mathcal{S}_d$  (often  $\mathcal{S}'_d = \mathcal{S}_d$ ),  $\mathbf{p}^A(\rho)$  and  $\mathbf{p}^B(\rho)$  denote the probability distributions over the outcomes of measurements  $A$  and  $B$  on state  $\rho$ , and  $c_{AB}^u > 0$  is a constant that does not depend on  $\rho \in \mathcal{S}'_d$ . As an example, consider the well-known Maassen-Uffink uncertainty relation, see Eq. (5.2),

where  $u$  is chosen to be the Shannon entropy  $H$ ,  $\mathcal{S}'_d = \mathcal{S}_d$ , and

$$c_{AB}^H = -2 \ln \left( \max_{ij} |\langle a_i | b_j \rangle| \right), \quad (6.7)$$

with  $|a_i\rangle$  and  $|b_j\rangle$  denoting eigenstates of  $A$  and  $B$ , respectively.

### 6.2.1 Universal minimum uncertainty states

States minimising the left-hand side of Eq. (6.6) for some choice of  $u$  and  $\mathcal{S}'_d$  will be called minimum uncertainty states (MUSs). These have been found in the case of  $u = H_1$  (the Shannon entropy) [101, 102] and  $u = H_2$  (collision entropy) [104] for  $\mathcal{S}'_d = \mathcal{S}_2$  (qubit systems). However, if we restrict the study of uncertainty relations to a particular uncertainty function, then anything we can say about the structure of MUSs will, in general, not hold for a different measure. Instead, having argued for the general sets of bona-fide uncertainty measures  $\mathcal{U}_0$ ,  $\mathcal{U}$  and  $\mathcal{U}_+$ , we can introduce the notion of *universal* minimum uncertainty state: a state that minimises all  $u \in \mathcal{F}$  simultaneously, with  $\mathcal{F}$  being one of the three sets of uncertainty measures introduced in the previous section.

#### Definition 6.1: Universal MUS

A universal minimum uncertainty state within a subset  $\mathcal{S}'_d \subseteq \mathcal{S}_d$  is a state  $\rho \in \mathcal{S}'_d$  that is universally less uncertain than any other element in  $\mathcal{S}'_d$  (modulo permutations of the outcomes). More precisely,

$$u(\mathbf{p}^A(\rho) \otimes \mathbf{p}^B(\rho)) < u(\mathbf{p}^A(\sigma) \otimes \mathbf{p}^B(\sigma)),$$

for all  $u \in \mathcal{F}$  and for all  $\sigma \in \mathcal{S}'_d$  such that  $\pi[\mathbf{p}^A(\rho) \otimes \mathbf{p}^B(\rho)] \neq \mathbf{p}^A(\sigma) \otimes \mathbf{p}^B(\sigma)$ , with  $\pi$  being an arbitrary permutation.<sup>a</sup>

<sup>a</sup>Note that there are no probability distributions, other than those linked by permutations, that have the same uncertainty with respect to all  $u \in \mathcal{F}$ .

The existence of such special states is conceptually very intriguing. Does quantum mechanics permit their existence for some natural choice of  $\mathcal{S}'_d$ ? An obvious choice that we will consider is the full state space  $\mathcal{S}'_d = \mathcal{S}_d$ . However, we will also focus on another physically motivated subset of states  $\mathcal{S}'_d$  arising while studying the uncertainty relations in the presence of classical noise (where  $\mathcal{S}'_d \subset \mathcal{S}_d$  is chosen to be a subset with a given level of mixedness [73, 86, 105, 106]).

Investigating the existence of a universal MUS, or even some approximate version of it – briefly discussed later in Sec. 6.4.3 – can also be practically relevant. This is because a universal MUS simultaneously minimises all possible uncertainty measures over a considered set of states, and different  $u$  are operationally relevant in different situations. As uncertainty relations have a range of applications in cryptography and quantum information [35], we conjecture that universal a MUS may be useful when we want to perform a protocol, but we do not know in advance what the rules of the probabilistic game are.

In what follows we first provide a general no-go theorem forbidding the existence of a universal MUS within the full state space  $\mathcal{S}_d$ . This shows that the best (least uncertain) state always depends on the details of the probabilistic game being played; no ultimate top element exists. However, in many physically relevant scenarios the only available quantum states are mixed. Hence, in the next section, we will explore features emerging from the interplay between noncommutativity and noise.

## 6.2.2 No-go theorem for a pure universal MUS

In the case of two commuting observables (or more generally observables sharing an eigenstate) the existence of a universal MUS is trivial: any common eigenstate has a sharp distribution with respect to both measurements. However, the problem is non-trivial for observables that do not share an eigenstate. In fact, the following result shows that in this case no top element exists within the full unconstrained state space  $\mathcal{S}_d$  for all three choices of  $\mathcal{F}$ :

### Theorem 6.1: No pure universal MUS

Given observables  $A$  and  $B$  acting on  $d$ -dimensional Hilbert space  $\mathcal{H}_d$  and not sharing any common eigenstate, no universal MUS within the full state space  $\mathcal{S}_d$  exists.

*Proof.* The proof consists of two steps. First, we prove that if there exists a universal MUS  $\rho \in \mathcal{S}_d$  then it must be pure. Next, we find all pure states  $\{|\psi_\infty^m\rangle\}$  that minimise  $H_\infty$  and show that there exist pure states that have smaller  $H_\alpha$  than any of the  $\{|\psi_\infty^m\rangle\}$  for some  $\alpha > 0$ . Therefore we conclude that no state can simultaneously minimise all uncertainty measures  $u \in \mathcal{F}$  over the full state space  $\mathcal{S}_d$ .

**Step 1.** Given any mixed state  $\rho$ , let us decompose it in its own eigenbasis  $\{|\psi_i\rangle\}$ :

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| := \sum_i \lambda_i \rho_i.$$

The strict concavity and additivity of the Shannon entropy  $H$  implies that

$$\begin{aligned} H(\mathbf{p}^A(\rho) \otimes \mathbf{p}^B(\rho)) &> \sum_{i,j} \lambda_i \lambda_j H(\mathbf{p}^A(\rho_i) \otimes \mathbf{p}^B(\rho_j)) = \sum_i \lambda_i H(\mathbf{p}^A(\rho_i) \otimes \mathbf{p}^B(\rho_i)) \\ &\geq \min_i H(\mathbf{p}^A(\rho_i) \otimes \mathbf{p}^B(\rho_i)), \end{aligned}$$

so that for every mixed state  $\rho$  there exists a pure state  $\rho_k$  that is characterised by lower Shannon entropy. This immediately implies that no mixed state can be a universal MUS and hence we can consider only pure states.

**Step 2.** We will now find the set of pure states  $\{|\psi_\infty^m\rangle\}$  that minimise  $H_\infty(\mathbf{p}^A(|\psi\rangle) \otimes \mathbf{p}^B(|\psi\rangle))$  among all pure states  $|\psi\rangle$ . Let us remind that  $H_\infty(\mathbf{p}) = -\ln \max_i p_i$ , so that we are actually looking for states that maximise the largest entry of the probability vector  $\mathbf{p}^A(|\psi\rangle) \otimes \mathbf{p}^B(|\psi\rangle)$ . Let  $V$  be the unitary connecting eigenbases of  $A$  and  $B$ , i.e.,  $|a_i\rangle = V |b_i\rangle$  for all  $i = 1, \dots, d$ . Since, by assumption,  $A$  and  $B$  do not share an eigenstate we have  $\forall i, j \quad |V_{ij}| := |\langle b_i | V | b_j \rangle| < 1$ . Let  $c$  denote the absolute value of the matrix element of  $V$  that has the largest absolute value, i.e.,

$$c := \max_{i,j} |V_{ij}| := |V_{i_M j_M}| < 1, \quad (6.8)$$

where  $(i_M, j_M)$  denotes the indices corresponding to one of such largest elements of  $V$ . Now, let's decompose a general normalised pure state  $|\psi\rangle$  into the eigenstates of  $A$ :

$$|\psi\rangle = \sum_{k=1}^d \alpha_k |a_k\rangle = \sqrt{p} |a_{k_M}\rangle + \sqrt{1-p} |a_{k_M}^\perp\rangle, \quad (6.9)$$

where  $\max_k |\alpha_k| = |\alpha_{k_M}| := \sqrt{p}$ , and we absorbed a phase in the definition of  $|a_{k_M}\rangle$ . Also,

$$|a_{k_M}^\perp\rangle := \frac{1}{\sqrt{1-p}} \sum_{k \neq k_M} \alpha_k |a_k\rangle. \quad (6.10)$$

Let  $p_{\max}(|\psi\rangle)$  denote the maximal element of the joint probability distribution  $\mathbf{p}^A(|\psi\rangle) \otimes \mathbf{p}^B(|\psi\rangle)$ :

$$p_{\max}(|\psi\rangle) := \max_{k,l} p_k^A(|\psi\rangle) p_l^B(|\psi\rangle) = p \max_l p_l^B(|\psi\rangle). \quad (6.11)$$

We have

$$\begin{aligned}
 p_{\max}(|\psi\rangle) &= \max_l p \left| \sqrt{p} \langle b_l | a_{k_M} \rangle + \sqrt{1-p} \langle b_l | a_{k_M}^\perp \rangle \right|^2 \\
 &:= p \left| \sqrt{p} \langle b_{l_M} | a_{k_M} \rangle + \sqrt{1-p} \langle b_{l_M} | a_{k_M}^\perp \rangle \right|^2 \\
 &= p \left| \sqrt{p} |V_{l_M k_M}| + e^{ix} \sqrt{1-p} \sqrt{1 - |V_{l_M k_M}|^2} \right|^2 \\
 &\leq p \left( \sqrt{p} |V_{l_M k_M}| + \sqrt{1-p} \sqrt{1 - |V_{l_M k_M}|^2} \right)^2,
 \end{aligned}$$

where  $l_M$  is the index  $l$  satisfying the first maximisation problem and

$$x := \arg \langle b_{l_M} | a_{k_M} \rangle - \arg \langle b_{l_M} | a_{k_M}^\perp \rangle. \quad (6.12)$$

The inequality is tight only if  $x = 0$  and then we have

$$p_{\max}(|\psi\rangle) \leq \frac{(1 + |V_{l_M k_M}|)^2}{4} \leq \frac{(1+c)^2}{4}, \quad (6.13)$$

where the first inequality is attained for  $p = (1 + |V_{l_M k_M}|)/2$  and the second inequality is attained only if  $l_M = i_M$  and  $k_M = j_M$ . One easily finds that the tightness of all of the above inequalities implies that

$$|b_{i_M}\rangle = e^{i\phi} \left( c |a_{k_M}\rangle + \sqrt{1-c^2} |a_{k_M}^\perp\rangle \right), \quad (6.14)$$

where  $\phi = \arg \langle a_{k_M} | b_{i_M} \rangle$ . We can now solve the above equation for  $|a_{k_M}^\perp\rangle$  and substitute the result to Eq. (6.9). Finally using  $k_M = j_M$  and optimal  $p = (1+c)/2$  one finds that states maximising  $p_{\max}(|\psi\rangle)$  are of the form

$$|\psi_\infty^m\rangle = \frac{|a_{j_M}\rangle + e^{-i\phi} |b_{i_M}\rangle}{\sqrt{2(1+c)}}, \quad (6.15)$$

where  $m$  enumerates all pairs  $(i_M, j_M)$  for which  $|V_{ij}|$  attains maximum. It is also worth noting that states of the above form actually saturate the bound found by Landau and Pollak [107], for the product of maximum outcome probabilities for noncommuting observables (see Eq. (9) of Ref. [34]).

Now, we just need to show that there exists a pure state  $|\psi\rangle$  for which  $H_\alpha(\mathbf{p}^A(|\psi\rangle) \otimes \mathbf{p}^B(|\psi\rangle))$  is smaller than for any of the states  $\{|\psi_\infty^m\rangle\}$  for some  $\alpha > 0$ . This can be proven in the following way. First, define  $\tilde{H}_0(\mathbf{p}) = \ln |\text{supp } \mathbf{p}|$ , where  $|\text{supp } \mathbf{p}|$  denotes the number of non-zero elements of  $\mathbf{p}$ . Then, note that the distribution  $\mathbf{p}^A(|\psi_\infty^m\rangle) \otimes \mathbf{p}^B(|\psi_\infty^m\rangle)$  has a higher  $\tilde{H}_0$  than the probability distribution corresponding to  $|b_{i_M}\rangle$ . Finally, note that  $\lim_{\alpha \rightarrow 0} H_\alpha(\mathbf{p}) = \tilde{H}_0(\mathbf{p})$  and  $H_\alpha$  is continuous in  $\alpha > 0$ , which means that there exists  $\alpha > 0$  such that  $H_\alpha$  is bigger for any of  $\{|\psi_\infty^m\rangle\}$  than for  $|b_{i_M}\rangle$ .  $\square$

Although a universal MUS does not exist within the full state space, it may still appear in many physical scenarios where some degree of noise is unavoidable. Noise may be present due to inevitable imperfections in the experimental apparatus, or because the system under scrutiny is entangled with some other degrees of freedom we do not have access to. Hence, one is left to wonder whether the no-go result we derived is robust to noise and, more generally, what is the effect of noise on the structure of minimum uncertainty states. We will discuss this in the next section, starting from some conceptual remarks about assessing uncertainty in the presence of classical uniform noise.

## 6.3 Noise and uncertainty

### 6.3.1 Role of noise and $H_2$ in the classical case

To build up intuition as to why the introduction of noise can make a difference in assessing uncertainty, it is useful to start with a simple yet suggestive example. Whereas the importance of  $H_0$ ,  $H_1$  and  $H_\infty$  Rényi entropies has been previously stressed [108], here we emphasise the special role played by the collision entropy  $H_2$  in the presence of noise.

#### Example 6.1: The effect of noise on compression rates

Consider two probability distributions

$$\mathbf{p} = (0.77, 0.10, 0.10, 0.03), \quad \mathbf{q} = (0.63, 0.35, 0.01, 0.01),$$

and two sources  $P$  and  $Q$  that produce messages by drawing from a four-element alphabet according to these probability distributions. One can immediately check that  $H_1(\mathbf{p}) > H_1(\mathbf{q})$ , which means that the messages produced by  $Q$  will have a higher compression rate than those produced by  $P$ . However, now assume that there is an additional noise channel that affects the messages produced by sources  $P$  and  $Q$ , so that the effective probability distributions become

$$\mathbf{p}^\epsilon = \epsilon \boldsymbol{\eta} + (1 - \epsilon) \mathbf{p}, \tag{6.16}$$

and similarly for  $\mathbf{q}^\epsilon$ , where  $\boldsymbol{\eta} = (0.25, 0.25, 0.25, 0.25)$  is a uniform distribution. It is then easy to verify that, for  $\epsilon \geq 0.05$ , we have  $H_1(\mathbf{p}^\epsilon) < H_1(\mathbf{q}^\epsilon)$ . According to the Shannon entropy,  $\mathbf{p}$  is more uncertain than  $\mathbf{q}$  (and hence more difficult to compress),

but the situation is reversed once enough noise is introduced. This shows that the noise more strongly affects the information content of the message produced by  $Q$ , as measured by  $H_1$ .

The above example leads us to the following question: under what conditions the information content encoded in a source  $P$  (as measured by a generic  $H_\alpha$ ) is more strongly affected by uniform noise than the information encoded in  $Q$ ? The answer is provided by the following result:

**Observation 6.2: Classical ordering by  $H_2$**

Given two probability distributions  $\mathbf{p}$  and  $\mathbf{q}$ , with  $H_2(\mathbf{p}) \neq H_2(\mathbf{q})$ , the following statements are equivalent for any given  $\alpha \in (-\infty, \infty)$ :

1. There exists  $\epsilon_\alpha \in [0, 1)$  such that

$$H_\alpha(\mathbf{p}^\epsilon) < H_\alpha(\mathbf{q}^\epsilon) \quad \forall \epsilon \geq \epsilon_\alpha.$$

2.  $H_2(\mathbf{p}) < H_2(\mathbf{q})$ .

*Proof.* The proof is based on the fact that, for every  $\alpha$ ,  $H_\alpha(\mathbf{p}^\epsilon)$  has a maximum at  $\epsilon = 1$  (as  $\mathbf{p}^\epsilon$  then corresponds to a uniform distribution), so that the dominating term in the Taylor expansion around this maximum is quadratic in the probabilities  $\{p_i\}$ . Then, for a finite  $\alpha$ , the ordering between  $H_\alpha(\mathbf{p}^\epsilon)$  and  $H_\alpha(\mathbf{q}^\epsilon)$  for  $\epsilon$  close enough to 1 depends solely on the ordering of the  $\alpha = 2$  Rényi entropy. To see this let us first calculate

$$H_\alpha(\mathbf{p}^\epsilon) - H_\alpha(\mathbf{q}^\epsilon) = \frac{\text{sgn}(\alpha)}{1 - \alpha} \ln A, \quad (6.17)$$

where

$$A = \frac{\sum_i [rp_i d + 1 - r]^\alpha}{\sum_i [rq_i d + 1 - r]^\alpha}, \quad (6.18)$$

and  $r := 1 - \epsilon$ . Hence  $H_\alpha(\mathbf{p}^\epsilon) < H_\alpha(\mathbf{q}^\epsilon)$  is equivalent to  $A > 1$  for  $\alpha < 0$  and  $\alpha > 1$ , whereas for  $\alpha \in (0, 1)$  it is equivalent to  $A < 1$ . Expanding around  $r = 0$  one gets

$$\sum_i [rx_i d + 1 - r]^\alpha = d + \frac{\alpha(\alpha - 1)}{2} \left( \sum_i x_i^2 d^2 - d \right) r^2 + O(r^3), \quad (6.19)$$

where  $\{x_i\}$  denotes the entries of either  $\mathbf{p}$  or  $\mathbf{q}$ . Hence, for any given  $\alpha$ , we can rewrite

$H_\alpha(\mathbf{p}^\epsilon) < H_\alpha(\mathbf{q}^\epsilon)$  as

$$\sum_i (p_i^2 - q_i^2) d^2 r^2 + O(r^3) > 0. \quad (6.20)$$

Then, it is clear that if  $\sum_i p_i^2 > \sum_i q_i^2$ , i.e.,  $H_2(\mathbf{p}) < H_2(\mathbf{q})$ , the above inequality is satisfied for  $r_\alpha > 0$  small enough. An analogous proof can be used to show that the statement is valid for  $\alpha = 0$  and  $\alpha = 1$ . On the other hand, for a given  $\alpha$  let  $H_\alpha(\mathbf{p}^\epsilon) < H_\alpha(\mathbf{q}^\epsilon)$  for all  $r \leq r_\alpha$  with  $r_\alpha > 0$ . Then it must be that  $\sum_i p_i^2 \geq \sum_i q_i^2$  and so  $H_2(\mathbf{p}) \leq H_2(\mathbf{q})$ . By assumption however  $H_2(\mathbf{p}) \neq H_2(\mathbf{q})$ , hence  $H_2(\mathbf{p}) < H_2(\mathbf{q})$ .  $\square$

### Remark 6.2

The ordering for the limiting cases of  $\alpha = \pm\infty$ , as well as  $\alpha = 2$ , can never be changed by introducing noise.

To see this, first note that  $H_2(\mathbf{p}) < H_2(\mathbf{q})$  is equivalent to  $\sum_i p_i^2 > \sum_i q_i^2$ . Recall that we denote by  $p_i^\epsilon$  the elements of  $\mathbf{p}^\epsilon = (1 - \epsilon)\mathbf{p} + \epsilon\boldsymbol{\eta}$ . Then, introducing  $r := 1 - \epsilon$ , we have

$$\sum_i (p_i^\epsilon)^2 = \sum_i \left( \frac{1-r}{d} + rp_i \right)^2 = \frac{1-r^2}{d} + r^2 \sum_i p_i^2.$$

It is then immediate to see that  $H_2(\mathbf{p}) < H_2(\mathbf{q}) \Leftrightarrow H_2(\mathbf{p}^\epsilon) < H_2(\mathbf{q}^\epsilon)$  for all  $r \in (0, 1]$ , i.e., for every  $\epsilon \in [0, 1)$ .

Similarly note that  $H_\infty(\mathbf{p}) < H_\infty(\mathbf{q})$  is equivalent to  $\max_i p_i > \max_i q_i$ . We will thus consider

$$\max_i p_i^\epsilon = \max_i \left( \frac{1-r}{d} + rp_i \right) = \frac{1-r}{d} + r \max_i p_i.$$

Again, it is easy to see that  $H_\infty(\mathbf{p}) < H_\infty(\mathbf{q})$  is equivalent to  $H_\infty(\mathbf{p}^\epsilon) < H_\infty(\mathbf{q}^\epsilon)$  for all  $\epsilon \in [0, 1)$ . An analogous reasoning works for  $H_{-\infty}$ . Moreover, from Observation 6.2, no other  $H_\alpha$  has this property.

Observation 6.2 shows that noise can indeed play a crucial role in uncertainty relations: it induces an order within the set of probability distributions, with  $H_2$  playing a leading role.<sup>3</sup> However, the problem is more complicated than one might initially expect. Recall that the crucial question unanswered by Observation 6.2 is whether a “finite” amount of noise is sufficient to induce an ordering between all Rényi entropies; in other words, if there

<sup>3</sup>Note that the Rényi entropy of order two is also distinguished from a geometric point of view: the isoentropic states (with fixed value of  $H_2$ ) form hyperspheres inside the probability simplex [65]. This means that for any two states the one that is closer to the uniform distribution  $\boldsymbol{\eta}$  with respect to the standard Euclidian metric will be described by larger entropy  $H_2$ .

exists an  $\tilde{\epsilon} < 1$  independent of  $\alpha$  such that  $H_\alpha(\mathbf{p}^{\tilde{\epsilon}}) < H_\alpha(\mathbf{q}^{\tilde{\epsilon}})$  for all  $\alpha$ . In fact, given two generic probability distributions, the condition  $H_\alpha(\mathbf{p}) < H_\alpha(\mathbf{q})$  for  $\alpha = 2$  and  $\alpha = \pm\infty$  is necessary, but *not* sufficient, to induce ordering between Rényi entropies for all  $\alpha$ . A counterexample is given by  $\mathbf{p} = (0.37, 0.32, 0.24, 0.07)$  and  $\mathbf{q} = (0.36, 0.35, 0.19, 0.10)$ . A direct calculation shows that  $H_2(\mathbf{p}) < H_2(\mathbf{q})$ , but for any amount of noise  $\epsilon$  we have  $H_\alpha(\mathbf{p}) > H_\alpha(\mathbf{q})$  for  $\alpha = 4/(1 - \epsilon)$ . This implies that for  $\alpha \rightarrow \infty$  the required amount of noise must go to 1, so there is no single  $\tilde{\epsilon} < 1$  that ensures the relation  $H_\alpha(\mathbf{p}^{\tilde{\epsilon}}) < H_\alpha(\mathbf{q}^{\tilde{\epsilon}})$  is satisfied for all  $\alpha$ .

### 6.3.2 Role of noise and $H_2$ in the quantum case

Given the discussion and results above, it is natural to define a quantum analogue of Eq. (6.16) by the set  $\epsilon$ -noisy states  $\mathcal{S}_d^\epsilon$  that can be written in the form

$$\mathcal{S}_d^\epsilon := \{\rho^\epsilon : \rho^\epsilon = \epsilon\mathbb{I}/d + (1 - \epsilon)\rho, \rho \in \mathcal{S}_d\}, \quad (6.21)$$

for a generic state  $\rho$  and a fixed  $\epsilon \in [0, 1]$ . However, by the same reasoning as in Theorem 6.1, i.e., using the strong concavity and additivity of the Shannon entropy, one can show that among all states in  $\mathcal{S}_d^\epsilon$  only the ones for which  $\rho$  is pure can be a universal MUS. This means that considerations concerning universal MUS can be restricted to the set of pseudo-pure states,<sup>4</sup> first introduced in the field of NMR spectroscopy [109]:

#### Definition 6.2: Pseudo-pure states

A state belongs to the subset of  $\epsilon$ -pseudo-pure states if it can be written in the form

$$\rho_\psi^\epsilon = \epsilon\mathbb{I}/d + (1 - \epsilon)|\psi\rangle\langle\psi|, \quad \epsilon \in [0, 1]. \quad (6.22)$$

Note that, unlike in Chapter 5 when the amount of classical noise was quantified by the von Neumann entropy, here we do not conform to any particular measure. Instead, the amount of noise is quantified by the admixture of a maximally mixed state.

We now provide a modified version of Observation 6.2 suited for probability distributions arising from the measurement of two noncommuting observables in the presence of

<sup>4</sup>Note that instead of considering a projective measurement described by projectors  $\{|a_i\rangle\langle a_i|\}$  on pseudo-pure states, one can equivalently consider a *noisy* positive operator valued measure (POVM) with POVM elements  $\{\epsilon\mathbb{I}/d + (1 - \epsilon)|a_i\rangle\langle a_i|\}$ .

noise. Let us define

$$\Delta H_\alpha := H_\alpha(\mathbf{p}^A(\rho^\epsilon) \otimes \mathbf{p}^B(\rho^\epsilon)) - H_\alpha(\mathbf{p}^A(\sigma^\epsilon) \otimes \mathbf{p}^B(\sigma^\epsilon)). \quad (6.23)$$

We then have the following:

**Observation 6.3: Quantum ordering by  $H_2$**

Let  $\rho$  and  $\sigma$  denote any two quantum states and  $A$  and  $B$  any two observables. If  $e^{-H_2(\mathbf{p}^A(\rho))} + e^{-H_2(\mathbf{p}^B(\rho))} \neq e^{-H_2(\mathbf{p}^A(\sigma))} + e^{-H_2(\mathbf{p}^B(\sigma))}$ , the following two conditions are equivalent for any given  $\alpha$ :

1. There exists  $\epsilon_\alpha$ :

$$\Delta H_\alpha < 0, \quad \forall \epsilon \geq \epsilon_\alpha,$$

2.  $e^{-H_2(\mathbf{p}^A(\rho))} + e^{-H_2(\mathbf{p}^B(\rho))} > e^{-H_2(\mathbf{p}^A(\sigma))} + e^{-H_2(\mathbf{p}^B(\sigma))}$ .

*Proof.* Introducing  $r := 1 - \epsilon$  one can compute

$$\mathbf{p}^A(\rho^\epsilon) \otimes \mathbf{p}^B(\rho^\epsilon) = \frac{(1-2r)}{d^2} + 2r\mathbf{P}^{AB}(\rho) + r^2\mathbf{Q}^{AB}(\rho), \quad (6.24)$$

where

$$\mathbf{P}^{AB}(\rho) = \frac{\mathbf{p}^A(\rho) + \mathbf{p}^B(\rho)}{2d}, \quad (6.25a)$$

$$\mathbf{Q}_{ij}^{AB}(\rho) = p_i^A(\rho)p_j^B(\rho) - \frac{p_i^A(\rho) + p_j^B(\rho)}{d} + \frac{1}{d^2}. \quad (6.25b)$$

Notice that the same expression holds for  $\sigma^\epsilon$ . We now proceed as in the proof of Observation 6.2. We have  $\Delta H_\alpha = \text{sgn}(\alpha) \ln B / (1 - \alpha)$  with

$$B = \frac{\sum_{ij} [2rP_{ij}^{AB}(\rho)d^2 + 1 - 2r + r^2d^2Q_{ij}^{AB}(\rho)]^\alpha}{\sum_{ij} [2rP_{ij}^{AB}(\sigma)d^2 + 1 - 2r + r^2d^2Q_{ij}^{AB}(\sigma)]^\alpha} := \frac{g_\alpha(\rho)}{g_\alpha(\sigma)}. \quad (6.26)$$

Hence,  $\Delta H_\alpha \leq 0$  is equivalent to  $B > 1$  for  $\alpha < 0$  and  $\alpha > 1$ , whereas for  $\alpha \in (0, 1)$  it is equivalent to  $B < 1$ . Expanding around  $r = 0$  one gets

$$g_\alpha(\rho) = d^2 + 2\alpha(\alpha - 1) \left( \sum_{ij} P_{ij}^{AB}(\rho)^2 d^4 - d^2 \right) r^2 + O(r^3). \quad (6.27)$$

Therefore, for any  $\alpha$  we can rewrite  $\Delta H_\alpha < 0$  as

$$\sum_{ij} (P_{ij}^{AB}(\rho)^2 - P_{ij}^{AB}(\sigma)^2) r^2 d^4 + O(r^3) > 0. \quad (6.28)$$

Let us fix  $\alpha$ . If Condition 2 holds, then  $\sum_{ij} P_{ij}^{AB}(\rho)^2 > \sum_{ij} P_{ij}^{AB}(\sigma)^2$ . Hence, there exists  $r_\alpha$  small enough such that Eq. (6.28) is satisfied, i.e., Condition 1 holds. On the other hand, if for a given  $\alpha$  Condition 1 holds for all  $\epsilon \geq \epsilon_\alpha$  (i.e., for all  $r \leq r_\alpha$ ) then one must have  $\sum_{ij} P_{ij}^{AB}(\rho)^2 \geq \sum_{ij} P_{ij}^{AB}(\sigma)^2$ , which is equivalent to

$$e^{-H_2(\mathbf{p}^A(\rho))} + e^{-H_2(\mathbf{p}^B(\rho))} \geq e^{-H_2(\mathbf{p}^A(\sigma))} + e^{-H_2(\mathbf{p}^B(\sigma))}.$$

However, by assumption the equality does not hold, so we obtain Condition 2.  $\square$

For any given measure of uncertainty  $H_\alpha$ , Observation 6.3 shows that the knowledge of  $H_2$  is sufficient to answer the following question: “which of two states has outcome statistics of two noncommuting measurements more uncertain in the presence of large enough uniform noise?”. Nevertheless, similarly to the classical case, we have no guarantee that a finite amount of noise will generate an ordering between all Rényi entropies. In the next section we will explore this question.

## 6.4 Existence of a noisy universal MUS

### 6.4.1 General results

Let us start by presenting three general results concerning a noisy universal MUS that are valid for all three choices of  $\mathcal{F}$ . One will give us an explicit candidate for such state; the other two prevent the existence of a universal MUS in a broad set of situations. First, however, we provide a technical lemma that may be of interest independently from the question of finding a noisy universal MUS:

#### Lemma 6.4: $\epsilon$ -pseudopure state minimising $H_\infty$

Given observables  $A$  and  $B$ , the  $\epsilon$ -pseudopure state  $\rho_{\psi_\infty}^\epsilon$  minimising  $H_\infty(\mathbf{p}^A(\cdot) \otimes \mathbf{p}^B(\cdot))$  is given by

$$\rho_{\psi_\infty}^\epsilon = \epsilon \mathbb{I}/d + (1 - \epsilon) |\psi_\infty\rangle\langle\psi_\infty|, \quad (6.29)$$

with

$$|\psi_\infty\rangle \propto |a_i\rangle + e^{-i\phi} |b_j\rangle, \quad (6.30)$$

where  $|a_i\rangle$  and  $|b_j\rangle$  are the eigenstates of  $A$  and  $B$  that maximise  $|\langle a_i|b_j\rangle|$  and  $\phi = \arg\langle a_i|b_j\rangle$ .

*Proof.* The proof follows a route similar to the one used in proving Theorem 6.1. We want to show that among  $\epsilon$ -pseudo pure states  $\rho_\psi^\epsilon$  defined in Eq. (6.22), the ones with  $|\psi\rangle = |\psi_\infty^m\rangle$  for some  $m$  (see Eq. (6.15)) are those minimising the quantity  $H_\infty(\mathbf{p}^A(\cdot) \otimes \mathbf{p}^B(\cdot))$ . Instead of minimising this entropic quantity we can equivalently maximise over all pure states  $|\psi\rangle$  the maximal element  $p_{\max}(\rho_\psi^\epsilon)$  of the probability vector  $\mathbf{p}^A(\rho_\psi^\epsilon) \otimes \mathbf{p}^B(\rho_\psi^\epsilon)$ . We have

$$p_{\max}(\rho_\psi^\epsilon) = \frac{\epsilon^2}{d^2} + \frac{\epsilon(1-\epsilon)}{d} \max_{k,l} (p_k^A(|\psi\rangle) + p_l^B(|\psi\rangle)) + (1-\epsilon)^2 \max_{k,l} (p_k^A(|\psi\rangle)p_l^B(|\psi\rangle)).$$

From the proof of Theorem 6.1 we know that the last term is maximised for  $|\psi\rangle = |\psi_\infty^m\rangle$ . We will now show that the second term is also maximised for the same state and, therefore, the whole expression for  $p_{\max}(\rho_\psi^\epsilon)$  is maximised for this choice of  $|\psi\rangle$ .

To shorten the notation let us introduce

$$s_{\max}(|\psi\rangle) = \max_{k,l} (p_k^A(|\psi\rangle) + p_l^B(|\psi\rangle)). \quad (6.31)$$

Using the same reasoning that led us in the proof of Theorem 6.1 to the bound on  $p_{\max}(|\psi\rangle)$ , we can obtain a bound on  $s_{\max}(|\psi\rangle)$ . More precisely we have

$$s_{\max}(|\psi\rangle) \leq p + \left( \sqrt{p}|V_{l_M k_M}| + \sqrt{1-p}\sqrt{1-|V_{l_M k_M}|^2} \right)^2, \quad (6.32)$$

where we use the same notation as in the proof of Theorem 6.1. It is straightforward to show that the above expression is maximised for  $|V_{l_M k_M}| = c$  and  $p = (1+c)/2$ . Similarly as in the proof of Theorem 6.1, this leads to the conclusion that  $s_{\max}(|\psi\rangle)$  is maximised by the states  $|\psi_\infty^m\rangle$ .  $\square$

### Corollary 6.5: Explicit form of a noisy universal MUS

Given observables  $A$  and  $B$  that do not share a common eigenstate, if there exists a universal MUS in  $\mathcal{S}_d^\epsilon$ , then it must be  $\epsilon$ -pseudo-pure with the pure state  $|\psi_\infty\rangle$ .

We now proceed to the first no-go result concerning mutually unbiased observables:

### Theorem 6.6: No-go for mutually unbiased observables

Given observables  $A$  and  $B$  that are mutually unbiased, no universal MUS exists within  $\mathcal{S}_d^\epsilon$ , for any  $\epsilon \in [0, 1)$ .

*Proof.* Note that the Shannon entropy  $H_1(\mathbf{p}^A(\rho) \otimes \mathbf{p}^B(\rho))$  is minimised among  $\epsilon$ -noisy states by  $\epsilon$ -pseudo-pure state given in Eq (6.22) with  $|\psi\rangle$  being an eigenstate of either  $A$

or  $B$ . This is because such states saturate the tight bound  $\ln d + S(\rho)$  found in Sec. 5.3 of Chapter 5 [see Eq. (5.23)] for mixed states in the case of mutually unbiased observables. Then, by direct calculation, one can check that for an  $\epsilon$ -pseudo-pure state defined in Lemma 6.4, the Shannon entropy is higher. Hence such a state cannot be a universal MUS and so, from Corollary 6.5, no universal MUS exists.  $\square$

For higher dimensional systems,  $d \geq 3$ , we can provide a general no-go result severely limiting the existence of a universal MUS in the presence of noise for the choices  $\mathcal{F} = \mathcal{U}_0$  and  $\mathcal{F} = \mathcal{U}$ . Specifically, we have:

**Theorem 6.7: No-go for  $d \geq 3$**

Consider two observables  $A$  and  $B$  with eigenstates  $\{|a_i\rangle\}$  and  $\{|b_j\rangle\}$  such that  $V_{ij} = \langle a_i | b_j \rangle \neq 0$ . If the dimension of the system  $d \geq 3$  then the introduction of noise does not lead to the emergence of a universal MUS for  $\mathcal{F} = \mathcal{U}_0$  and  $\mathcal{F} = \mathcal{U}$ .

*Proof.* From Corollary 6.5 we know that a noisy universal MUS must be of the form specified by Eq. (6.29), so that

$$H_{-\infty}(\mathbf{p}^A(\rho_{\psi_\infty}^\epsilon) \otimes \mathbf{p}^B(\rho_{\psi_\infty}^\epsilon)) > 2 \ln \epsilon / d, \quad (6.33)$$

the inequality being strict because  $V_{ij} \neq 0$ . Now consider the state  $\rho_\xi^\epsilon = \epsilon \mathbb{I} / d + (1 - \epsilon) |\xi\rangle\langle\xi|$ , where  $|\xi\rangle = a_1 |a_1\rangle + a_2 |a_2\rangle$  and  $a_1, a_2$  are chosen such that  $|\xi\rangle$  is orthogonal to  $|b_1\rangle$ . Then computing the left hand side of Eq. (6.33) for  $\rho_\xi^\epsilon$  gives exactly  $2 \ln(\epsilon/d)$ . Hence, the state minimising  $H_{-\infty}$  does not coincide with  $\rho_{\psi_\infty}^\epsilon$ . From Corollary 6.5 it implies that no universal MUS exists for  $\mathcal{F} = \mathcal{U}_0$  and  $\mathcal{F} = \mathcal{U}$ .  $\square$

### 6.4.2 Universal MUS for qubit systems

The general results we have presented cannot be applied to the scenario with two distinct qubit observables when they are not mutually unbiased. In fact, as it often happens, qubits are special and thus require separate treatment. As all considered uncertainty measures depend only on the eigenstates of the observables and not on their eigenvalues, without loss of generality we can choose qubit observables as in Chapter 5. Hence, we choose  $A = \mathbf{a} \cdot \boldsymbol{\sigma}$  and  $B = \mathbf{b} \cdot \boldsymbol{\sigma}$ , where  $\boldsymbol{\sigma}$  denotes the vector of Pauli operators, while  $\mathbf{a}$

and  $\mathbf{b}$  are the Bloch vectors. Let us also denote the angle between the two Bloch vectors by  $\gamma$ , so that  $\mathbf{a} \cdot \mathbf{b} = \cos \gamma$ .

The detailed analysis and calculations can be found in Appendix D, and here we will only state the main results. First of all, for the choice  $\mathcal{F} = \mathcal{U}_0$  (the framework of majorisation uncertainty relations) there exists no universal MUS in the presence of noise, i.e., no amount of noise  $\epsilon < 1$  can lead to the emergence of such state. However, for the two other choices of the family of uncertainty functions ( $\mathcal{F} = \mathcal{U}$  and  $\mathcal{F} = \mathcal{U}_+$ ) a universal MUS may emerge after introducing a threshold amount of noise. Specifically, in Appendix D we prove that for  $\gamma = \pi/4$  the amount of noise  $\epsilon = 1/2$  leads to the emergence of a universal MUS.

In panel (a) of Fig. 6.2 we illustrate on the Bloch sphere this emergence of a universal MUS with the introduction of noise. We plot the states that for a given purity minimise different  $\alpha$ -Rényi entropies. As can be seen in the figure, above a threshold level of noise all the Rényi entropies are minimised by a state described by a Bloch vector lying on the bisection of the angle  $\gamma$ . Note that, according to a numerical investigation, the level of noise  $\epsilon = 1/2$  used in the proof is actually much larger than required. In panel (b) of Fig. 6.2 we similarly plot the position of MUSs for different  $\alpha$ , but in the case of qubit observables separated by  $\gamma = 5\pi/12$ . Notice that now the amount of noise required for a universal MUS to appear is larger and, from Theorem 6.6, one can expect that it grows with  $\gamma$  up to the point when  $\epsilon = 1$  for  $\gamma = \pi/2$  (corresponding to mutually unbiased bases). A numerical investigation supports the conjecture that a universal MUS exists for generic qubit observables.

### 6.4.3 Approximate notion of universality for higher dimensions

For dimensions  $d \geq 3$  the no-go theorems presented so far do not apply to the choice  $\mathcal{F} = \mathcal{U}_+$  (corresponding to uncertainty measures that are both continuous and context-independent) when the considered observables are not mutually unbiased. As in higher dimensions the analytical verification of the emergence of an  $\epsilon$ -noisy universal MUS becomes extremely complicated, we numerically verify whether a universal MUS could emerge with the introduction of noise. The following lemma provides an additional necessary condition for such emergence that can be checked numerically by only investigating pure states:

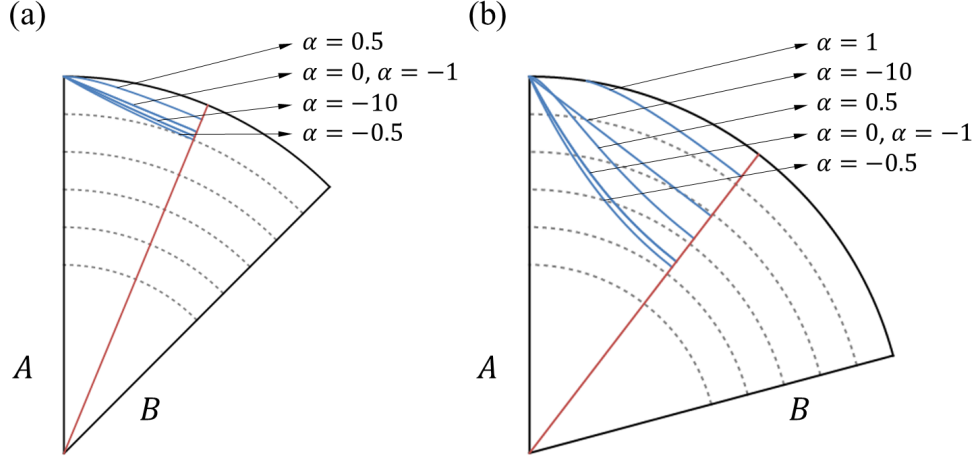


Figure 6.2: *Emergence of a universal MUS for qubits.* The blue curves described by arrows show the position of MUSs with respect  $\alpha$ -Rényi entropies of different orders, as a function of the purity (distance from the origin) for qubit observables  $A$  and  $B$ .  $A$  and  $B$  are separated by angle  $\gamma = \pi/4$  [panel (a)] and  $\gamma = 5\pi/12$  [panel (b)], and the plot presents the Bloch sphere in the first quadrant of the plane spanned by eigenvectors of  $A$  and  $B$ . The bisecting line (in red) denotes the set of states approached by all MUSs and, ultimately, containing the universal MUS. For simplicity we only plot MUSs for  $\theta \in [0, \gamma/2]$ , as the case  $\theta \in [\gamma/2, \gamma]$  is symmetric. Note that MUSs with respect to  $\alpha \geq 1$  [panel (a)] and  $\alpha \geq 2$  [panel (b)] lie on the bisecting line for all purities. Dashed grey lines correspond to states with fixed purity (length of the Bloch vector equal to  $1 - n/10$  for  $n \in \{1, \dots, 5\}$ ).

### Lemma 6.8: Necessary condition for a noisy universal MUS

A necessary condition for a universal MUS to emerge with the introduction of noise is that the expression

$$e^{-H_2(\mathbf{p}^A(\rho))} + e^{-H_2(\mathbf{p}^B(\rho))} \quad (6.34)$$

is maximised among pure states by  $\rho = |\psi_\infty\rangle\langle\psi_\infty|$  (for the definition of  $|\psi_\infty\rangle$  see Lemma 6.4).

*Proof.* By contradiction, suppose that there exists a pure state  $\phi \equiv |\phi\rangle\langle\phi|$  such that

$$e^{-H_2(\mathbf{p}^A(\psi_\infty))} + e^{-H_2(\mathbf{p}^B(\psi_\infty))} < e^{-H_2(\mathbf{p}^A(\phi))} + e^{-H_2(\mathbf{p}^B(\phi))},$$

where  $\psi_\infty \equiv |\psi_\infty\rangle\langle\psi_\infty|$ . Let  $\rho_{\psi_\infty}^\epsilon = \epsilon\mathbb{I}/d + (1 - \epsilon)\psi_\infty$  and  $\rho_\phi^\epsilon = \epsilon\mathbb{I}/d + (1 - \epsilon)\phi$  be the

corresponding  $\epsilon$ -pseudo-pure states and fix  $\alpha = 1$ . Then by Observation 6.3 there exists  $\tilde{\epsilon}$  such that for all  $\epsilon \geq \tilde{\epsilon}$ , one has  $\Delta H_1 > 0$ . This implies that, for any of the three choices of  $\mathcal{F}$ , there is no  $\epsilon$  close enough to 1 such that  $\rho_{\psi_\infty}^\epsilon$  is a universal MUS in  $B_d^\epsilon$ . However, by Corollary 6.5, this implies that a universal MUS does not exist within  $B_d^\epsilon$  for any amount  $\epsilon$  of noise introduced.  $\square$

Given two observables  $A$  and  $B$  we can now use the above result to numerically verify whether the universal MUS emerges with the introduction of noise. This can be done in the following way. First, we need to find  $|\psi_\infty\rangle$ . Then, using a numerical optimisation procedure we search for a state  $|\psi_{\text{opt}}\rangle$  that minimises Eq. (6.34). Finally we can compare  $|\psi_\infty\rangle$  with  $|\psi_{\text{opt}}\rangle$  and if these states differ we can conclude that no universal MUS exists for  $A$  and  $B$ , even for the choice  $\mathcal{F} = \mathcal{U}_+$ .

We numerically investigate  $d \in \{3, 4, 5\}$ , each time generating 1000 pairs of observables  $(A, B)$  whose eigenvectors are connected by a unitary, randomly chosen according to the Haar measure. Our analysis shows that  $|\psi_\infty\rangle$  does not coincide with  $|\psi_{\text{opt}}\rangle$ , showing that in general no universal MUS exists in higher dimension even with the choice  $\mathcal{F} = \mathcal{U}_+$ .<sup>5</sup> However, we also observe that the two states are very close. More precisely, we numerically found that their average overlap  $|\langle\psi_{\text{opt}}|\psi_\infty\rangle|$  is approximately equal to 0.9996, 0.9904 and 0.9842 for dimension  $d$  equal to 3, 4 and 5, respectively. From Observation 6.3 we know that for any given  $\alpha \in (0, +\infty)$ ,  $\rho_{\text{opt}}^\epsilon = \epsilon\mathbb{I}/d + (1 - \epsilon)|\psi_{\text{opt}}\rangle\langle\psi_{\text{opt}}|$  has smaller  $H_\alpha(p^A \otimes p^B)$  than any other given pseudo-pure state if  $\epsilon$  is taken to be bigger than some  $\epsilon_\alpha < 1$ . So, for any arbitrarily fine sample of  $\alpha$ 's and pseudo-pure states, there would be some  $\tilde{\epsilon}$  large enough such that  $\rho_{\text{opt}}^{\tilde{\epsilon}}$  is the best pseudo-pure state. At the same time, the case  $\alpha = \infty$  is optimised by the pseudo-pure state  $\rho_{\psi_\infty}^\epsilon$  that, as we said above, has pure component with large overlap with  $|\psi_{\text{opt}}\rangle$ . This leads naturally to the conjecture that noise can lead to the emergence of a universal MUS in an approximate sense. We leave this as an interesting open question for future work.

In Fig. 6.3 we present an example of the emergence of such approximate universal MUS for a qutrit system. We choose the observables  $A$  and  $B$  such that the eigenstates of  $B$  are connected to the eigenstates of  $A$  by a rotation around  $(1, 1, 1)$  axis by angle  $\pi/6$ . As can be seen in panel (a) of Fig. 6.3, without noise the candidate universal MUS  $\rho_{\psi_\infty}^\epsilon$

<sup>5</sup>Even more strongly, no universal MUS exist whenever  $\mathcal{F}$  contains both  $H_\infty$  and any  $H_\alpha$  for finite  $\alpha$ , e.g.,  $\mathcal{F} = \{H_1, H_\infty\}$ .

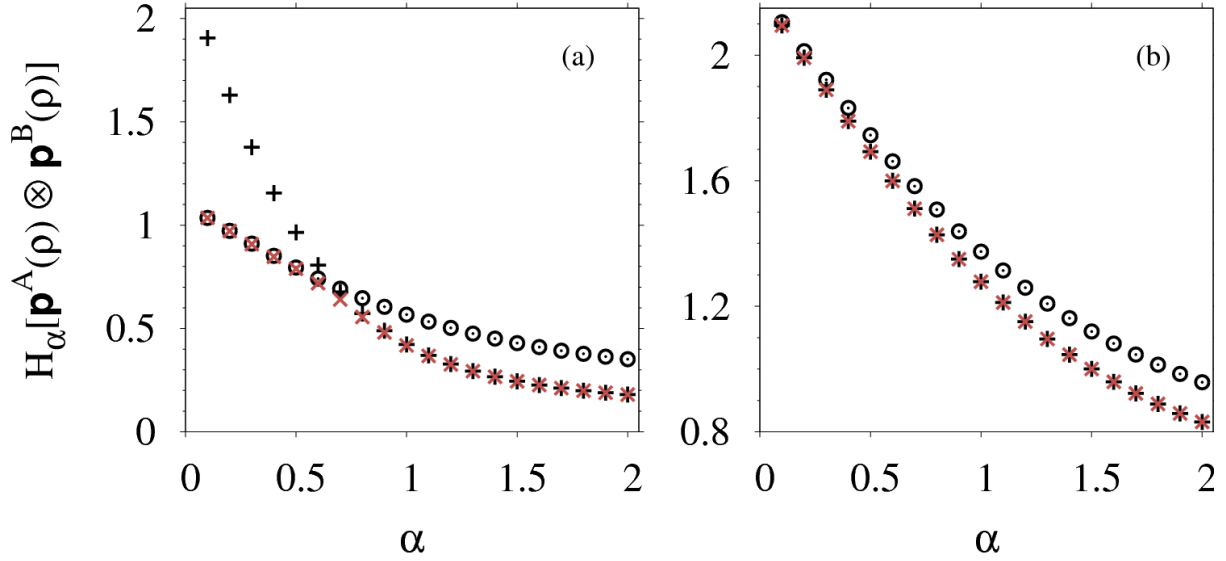


Figure 6.3: *Emergence of an approximate universal MUS for qutrits.* Rényi entropies for  $\rho_\infty^\epsilon$  (black pluses), the  $\epsilon$ -noisy eigenstate of  $A$  (black circles) and numerically optimised  $\epsilon$ -noisy states minimising a given  $\alpha$ -Rényi entropy (red X's), for the qutrit system and observables  $A$  and  $B$  as described in the main text. (a) Level of noise  $\epsilon = 0$ ; (b) Level of noise  $\epsilon = 0.25$  (note the change of vertical scale).

has larger Rényi entropies of order  $\alpha < 1$  than the optimal state. For example, for zero noise the optimal states for  $H_{0,1}$  are close to the eigenstate of either  $A$  or  $B$ . However, the introduction of noise  $\epsilon = 0.25$  results in the approximate equality (discrepancy on the order of  $10^{-4}$  at worst) between Rényi entropies of the candidate and optimal state for the investigated region of  $\alpha \in [0, 2]$ . Numerical investigations also show that this approximate equality holds for  $\alpha > 2$ .

## 6.5 Thermodynamic considerations

In this final section of Chapter 6, we will exploit the links between measures of uncertainty in information theory and measures of the departure from equilibrium in thermodynamics, to get some insight into the thermodynamic ordering of classical states discussed in Chapter 1. Namely, we will extend our observation on the role of collision entropy to the field of thermodynamics and point out some interesting properties of thermalisation.

Let us recall (see Sec. 1.2.3 of Chapter 1) that given a system described by a free Hamiltonian  $H$ , and in contact with a heat bath at inverse temperature  $\beta$ , the free ther-

modynamic transformations between classical states (recall Definition 1.6) are governed by the  $\alpha$ -Rényi divergences  $S_\alpha(\cdot||\cdot)$ , see Eq. (1.21). More precisely, a classical state described by a probability distribution  $\mathbf{p}$  can be freely transformed to another state  $\mathbf{q}$ , if and only if  $S_\alpha(\mathbf{p}||\gamma) \geq S_\alpha(\mathbf{q}||\gamma)$  for all  $\alpha$ , where  $\gamma$  is the equilibrium Gibbs state of the system. Hence, each  $S_\alpha(\mathbf{p}||\gamma)$  is a thermodynamic monotone and can be used as a measure of non-equilibrium.

A similar reasoning to the one showing that  $H_\alpha(\cdot)$  constitute a canonical family of uncertainty functions satisfying two simple and natural axioms can be applied to the quantification of non-equilibrium, with  $S_\alpha(\cdot||\gamma)$  playing the role of  $H_\alpha(\cdot)$ . In order to introduce the thermodynamic analogue of the Axioms 1 and 2 from Sec. 6.1, let us denote by  $\Lambda_{\text{GP}}$  any Gibbs-preserving map, i.e., any stochastic map satisfying  $\Lambda_{\text{GP}}(\gamma) = \gamma$ . As discussed in Part I, Gibbs-preserving maps are the most general set of transformations between classical states that can be performed without using work [61, 110]. In fact, any operation outside this set brings an initially thermal state out of equilibrium, which would allow for building a perpetuum mobile of the second kind by extracting work from a single heat bath, thus violating the second law of thermodynamics. We then require all functions  $f$  that quantify departure of the system from thermodynamic equilibrium to satisfy the following two axioms:

#### Axioms for measures of the departure from equilibrium

1. Non-increasing under Gibbs-preserving maps:  $f(\Lambda_{\text{GP}}(\mathbf{p})) \leq f(\mathbf{p})$ .
2. Additivity for independent systems:  $f(\mathbf{p} \otimes \mathbf{q}) = f(\mathbf{p}) + f(\mathbf{q})$ .

As we have shown in Sec. 1.2.2 of Chapter 1, using the reasoning analogous to the one presented in Sec. 6.1.1 together with the embedding map (see Definition 1.7), one can prove that the first of the above axioms implies that  $f$  must respect the ordering induced by a thermodynamic generalisation of the notion of majorisation – thermo-majorisation. The second requirement then leads us to a thermodynamic analogue of the notion of trumping, which in Ref. [8] was proven to be characterised exactly by the  $\alpha$ -Rényi divergences  $S_\alpha(\cdot||\gamma)$ . Hence  $S_\alpha(\cdot||\gamma)$  play the same canonical role in quantifying the departure from equilibrium as the Rényi entropies in the case of measuring uncertainty. Finally, notice that by requiring continuity we would restrict to Rényi divergences of positive order.

We will now translate Observation 6.2 into the language of thermodynamics and analyse the consequences for near-equilibrium processes. In order to do this let us introduce the set of classical  $\epsilon$ -thermal states (compare with the set of  $\epsilon$ -noisy states  $\mathcal{S}_d^\epsilon$ ):

$$\mathcal{T}_d^\epsilon := \{\mathbf{q}^\epsilon = \epsilon\boldsymbol{\gamma} + (1 - \epsilon)\mathbf{p}, \mathbf{p} \in \mathcal{P}_d\}. \quad (6.35)$$

Notice that Eq. (6.35) describes classical states that are the outcome of an elementary model of thermalisation [111]. We then have the following result:

**Observation 6.9: Thermal ordering by  $S_2(\cdot||\boldsymbol{\gamma})$**

Consider two classical states described by probability vectors  $\mathbf{p}$  and  $\mathbf{q}$ , such that  $S_2(\mathbf{p}||\boldsymbol{\gamma}) \neq S_2(\mathbf{q}||\boldsymbol{\gamma})$ . Then the following statements are equivalent:

1. For every  $\alpha \in \mathbb{R}$  there exists  $\epsilon_\alpha \in [0, 1)$  such that

$$S_\alpha(\mathbf{p}^\epsilon||\boldsymbol{\gamma}) > S_\alpha(\mathbf{q}^\epsilon||\boldsymbol{\gamma}) \quad \forall \epsilon \geq \epsilon_\alpha.$$

2.  $S_2(\mathbf{p}||\boldsymbol{\gamma}) > S_2(\mathbf{q}||\boldsymbol{\gamma})$ .

The proof is a trivial generalisation of Observation 6.2 and it is hence omitted.

Observation 6.9 provides operational meaning to a so-far neglected thermodynamic quantity: the  $\alpha = 2$  Rényi divergence defined by

$$S_2(\mathbf{p}||\boldsymbol{\gamma}) = \ln \sum_i \frac{p_i^2}{\gamma_i}. \quad (6.36)$$

Notice that  $S_2(\cdot||\boldsymbol{\gamma})$  is linked to the thermal average of  $(p_i/\gamma_i)^2$ . Given any valid measure of non-equilibrium, for  $\epsilon$  large enough  $S_2(\cdot||\boldsymbol{\gamma})$  determines which of two the states,  $\mathbf{p}^\epsilon$  or  $\mathbf{q}^\epsilon$  (partially thermalised versions of states  $\mathbf{p}$  and  $\mathbf{q}$ ), is farther from thermal equilibrium. Hence,  $S_2(\cdot||\boldsymbol{\gamma})$  provides an ordering between different near-equilibrium states.

As an application of Observation 6.9, consider a system in state  $\mathbf{p}$  that we want to transform into a target state  $\mathbf{q}$  by putting it in contact with a heat bath. Let us assume that Rényi divergences satisfy  $S_1(\mathbf{p}||\boldsymbol{\gamma}) < S_1(\mathbf{q}||\boldsymbol{\gamma})$ , so that such transformation is forbidden by the second law of thermodynamics. However, it is still possible that  $S_2(\mathbf{p}||\boldsymbol{\gamma}) > S_2(\mathbf{q}||\boldsymbol{\gamma})$ . If this is the case then close enough to equilibrium we have  $S_1(\mathbf{p}^\epsilon||\boldsymbol{\gamma}) > S_1(\mathbf{q}^\epsilon||\boldsymbol{\gamma})$ . This implies that we can transform many copies of  $\mathbf{p}^\epsilon$  into many copies of  $\mathbf{q}^\epsilon$  [6]. It also means that we can extract on average a positive amount of work

by transforming  $\mathbf{p}^\epsilon$  into  $\mathbf{q}^\epsilon$ , even though work is required to transform  $\mathbf{p}$  to  $\mathbf{q}$ . Moreover, as shown in Ref. [112], for  $\epsilon$  large enough we can transform  $\mathbf{p}^\epsilon$  into  $\mathbf{q}^\epsilon$  by thermal operations using a source of stochastic independence. Finally, note that taking into account the thermalisation interpretation of Eq. (6.35), the reversal of free energy  $S_1(\cdot||\gamma)$  ordering between states  $\mathbf{p}$  and  $\mathbf{q}$  can arise from a thermalisation process. For example, if initial states  $\mathbf{p}(0)$  and  $\mathbf{q}(0)$ , with  $S_2(\mathbf{p}(0)||\gamma) > S_2(\mathbf{q}(0)||\gamma)$ , thermalise according to Eq. (6.35) at the same rate  $\epsilon(t)$  [ $\epsilon(t)$  monotonically increases with  $t$ ], then for all times  $t > t_r$  for some finite  $t_r$  we will have  $S_1(\mathbf{p}(t)||\gamma) > S_1(\mathbf{q}(t)||\gamma)$ .

# Chapter 7

## Error-disturbance relations

The best known modern formulation of Heisenberg uncertainty relation (also known as the Heisenberg-Robertson UR [28]) concerns the outcome statistics of two independent measurements of noncommuting observables performed on an ensemble of identically prepared quantum states. As explained in Chapter 5, it states that the product of variances of these two outcome statistics is lower-bounded by the mean value of the commutator of measured observables in a given quantum state. Although this formulation says nothing about the effect of one measurement on the outcome statistics of the other, it is often misinterpreted in the spirit of the original Heisenberg microscope thought experiment [29], i.e., that the bigger the precision of the measurement of one observable, the bigger the disturbance to a subsequent measurement of the other one (with which it does not commute). Indeed, the formulation of uncertainty relation in terms of precision and disturbance of sequential measurements was Heisenberg's original concept and is often referred to as the *error-disturbance uncertainty relation* (see Fig. 7.1). The problem has its origin in the fact that the act of measuring observable  $A$  collapses the measured system into one of the eigenstates of  $A$ . Thus, after discarding the measurement outcome, the uncertainty about observable  $A$  is purely classical, i.e., a non-selective measurement of  $A$  can be seen as transforming quantum uncertainty about  $A$  into classical uncertainty. However, this collapse into a mixture of eigenstates of  $A$  may disturb the information (affect the uncertainty) one has about another observable  $B$  that is not commuting with  $A$ . In the extreme case the measurement of  $A$  may even change the complete certainty about observable  $B$  (when the system is prepared in one of the eigenstates of  $B$ ) into complete uncertainty (described by the maximally mixed state). Interestingly, despite almost a century of re-

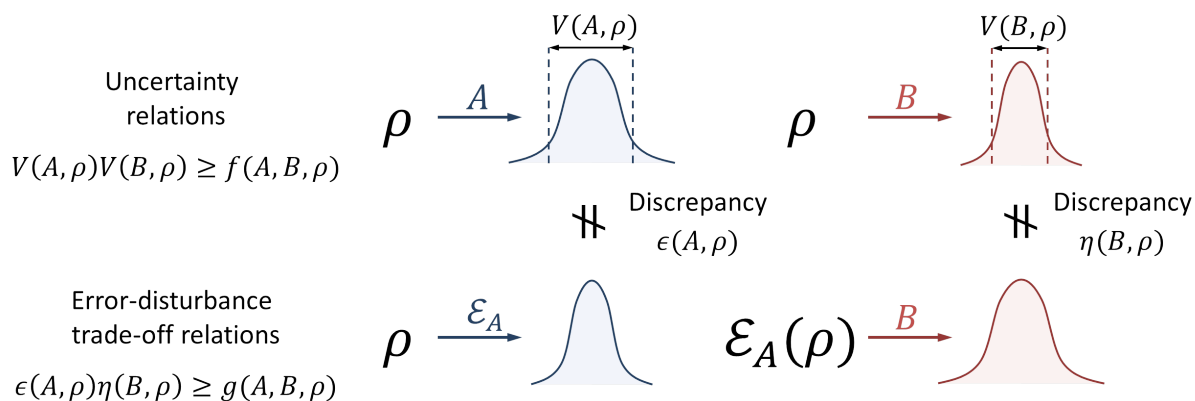


Figure 7.1: *Uncertainty relations vs error-disturbance trade-off relations.* The role of the uncertainty relations is to find a bound  $f(A, B, \rho)$  on the spread of the outcomes (measured, e.g., by the variance) of two noncommuting measurements,  $A$  and  $B$ , performed on independent systems prepared in the same state  $\rho$ . In contrast, the scenarios studied by error-disturbance relations involve both measurements sequentially performed on the same system. The first measurement  $\mathcal{E}_A$  approximates the measurement of  $A$  with an error  $\epsilon(A, \rho)$ , and modifies the state of the system, so that the statistics of the subsequent measurement of  $B$  will be affected by disturbance  $\eta(B, \rho)$ . One is then interested in the trade-off between  $\epsilon(A, \rho)$  and  $\eta(B, \rho)$ , looking for a bound  $g(A, B, \rho)$ .

search on quantum theory, in the last few years the error-disturbance trade-off relations have become again the topic of much discussion [113–123].

Most of the controversies around the error-disturbance relation arose due to disagreement about proper definitions of error (noise) and disturbance. Let us recall that the original Heisenberg argument suggests that for a system prepared in any state  $\rho$  one has  $\epsilon(X, \rho)\eta(P, \rho) \sim h/2$ , where  $\epsilon(X, \rho)$  denotes the error of the approximate position measurement and  $\eta(P, \rho)$  is the disturbance to the subsequent measurement of momentum. It was suggested by Ozawa [113] that more generally a Robertson-like relation  $\epsilon(A, \rho)\eta(B, \rho) \geq |\text{Tr}(\rho[A, B])|/2$  should bound the error and disturbance of sequential measurements of arbitrary operators  $A$  and  $B$  performed in a given state  $\rho$ . This gave rise to state-dependent trade-off relations that use the expectation value of the commutator in a given state to bound some function of error and disturbance for that state. [113, 118, 124–128].

In this chapter we critically review such approaches by examining the consequences of the accepted and operationally motivated requirement that all physically meaningful

notions of disturbance should satisfy [114, 121]. Specifically, we focus on operationally detectable disturbances for which it is natural to define the following:

**Definition 7.1: Operational disturbance**

Consider a non-selective measurement of an observable  $A$  on a system in state  $\rho$  that is described by a CPTP map  $\mathcal{E}_A$ , so that the resulting final state is given by  $\mathcal{E}_A(\rho)$ . We say the measurement of  $A$ , given  $\rho$ , is operationally disturbing to a subsequent measurement of  $B$  if and only if the statistics of  $B$  differ for  $\rho$  and  $\mathcal{E}_A(\rho)$ .

Moreover, we require that any measure of disturbance should assign the value 0 to operationally non-disturbing measurements. This is the central *operational requirement* (OR) around which we will build results in this chapter. Although it is clearly an uncontroversial demand, we will show that OR rules out a broad class of “natural” error-disturbance relations, including recent prominent examples in the literature [118, 124, 127] that fail to adhere to this basic requirement.

To show this, we shall prove that for any finite dimensional quantum system, and any two noncommuting observables  $A$  and  $B$  there always exist pure states  $\{|\psi_{AB}^*\rangle\}$ , such that a perfect (projective and sharp) measurement of  $A$  can be performed (so there is no error in the statistics of  $A$ ) and the disturbance (in the subsequent statistics of  $B$ ) vanishes. Moreover, we show that the expectation value of the commutator  $[A, B]$  for such a state  $|\psi_{AB}^*\rangle$  generically does not vanish. Therefore, any state-dependent error-disturbance relation for sequential measurements of  $A$  and  $B$  that uses the expectation value of  $[A, B]$  in a given state as a lower bound must violate the operational requirement. In other words, the measures of disturbance used in all such relations must take non-zero values even in the situations, when the measurement statistics have not been changed, which is an unphysical conclusion. We will illustrate this explicitly by analysing the “universally valid error-disturbance uncertainty relation” obtained by Ozawa [113].

## 7.1 General framework

### 7.1.1 State-dependent notions

A state-dependent approach to error and disturbance is based on the following scenario. Given an initial quantum state  $\rho$  of the system we look for an approximate POVM mea-

surement  $\mathcal{E}_A$  that tries to reproduce the perfect measurement of an observable  $A$  on this state. The failure of reproducing the proper statistics is quantified by error  $\epsilon(A, \rho)$ . At the same time a given measurement  $\mathcal{E}_A$  disturbs the system and leaves it in the final state  $\mathcal{E}_A(\rho)$ . Hence, the subsequent measurement of the observable  $B$  is performed on  $\mathcal{E}_A(\rho)$  and the outcome statistics may differ from the one obtained from  $\rho$ . This disagreement between the two statistics is quantified by the disturbance  $\eta(B, \rho)$ . We see that any state-dependent measures of error,  $\epsilon(A, \rho)$ , and disturbance,  $\eta(B, \rho)$ , depend on three aspects: the approximate measurement  $\mathcal{E}_A$  used, the observables to be measured ( $A$  and  $B$ ) and the initial state of the system  $\rho$ . This is a broad setting that also may be used to define state-independent notions of error and disturbance. This can be done, e.g., through averaging  $\epsilon(A, \rho)$  and  $\eta(B, \rho)$  over all possible initial states or by finding the maximum and minimum values of  $\epsilon(A, \rho)$  and  $\eta(\rho, B)$  over the full set of states [120]. Additionally, we can utilise the prior knowledge of the measured state. As an example consider a qubit system prepared in an unknown state. On average, the projective measurement  $\sigma_z$  will then disturb the subsequent measurement of  $\sigma_x$ . However, if we know that the investigated system is prepared in one of the two eigenstates of the measured  $\sigma_z$  operator, then clearly such a measurement does not change the system state. Hence, the subsequent measurement of  $\sigma_x$  will not be disturbed.

The goal of state-dependent trade-off relations between error and disturbance is to put a bound on some function of  $\epsilon(A, \rho)$  and  $\eta(B, \rho)$ , that holds for all approximate measurements  $\mathcal{E}_A$  performed on a system prepared in a given state  $\rho$ . One of such relations is the modified Heisenberg noise-disturbance uncertainty relation proposed by Ozawa [113]:

$$\epsilon(A, \rho)\eta(B, \rho) \geq \frac{|\text{Tr}(\rho[A, B])|}{2}. \quad (7.1)$$

We will refer to the above as to the restricted Ozawa's relation, as it is a special case of his "universally valid error-disturbance uncertainty relation" [113], which will be discussed in more detail in Sec. 7.3 [where we will also provide the explicit definitions of  $\epsilon$  and  $\eta$  used to derive Eq. (7.1)]. Here we use its simplified form to emphasise the main component present in all such relations. Namely, Eq. (7.1) states that an approximate measurement on  $\rho$  may reproduce the ideal projective measurement of observable  $A$  on this state with precision limited by noise  $\epsilon(A, \rho)$ , only if it also produces a disturbance  $\eta(B, \rho)$  of the subsequent projective measurement of  $B$ , such that the product of  $\epsilon(A, \rho)$  and  $\eta(B, \rho)$  is

lower-bounded by  $|C_{AB}|$ , where

$$C_{AB} := \frac{\text{Tr}(\rho[A, B])}{2i}. \quad (7.2)$$

### 7.1.2 Disturbance of a measurement

As the measurement outcomes of  $\mathcal{E}_A$  cannot be controlled, when one describes the disturbing effect of  $\mathcal{E}_A$  on the subsequent measurement of  $B$ , one focuses on the average over all outcomes. This means that one is considering the effect of a non-selective POVM measurement  $\mathcal{E}_A$  – in this way one captures the disturbing effect of the measurement itself, independently of which outcome is recorded. Furthermore, it is important to make a strict distinction between the disturbance of a quantum state and the disturbance to a subsequent quantum measurement. In general, performing a projective measurement of some observable  $A$  on a state  $\rho$  will affect (disturb) the state and change it into  $\rho' \neq \rho$  (apart from the special case, when  $\rho$  is diagonal in the basis of eigenstates of  $A$ ). The same holds true for a POVM measurement  $\mathcal{E}_A$ . The trade-off between information gain and state disturbance is itself a very subtle subject [129, 130], especially from the viewpoint of quantum information processing. However, let us stress that we are interested in the disturbance to a subsequent measurement of  $B$  and not of the system state. As even a perfect (projective and sharp) measurement of observable  $B$  gives us only insight into the probability distribution of a state  $\rho$  over the eigenstates of  $B$ , any state disturbance causing solely a change of the relative phases between eigenstates of  $B$  (the off-diagonal terms) should not be treated as disturbance to the measurement of  $B$ . In other words, disturbance of the measurement of  $B$  occurs if and only if diagonal elements of  $\rho$  in the basis of eigenstates of  $B$  change.

This is the essence of the OR, which is operationally motivated by the fact that only the change in the measurement statistics can be detected by the measurement (otherwise one would call a measurement disturbed even though it is indistinguishable from the perfect one). To be more precise let us denote the outcome probability distribution of a perfect measurement of  $B$  in a state  $\rho$  by  $\mathbf{p}^B(\rho)$  and the outcome probability distribution of a measurement of  $B$  on a state  $\mathcal{E}_A(\rho)$  (obtained after the measurement of an observable  $A$  performed on the original state  $\rho$ ) by  $\mathbf{p}^B(\mathcal{E}_A(\rho))$ . Then our requirement can be written as

$$\mathbf{p}^B(\rho) = \mathbf{p}^B(\mathcal{E}_A(\rho)) \Leftrightarrow \eta(B|\mathcal{E}_A, \rho) = 0, \quad (7.3)$$

which is simply the mathematical expression of Definition 7.1. Let us emphasise that Eq. (7.3) describes a state-dependent requirement for no disturbance. One might, however, require a stronger condition that a perfect (not disturbed) measurement should be state-independent, i.e., that the Born statistical formula should be satisfied for all initial states. As explained before, this can be achieved because state-independent notions can be recovered from the state-dependent ones.

## 7.2 Vanishing error and disturbance

As we have discussed, the definition of state-dependent disturbance should only depend upon operational distinguishability between the outcome statistics of a disturbed measurement and the ideal one. Although one can define an infinite number of different distance measures between probability distributions, all of them must assign zero for a pair of identical probability distributions. In this section we will investigate the consequences of this for notions of disturbance and constrain any state-dependent definitions of disturbance that fulfil the OR. We focus on the measurements in a finite  $d$ -dimensional Hilbert space and consider the following sequential measurement scenario. We perform a perfect (projective and sharp) measurement  $\mathcal{E}_A = \mathcal{D}_A$  of an observable  $A$ , and look at the disturbance to the measurement of an observable  $B$  that this projective measurement of  $A$  induces. The observables  $A$  and  $B$  are assumed to have non-degenerate spectra only for the sake of clarity, as this is not a necessary condition for the results presented in this section, unless stated otherwise. Within this framework we will analyse the possibilities for the existence of states for which a perfect measurement of observable  $A$  can be performed that causes no disturbance to the measurement of  $B$ . We will call such states *zero-noise, zero disturbance states* (ZNZD) and formally define them as follows:

### Definition 7.2: Zero-noise, zero-disturbance state

A state  $\rho$  is a zero-noise, zero-disturbance (ZNZD) state with respect to observables  $A$  and  $B$  if a perfect (projective and sharp) measurement  $\mathcal{E}_A = \mathcal{D}_A$  of an observable  $A$  does not change the probability distribution of a subsequent projective measurement of  $B$ .

Before presenting the main result of this section let us focus on two cases that one might consider trivial. Firstly, whenever a system state  $\rho$  is diagonal in the basis of the

eigenstates of  $A$ , the state after the non-selective projective measurement of  $A$  is clearly not disturbed as  $\mathcal{D}_A(\rho) = \rho$ , so also no subsequent measurement is disturbed. That means that in such a case for every  $B$  we have  $\eta(B, \rho) = \eta(B, \mathcal{D}_A(\rho)) = 0$ . The second trivial case is when the system is prepared in a maximally mixed state  $\rho = \mathbb{I}/d$ , so that it is diagonal in every basis. Since the maximally mixed state is unchanged by any measurement, and every measurement has the same uniform outcome probability distribution<sup>1</sup>, one has  $\eta(B, \mathbb{I}/d) = 0$  for all  $A$  and  $B$ . We also note that both cases trivially satisfy the restricted Ozawa's (and any commutator-based) relation, in the sense that the noise, disturbance and  $C_{AB}$  all simultaneously vanish for these states. The existence of these trivial zero-noise, zero-disturbance (ZNZD) states is therefore consistent with the commutator-based bounds.

We can now ask if there exist other ZNZD states, specifically ones for which the average value of the commutator  $[A, B]$  does not vanish. If the answer to this question is positive, then it means that state-dependent trade-off relations between error and disturbance satisfying the OR cannot be based only on the expectation value of the commutator,  $C_{AB}$ . In what follows we prove the existence of a family of pure ZNZD states for every pair of noncommuting observables  $A$  and  $B$ , and show that the expectation value of  $[A, B]$  generically does not vanish on these states.

### 7.2.1 Mutually unbiased ZNZD states

We start with the following straightforward lemma:

**Lemma 7.1: Sufficient condition for the existence of a pure ZNZD state**

If for any two observables,  $A$  and  $B$ , there exists a pure state  $|\psi_{AB}^*\rangle$  that is unbiased in eigenbases of both  $A$  and  $B$ ,

$$\forall i \quad |\langle a_i | \psi_{AB}^* \rangle|^2 = |\langle b_i | \psi_{AB}^* \rangle|^2 = \frac{1}{d}, \quad (7.4)$$

then for any two observables  $A$  and  $B$  there exists a pure ZNZD state given by  $|\psi_{AB}^*\rangle$ .

*Proof.* After the projective measurement of  $A$  the system initially in a state  $|\psi_{AB}^*\rangle$  will be transformed into a maximally mixed state  $\mathbb{I}/d$ . Therefore, the outcome probability

<sup>1</sup>Note, however, that the uncertainty here is entirely classical, and not associated with noncommutativity of  $A$  and  $B$ .

distribution of the subsequent measurement of  $B$  will be uniform, which is the same as before the measurement of  $A$ . Hence, the disturbance  $\eta(B, |\psi_{AB}^*\rangle)$  vanishes.  $\square$

We can now prove that, indeed, such pure ZNVD states  $|\psi_{AB}^*\rangle$  exist for every pair of observables  $A$  and  $B$ :

**Theorem 7.2: State unbiased with respect to two bases**

For any two bases  $\{|a_i\rangle\}$  and  $\{|b_i\rangle\}$  of a  $d$ -dimensional Hilbert space there exists a state  $|\psi_{AB}^*\rangle$  that is unbiased in both bases, i.e., that satisfies Eq. 7.4.

*Proof.* Let  $U^\dagger = \sum_i |b_i\rangle\langle a_i|$  denote the unitary connecting the  $\{|a_i\rangle\}$  basis to the  $\{|b_i\rangle\}$  basis. We need to show that there exists a pure quantum state  $|\psi_{AB}^*\rangle$  such that for every  $i$  we have

$$|\langle a_i | \psi_{AB}^* \rangle|^2 = \frac{1}{d}, \quad |\langle a_i | U | \psi_{AB}^* \rangle|^2 = \frac{1}{d}. \quad (7.5)$$

These two conditions imply that such a state must simultaneously satisfy the following

$$|\psi_{AB}^*\rangle = \frac{1}{\sqrt{d}} \sum_i e^{i\phi_i} |a_i\rangle, \quad U |\psi_{AB}^*\rangle = \frac{1}{\sqrt{d}} \sum_j e^{i\theta_j} |a_j\rangle, \quad (7.6)$$

with phases  $\{\phi_i\}$  and  $\{\theta_j\}$  belonging to the interval  $[0, 2\pi)$ . We will now show that one can always make a choice of  $\{\phi_i\}$  defining  $|\psi_{AB}^*\rangle$  so that  $U |\psi_{AB}^*\rangle$  has the desired form.

From Ref. [131] we know that every unitary matrix can be decomposed into Sinkhorn normal form given by  $U = LVR$ , where  $L$  and  $R$  are diagonal unitary matrices and  $V$  is a unitary matrix satisfying  $\sum_i V_{ij} = \sum_j V_{ij} = 1$ . By choosing the phases so that  $e^{i\phi_i} = R_{ii}^*$ , we get  $|\psi_{AB}^*\rangle = \frac{1}{\sqrt{d}} R^\dagger \sum_i |a_i\rangle$ . This, however, results in the following:

$$\begin{aligned} U |\psi_{AB}^*\rangle &= \frac{1}{\sqrt{d}} L V R R^\dagger \sum_i |a_i\rangle = \frac{1}{\sqrt{d}} L \sum_{i,j,k} V_{jk} |a_j\rangle \langle a_k | a_i\rangle = \frac{1}{\sqrt{d}} L \sum_j |a_j\rangle \sum_k V_{jk} \\ &= \frac{1}{\sqrt{d}} L \sum_j |a_j\rangle = \frac{1}{\sqrt{d}} \sum_j e^{i\theta_j} |a_j\rangle, \end{aligned}$$

where the phases  $\{\theta_j\}$  are given by  $e^{i\theta_j} = L_{jj}$ .  $\square$

Let us also provide a sketch of an alternative proof that does not involve the use of a decomposition of a unitary into Sinkhorn normal form (in fact, the reasoning presented below is also the main ingredient in proving that decomposition in Ref. [131]). The set of states of the form  $\frac{1}{\sqrt{d}} \sum_i e^{i\phi_i} |a_i\rangle$  defines a Clifford torus in the complex projective space  $\mathbb{C}P^{d-1}$  and, moreover, it is known [132] that the action of the unitary group induces a

Hamiltonian flow on  $\mathbb{C}P^{d-1}$ . However, it has been shown [133, 134] that this torus is not “Hamiltonian displaceable”, meaning that the image of the torus resulting from the action of  $U$  must intersect the original torus (in at least  $2^{d-1}$  points). This immediately implies the existence of at least  $2^{d-1}$  pure quantum states  $\{|\psi_{AB}^*\rangle\}$  that satisfy the required conditions. Let us illustrate the above reasoning on the simplest example of a qubit system:

**Example 7.1: Qubit states unbiased with respect to two bases**

For a qubit system the Clifford torus in  $\mathbb{C}P^1$  is simply given by the states lying on the equator of the Bloch sphere. Now, the unitary change of basis  $U$  corresponds to the rotation of the Bloch sphere. The fact that the considered torus is not displaced by  $U$  is equivalent to the statement that there does not exist a rotation resulting in a transformed great circle that would not intersect the original equator. Two states lying on the intersection points are unbiased in both the original basis and the one obtained by a unitary transformation  $U$ .

It is also intuitively clear that Theorem 7.2 cannot extend unconditionally to non-uniform distributions. Specifically, for any given state  $|\psi\rangle = \sum_i \sqrt{p_i} e^{i\phi_i} |a_i\rangle$ , such that not all of  $\{p_i\}$  are equal to  $1/d$ , there should exist a basis  $\{|b_n\rangle\}$  in which the probability distribution differs from the one given by  $\{p_i\}$ , no matter what the choices of  $\{\phi_i\}$  are made. Let us again use a qubit example to illustrate this statement. Consider a qubit system with the outcome probability distribution of  $\sigma_z$  measurement  $(p, 1-p)$ . States corresponding to this statistics form a circle on the Bloch sphere. Now if  $p \neq 1/2$ , i.e., if we are not dealing with the great circle, one can find a rotation of the Bloch sphere  $U$ , such that its action will transform the considered circle to the one not intersecting with the initial one. However, if we limit to “small rotations”, so that the “distance” between two bases is small enough (with respect to some appropriately defined distance measure, e.g.,  $\|\mathbb{I} - U\|$  in the operator norm) then the circles will intersect and there will exist a state that has the same statistics with respect to the two bases.

### 7.2.2 Existence of non-trivial ZNZZ states

We are now in the position that we know that for any observables  $A$  and  $B$  there exist at least  $2^{d-1}$  ZNZZ states  $|\psi_{AB}^*\rangle$ , but we lack the construction of such states. Therefore it is

not a priori obvious whether  $C_{AB}$ , defined in Eq. (7.2), is non-zero when  $[A, B] \neq 0$ . We will now show that  $C_{AB}$  is in fact non-vanishing unless a special choice of eigenvalues of  $A$  and  $B$  is made. In other words, we will prove that for two eigenbases  $\{|a_i\rangle\}$  and  $\{|b_i\rangle\}$  there always exists a choice of eigenvalues  $\{a_i\}$  and  $\{b_i\}$  such that  $C_{AB} \neq 0$ . Notice that this is in contrast with trivial ZNVD states (states diagonal in the eigenbasis of  $A$ ), where  $C_{AB}$  vanishes independently of the eigenvalues of  $A$  and  $B$ . Before proving this general result, however, we will consider a particular case of complementary observables, whose eigenstates form mutually unbiased bases (MUBs), as in this special case we can provide an explicit construction of  $|\psi_{AB}^*\rangle$  states.

For mutually unbiased observables  $A$  and  $B$  the eigenstates are related by

$$\forall i \quad |a_i\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d e^{i\phi_{ij}} |b_j\rangle, \quad \phi_{ij} \in [0, 2\pi]. \quad (7.7)$$

Now, it is known that for every  $d$ -dimensional Hilbert space there exist at least three mutually unbiased bases [135], which means that apart from  $\{|a_i\rangle\}$  and  $\{|b_i\rangle\}$  bases there also exist a basis  $\{|c_i\rangle\}$ , such that any  $|c_i\rangle$  can be taken as  $|\psi_{AB}^*\rangle$ . Also, the construction of such three bases is known for all  $d$  (e.g., by using the Heisenberg-Weyl group method [135]), so one can find explicit forms of  $|\psi_{AB}^*\rangle$ . For dimensions that are power of primes there actually exist  $(d + 1)$  MUBs. Hence, the existence of  $d(d - 1)$  states  $|\psi_{AB}^*\rangle$  is guaranteed, which is bigger than  $2^{d-1}$  for  $d \in \{3, 4, 5\}$ . Moreover, as has been found in Ref. [131], for  $d = 4$  and MUBs connected by the discrete Fourier transform unitary,

$$U = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}, \quad (7.8)$$

there exists a continuous family of  $|\psi_{AB}^*\rangle$  states given by

$$|\psi_{AB}^*\rangle = \frac{1}{2} (|a_1\rangle + ie^{-i\phi} |a_2\rangle + |a_3\rangle - ie^{-i\phi} |a_4\rangle), \quad (7.9)$$

for  $\phi \in [0, 2\pi)$ .

Since in the case of complementary observables we know the form of  $|\psi_{AB}^*\rangle$ , we can simply check if  $C_{AB}$  is non-zero. In the simplest case of  $d = 2$ , let us consider the mutually unbiased observables  $A = \sigma_x$  and  $B = \sigma_y$ . The state  $|\psi_{AB}^*\rangle$  is then the eigenstate of  $\sigma_z$

and, as we have  $[\sigma_x, \sigma_y] = 2i\sigma_z$ , we obtain  $|C_{AB}| = 1$ . For  $d = 3$  one can choose the following three unbiased bases

$$\{|a_i\rangle\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \quad (7.10a)$$

$$\{|b_i\rangle\} = \frac{1}{\sqrt{3}} \{(1, 1, 1), (1, \omega_3, \omega_3^2), (1, \omega_3^2, \omega_3)\}, \quad (7.10b)$$

$$\{|c_i\rangle\} = \frac{1}{\sqrt{3}} \{(1, \omega_3^2, \omega_3^2), (1, \omega_3, 1), (1, 1, \omega_3)\}, \quad (7.10c)$$

where  $\omega_3 = \exp(2\pi i/3)$ . In this case  $C_{AB}$  also does not vanish for at least one of  $|c_i\rangle$  states, unless  $[A, B] = 0$ , which can only be the case when  $A$  or  $B$  is completely degenerate and thus proportional to identity. As an example let us choose eigenvalues of  $A$  and  $B$  to be  $a_i = b_i = i - 2$ , for  $i \in \{1, 2, 3\}$ . Then  $|C_{AB}| = 1/2\sqrt{3}$  for  $|\psi_{AB}^*\rangle \in \{|c_1\rangle, |c_2\rangle\}$ , and  $|C_{AB}| = 1/\sqrt{3}$  for  $|\psi_{AB}^*\rangle = |c_3\rangle$ . Similarly, for  $d = 4$  one can choose the eigenstates of  $A$  to be the two-qubit computational basis, and the eigenstates of  $B$  to be defined by  $\{|b_i\rangle = H \otimes H |a_i\rangle\}$ , where  $H$  is the 2-dimensional Hadamard matrix. These two bases are mutually unbiased and, since for  $d = 4$  there exist five mutually unbiased bases, it leaves 12 states (four from each of the remaining three bases) that are ZNZZ states with respect to  $A$  and  $B$ . Again, unless  $[A, B] = 0$ , at least for one of this states we have  $C_{AB} \neq 0$ .

We have seen that for complementary observables  $C_{AB}$  does not vanish, unless a particularly special choice of eigenvalues  $\{a_i\}$  and  $\{b_i\}$  is made. We will now show that this is true for general observables. First of all note that

$$C_{AB} = \frac{1}{d} \sum_{ij} a_i b_j \operatorname{Im} (e^{i(\theta_j - \phi_i)} \langle a_i | b_j \rangle), \quad (7.11)$$

where the phases  $\{\phi_i\}$  and  $\{\theta_j\}$  are defined by Eq. (7.6). We will also make use of the following relation

$$e^{-i\theta_j} = \sum_i e^{-i\phi_i} \langle a_i | b_j \rangle. \quad (7.12)$$

We are now ready to prove that if  $C_{AB} = 0$  then by modifying only the eigenvalues of  $A$  and  $B$  we can obtain  $C_{AB} \neq 0$ .

We start by noting that  $\operatorname{Im} (e^{i(\theta_j - \phi_i)} \langle a_i | b_j \rangle)$  cannot vanish for all  $i$  and  $j$ , unless  $[A, B] = 0$ . To prove this let us assume that this expression vanishes for all  $i$  and  $j$ . Then we would have

$$\langle a_i | b_j \rangle = r_{ij} e^{i(\phi_i - \theta_j)}, \quad r_{ij} \in \mathbb{R}, \quad (7.13)$$

which, using Eq. (7.12), would mean that  $\sum_i r_{ij} = 1$ . However, from normalisation we have  $\sum_i r_{ij}^2 = 1$ , so that the only solution would be that the eigenstates of  $A$  and  $B$  coincide, and hence  $[A, B] = 0$ .

It is therefore guaranteed that there exist  $k$  and  $l$  such that  $\text{Im} (e^{i(\theta_l - \phi_k)} \langle a_k | b_l \rangle) \neq 0$ . Now, assume that  $C_{AB} = 0$ . By modifying the eigenvalue  $a_k$  to  $a_k + \epsilon_k$  and  $b_l$  to  $b_l + \epsilon_l$ , one gets

$$C_{AB} = \frac{1}{d} \left( \epsilon_l \tilde{a} + \epsilon_k \tilde{b} + \epsilon_k \epsilon_l \text{Im} (e^{i(\theta_l - \phi_k)} \langle a_k | b_l \rangle) \right), \quad (7.14)$$

with

$$\tilde{a} = \sum_i a_i \text{Im} (e^{i(\theta_l - \phi_i)} \langle a_i | b_l \rangle), \quad \tilde{b} = \sum_j b_j \text{Im} (e^{i(\theta_j - \phi_k)} \langle a_k | b_j \rangle). \quad (7.15)$$

We need to consider the following four possibilities:

1.  $\tilde{a} = 0$  and  $\tilde{b} = 0$ : the choice of any non-zero  $\epsilon_k$  and  $\epsilon_l$  guarantees  $C_{AB} \neq 0$ ;
2.  $\tilde{a} = 0$  and  $\tilde{b} \neq 0$ : the choice of  $\epsilon_l = 0$  any non-zero  $\epsilon_k$  guarantees  $C_{AB} \neq 0$ ;
3.  $\tilde{a} \neq 0$  and  $\tilde{b} = 0$ : the choice of  $\epsilon_k = 0$  any non-zero  $\epsilon_l$  guarantees  $C_{AB} \neq 0$ ;
4.  $\tilde{a} \neq 0$  and  $\tilde{b} \neq 0$ : either choices from points 2. and 3. guarantee  $C_{AB} \neq 0$ .

We have thus proven that for noncommuting observables,  $A$  and  $B$ , the average value of the commutator in a state  $|\psi_{AB}^*\rangle$  generically does not vanish.

### 7.2.3 Consequences for error-disturbance relations

As already mentioned, the existence of pure ZNVD states  $|\psi_{AB}^*\rangle$  for every pair of noncommuting observables  $A$  and  $B$ , such that the average of  $[A, B]$  does not vanish implies that any relation of the form

$$\sum_{m,n=0}^{\infty} c_{mn}(A, B) \epsilon^m(A, \rho) \eta^n(B, \rho) \geq \sum_{m=0}^{\infty} d_m(A, B) |C_{AB}|^m, \quad (7.16)$$

with  $c_{00} = 0$ ,  $d_m \geq 0$  and at least one non-zero  $d_m$ , must violate the operational requirement. This includes the relations developed and used in Refs. [113, 118, 124–128], in particular the restricted Ozawa's relation, Eq. (7.1), Ozawa's "universally valid error-disturbance uncertainty relation" given by

$$\epsilon(A, \rho) \eta(B, \rho) + \epsilon(A, \rho) \sqrt{V(B, \rho)} + \eta(B, \rho) \sqrt{V(A, \rho)} \geq |C_{AB}|, \quad (7.17)$$

and Branciard's trade-off relation:

$$\epsilon^2(A, \rho)V(B, \rho) + \eta^2(B, \rho)V(A, \rho) + 2\epsilon(A, \rho)\eta(B, \rho)\sqrt{V(A, \rho)V(B, \rho) - C_{AB}^2} \geq |C_{AB}|^2. \quad (7.18)$$

We conclude that the physical meaning of all such relations can be questioned on the basis of using disturbance measures that do not conform to the operational requirement.

### 7.3 Analysis of the Ozawa error-disturbance relation

In this final section we will analyse the well-known state-dependent Ozawa's trade-off relation [113], given by Eq. (7.17), focusing on the fact that it violates the OR, and emphasising that care should be taken in its interpretation. We have decided to discuss this particular relation separately from the general case presented in the previous section, due to the recent experimental investigations of Ozawa's error-disturbance trade-off relation with the use of qubit systems [124, 126]. These works claim to experimentally verify Ozawa's relation, which implies that any measurement of an observable  $A$  in a state  $\rho$  with error  $\epsilon(A, \rho)$  causes disturbance  $\eta(B, \rho)$  on another observable  $B$  satisfying Eq. (7.17). What we want to stress here is that the notion of disturbance being used has the non-operationally motivated properties, and so the sense in which it can describe how a disturbed measurement on any given state differs from the perfect one is debatable. If one insists on using what we consider better operationally motivated definition, then such a trade-off between error and disturbance will not be inevitable in general (i.e. applying for all states).

To see this more clearly let us analyse Ozawa's relation, specifically the experimentally investigated case of  $d = 2$ . Central to the relation are the error and disturbance terms which are defined for two observables  $A$  and  $B$ , and a pure state  $|\psi\rangle$  in the following way [136]:

$$\epsilon_O(A)^2 = \sum_k \|M_k(k - A)|\psi\rangle\|^2, \quad \eta_O(B)^2 = \sum_k \|[M_k, B]|\psi\rangle\|^2, \quad (7.19)$$

where  $\{M_k\}$  are the POVM elements induced by the actual measurement performed on the system, and  $k$  denote the corresponding eigenvalues of the observable. These terms, together with the variances  $V(A, \rho)$  and  $V(B, \rho)$  of  $A$  and  $B$  in the state  $|\psi\rangle$  can be shown to obey the error-disturbance relation given by Eq. (7.17), which is argued to be a rigorous generalisation of Heisenberg's microscope relation [113]. The above measures

of error and disturbance have been accused of being problematic, both in terms of what they quantify [114, 121] and in relation to interpretative issues [116]. Here, we address the non-operational meaning of the disturbance term  $\eta_{\mathcal{O}}$ , as well as its apparent state dependence.

First of all, let us note that if a perfect (projective and sharp) measurement of observable  $A$  is performed on a state  $|\psi\rangle$  then  $\epsilon_{\mathcal{O}}(A, |\psi\rangle) = 0$  and

$$\eta_{\mathcal{O}}(B, |\psi\rangle)^2 = \sum_k \| |a_k\rangle\langle a_k|, B | \psi \rangle \|^2. \quad (7.20)$$

Focusing on the disturbance for the initial state of the system being  $|a_l\rangle$ , i.e. the eigenstate of  $A$ , one has

$$\eta_{\mathcal{O}}(B, |\psi\rangle)^2 = \sum_{k \neq l} \| |a_k\rangle\langle a_k| B |a_l\rangle \|^2 + \| (\langle a_l| B |a_l\rangle - B) |a_l\rangle \|^2. \quad (7.21)$$

The sum on the right hand side of the above equation vanishes only when  $|a_l\rangle$  is the eigenstate of  $B$  (as  $B |a_l\rangle$  must be orthogonal to all  $|a_k\rangle$ ). Therefore, unless all the eigenstates of  $A$  coincide with the eigenstates of  $B$  (which implies  $[A, B] = 0$ ), for at least one of such eigenstates the disturbance is non-zero. We identify this as an unphysical property of the state-dependent disturbance measure, as the measurement of  $A$  performed on the eigenstate of  $A$  not only does not change the outcome probability distribution of the subsequent measurement of  $B$ , but also does not change the state of the system at all.

Let us now turn to the qubit scenario. Consider sequential projective measurements of  $A = \mathbf{a} \cdot \boldsymbol{\sigma}$  and  $B = \mathbf{b} \cdot \boldsymbol{\sigma}$  on a pure qubit state

$$|\psi\rangle\langle\psi| = \frac{\mathbb{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}, \quad (7.22)$$

where  $\boldsymbol{\sigma}$  denotes the vector of Pauli matrices and  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{r}$  are the Bloch vectors. It is then straightforward to compute [124] that  $\eta_{\mathcal{O}}$  is given by

$$\eta_{\mathcal{O}}(B, |\psi\rangle) = \sqrt{2} |\sin \gamma|, \quad (7.23)$$

where  $\gamma$  is the angle between the Bloch vectors  $\mathbf{a}$  and  $\mathbf{b}$ ,  $\cos \gamma = \mathbf{a} \cdot \mathbf{b}$ . For this elementary scenario, we find that although the definition of  $\eta_{\mathcal{O}}$  appears to be state-dependent, the resultant expression for a qubit system turns out to have no dependence on the system state  $|\psi\rangle$ . Further insight can be obtained by the following observation. Let us introduce

a state-dependent measure of disturbance  $\eta_K$  defined by the Kolmogorov distance between outcome probability distributions of a perfect and disturbed measurement, i.e.,

$$\eta_K(B, |\psi\rangle) := \frac{1}{2} \sum_i |p_i^B(|\psi\rangle\langle\psi|) - p_i^B(\mathcal{D}_A(|\psi\rangle\langle\psi|))|. \quad (7.24)$$

The operational meaning of the introduced measure of disturbance is as follows: the optimal success probability with maximum likelihood estimation for distinguishing between the perfect and disturbed probability distributions is given by  $(1 + \eta_K(B, |\psi\rangle))/2$ . Now it can be shown that the expression for Ozawa's disturbance  $\eta_O$  can be recovered by averaging the disturbance  $\eta_K$ , over all possible states of the system,

$$\langle \eta_K(B, |\psi\rangle) \rangle_{\text{Bloch}} = \frac{1}{4} |\sin \gamma| = \frac{1}{4\sqrt{2}} \eta_O(B, |\psi\rangle). \quad (7.25)$$

Thus, for  $d = 2$  the definition of disturbance proposed by Ozawa coincides with the average over the state-dependent notion defined here. It follows that  $\eta_O(B, |\psi\rangle)$  does not satisfy the operational requirement.

Finally, let us note that in a qubit case the set of non-trivial ZNVD states is not only limited to states  $|\psi_{AB}^*\rangle$ , i.e., the states unbiased in the bases of eigenstates of  $A$  and  $B$ . Without loss of generality one may choose the Bloch vectors representing considered observables to be  $\mathbf{a} = (0, 0, 1)$  and  $\mathbf{b} = (\sin \gamma, 0, \cos \gamma)$ . Consider then pure qubit states  $|\psi(\theta)\rangle$  described by Eq. (7.22) with Bloch vectors  $\mathbf{r} = (0, \sin \theta, \cos \theta)$  and  $\theta \in [0, \pi)$ . A projective measurement of  $B$  on such states yields a probability distribution  $\mathbf{p}^B(|\psi(\theta)\rangle\langle\psi(\theta)|) = (p, 1 - p)$ , with

$$p = \frac{1 + \mathbf{b} \cdot \mathbf{r}}{2} = \frac{1 + \cos \gamma \cos \theta}{2}. \quad (7.26)$$

If a projective measurement  $\mathcal{E}_A$  of  $A$  was performed before measuring  $B$  then one would get a probability distribution  $\mathbf{p}^B(\mathcal{E}_A(|\psi(\theta)\rangle\langle\psi(\theta)|)) = (p', 1 - p')$  with

$$p' = \frac{1 + \mathbf{a} \cdot \mathbf{r}}{2} \cdot \frac{1 + \mathbf{b} \cdot \mathbf{a}}{2} + \frac{1 - \mathbf{a} \cdot \mathbf{r}}{2} \cdot \frac{1 - \mathbf{b} \cdot \mathbf{a}}{2} = \frac{1 + \cos \gamma \cos \theta}{2}. \quad (7.27)$$

Since  $p = p'$ , meaning that a projective measurement of  $A$  on any of the states  $|\psi(\theta)\rangle$  does not change the statistics of the subsequent measurement of  $B$ , these states are ZNVD states for observables  $A$  and  $B$ . Moreover, the average value of commutator  $C_{AB} = \sin \theta \sin \gamma$ , so it does not vanish unless  $[A, B] = 0$  or  $|\psi(\theta)\rangle$  is the eigenstate of  $A$ .

# Conclusions

*Scientia potentia est*, knowledge is power, the latin aphorism goes. In fact, it is not just a maxim, but a statement that can be formalised and studied with scientific scrutiny. The pioneering works of Claude Shannon provided us with the tools to quantitatively assess the amount of information we have about a system: the further the probability distribution representing our knowledge is from being uniform, the more information we have. Moreover, looking through an information-theoretic lens on classical thermodynamics, we can see it as a theory of “biased” uncertainty. By exchanging the distance from a uniform distribution with a distance from a thermal distribution, our measures of uncertainty are replaced by free energy measures. These quantify how much mechanical work (“potentia”) we can perform by trading our knowledge (“scientia”) about a system state in the presence of a thermal bath. However, such considerations have been, until recently, restricted to classical information represented by probability distributions. Within the quantum realm the state of the system is represented by a density matrix, whose off-diagonal elements (coherences) lead to a new, purely quantum uncertainty. A proper understanding of its role in thermodynamics, as well as its interplay with classical uncertainty, was the main aim of this thesis.

In Part I we focused on thermodynamic scenarios, where the second law of thermodynamics enforces evolution towards thermal equilibrium. We have shown that the first law of thermodynamics leads to additional symmetry constraints on processing quantum coherences between different energy eigenstates. This in turn means that coherence is a thermodynamic resource and cannot be created for free. Moreover, without an external source of coherence (a reference system) no amount of work performed can overcome these constraints – coherence and energy are independent resources within thermodynamics. Thermodynamic processing of quantum information can thus be split into processing of its classical component (linked with the distribution over energy eigenstates) and purely

quantum component (coming from coherences between these states). As a result, coherence can be transferred only between pairs of energy states that belong to the same *mode of coherence*, and this transfer reflects thermodynamic irreversibility. Although we have formulated these restrictions in a mathematically rigorous way, there are still many open questions. First of all, our results concerning qubit systems suggest that it may be possible to infer the thermodynamic constraints on processing higher-dimensional quantum systems in the low-temperature limit [16] based solely on symmetry considerations. This would shed new light on the resource theory of thermodynamics in this particular limit, and would make a link with asymmetry theory even stronger. One could also try to use our bounds in order to solve the state interconversion problem under the evolution constrained to thermal operations (note that up till now this problem has been fully solved only for incoherent states [8]). In principle, this would require solving a maximisation problem over the constrained set of stochastic matrices. Thus, it seems (at least from a numerical point of view) that this problem should be tractable using semidefinite programming techniques. Finally, one could investigate how our results can be specialised to a particularly interesting case of Markovian evolution of coherences in an open quantum system, i.e., under the assumption of memoryless thermal bath [137]. In other words, how the bounds on coherence processing derived in this thesis can be translated into the Kossakowski-Linblad formalism [138, 139], and what they can tell us about the relation between relaxation and decoherence times.

One of the most important consequences of symmetry constraints in thermodynamics covered in this thesis is that the traditional Szilard argument [60] must be modified. The classical result links the information we have about a system with the amount of work we can perform using it – it is the free energy of the system that specifies its energetic value. Due to limitations of processing coherence this does not extend straightforwardly to the quantum case as was initially suggested [67]. Instead, one needs to use an ancillary reference system carrying coherence [14] in order to access the quantum information encoded in the system. Moreover, to fully exploit this information, one has to use a reference with unbounded coherence resources, the access to which should be questioned in the nanoscale regime. We showed that a bounded reference can still improve the performance of work extraction protocols (by either increasing the average extracted work or decreasing the failure probability of single-shot protocols), however the optimal way to achieve this re-

mains one of the important open problems. Another issue that requires deeper analysis is the transition between processing a single copy of a system and asymptotically many copies [6]. Although the two extremal scenarios have been discussed in detail in this thesis, the intermediate and most realistic regime is still not fully explored. Finally, taking into account recent experimental realisations and verifications of the generalised Jarzynski equality [21], Landauer’s erasure [140] and Szilard’s engine [141], one may try to devise an experimental setup that would demonstrate the extraction of work from coherence of a quantum system.

In this thesis we have also presented evidence that coherence can affect the structure of the thermodynamic arrow of time. Unlike previous direct studies on the state interconversion problem in thermodynamics [6, 8, 15], here we have focused on structural properties of the induced ordering of states. More precisely, the second law of thermodynamics induces an ordering between states that allows us to distinguish past from future. In the infinite temperature limit this ordering has a particular lattice structure: for any two states there exists a unique state in the past consistent with the considered states at present and with all possible joint pasts of those states; and there exists a unique state in the future consistent with the considered states at present and with all possible joint futures of those states. We proved that whereas this structure is lost at finite temperatures for incoherent states, it is preserved for states with coherence in the simplest quantum scenario of a two-dimensional system. Since partial order is the basic notion used within all resource theories, one may wonder whether a lattice structure appears and could be employed in the studies on resource theories of entanglement [1], coherence [93] or asymmetry [51]. In the first two cases it is suggested by the fact that the majorisation lattice describes the ordering between pure entangled [142] and pure coherent [143] states. In fact, very recently the lattice structure was used to investigate approximate transformations between pure bipartite entangled states [144]. On the other hand, already in the simplest case of a qubit system under the restriction of time-translation covariant evolution there is no top state (i.e., the one that could be freely transformed to all other states). This means that it is unlikely that the ordering of states in the resource theory of asymmetry exhibits a lattice structure.

The results we have presented in Part I suggest that a full theory of thermodynamics in the quantum regime will require a better understanding of the accounting of coherence resources. Despite our improving description of the role of coherence in quantum heat

engines [145], biological systems [146], and fluctuation theorems [147], we are still far from having a complete picture. Due to the fact that within thermodynamics the processing of coherence is constrained by time-translation covariance, it seems that we could get more insight into the problem by studying the role of references – systems that, like clocks, transform non-trivially under time translations. Most thermodynamic transformations rely on proper timing: whether it is time-dependence of the Hamiltonian controlled by some external parameter, performing an operation at a particular time of the evolution or even precisely performing a given operation (which, after all, comes from interaction with an external ancillary system and depends on the coupling strength and the interaction time). As we have shown in Part I, the only way to ignore the potential thermodynamic cost of such operations is to assume access to a classical clock. Instead, by introducing and investigating quantum clocks, we could obtain a more fundamental description of thermodynamic processes.

In Part II we turned to scenarios where quantum uncertainties have been traditionally investigated – the uncertainty relations. Although seemingly unrelated to thermodynamics, they are actually strongly linked with problems studied in Part I, as the measures of quantum uncertainty are mathematically equivalent to the measures of coherence used within the resource theory of thermodynamics. We studied the possibility of decomposing total uncertainty about a given observable into quantum and classical parts (which parallels the decomposition of free energy into these two parts). Unlike for variance-based measures [91], the sum of quantum uncertainties about two different observables measured by any entropic function satisfying basic desiderata, is not subject to state-independent constraints. Instead, it can be bounded relative to the degree of von Neumann entropy present in the state. When the total uncertainty is measured by the Shannon entropy, it can be cleanly separated into a sum of quantum and classical components. We proved that this natural decomposition leads to fixed-entropy bounds, which provide generalisations of the traditional entropic relations (that has recently been extensively studied [38, 85, 86, 88]), and emphasise the role of classical noise. In particular we have observed that the number of distinct states that minimise uncertainty with respect to two noncommuting observables  $A$  and  $B$  depends on the “distance” between the eigenbases of  $A$  and  $B$ , as well as on the amount of classical noise present in a state. The rigorous formulation of this simple observation seems to be a promising avenue in further studies

of uncertainty relations.

We also investigated the interplay between quantum and classical uncertainty manifested in the non-trivial structure of minimum uncertainty states (MUSs). These are quantum states that minimise the total uncertainty of two given observables. Although, there are a plethora of uncertainty measures resulting in many different MUSs (as many as one for each measure), we have shown that a particular entropic function, Rényi entropy  $H_2$ , plays a crucial role in understanding the structure of all MUSs when classical uncertainty is present. Namely, a strong enough uniform noise transforms all MUSs and brings them “closer” to the MUS linked with  $H_2$ . Therefore, even though each uncertainty measure may have its own operational meaning relevant in a specific scenario, for qubit systems (and, in a weaker sense, also for higher dimensions), there exist states that are universally optimal in a noisy enough environment. In order to fully exploit our results, however, further studies on the operational meaning of each uncertainty measure are needed. In Chapter 6 we also found that close to thermal equilibrium the interconversion of incoherent states is approximately ruled by a single quantity - the Rényi divergence of order two between a given state and a thermal Gibbs state. An interesting open question is whether a similar relation holds also for states with coherence. Many “coherent” out-of-equilibrium measures are known [9], however a complete set has not yet been found. Based on the incoherent result it is tempting to conjecture that close to equilibrium also all coherence measures collapse to a single quantity.

Finally, this thesis also addressed the problem of state-dependent trade-off relations between error and disturbance. We proved the existence of a state  $|\psi_{AB}^*\rangle$  that simultaneously maximises quantum uncertainty with respect to any two given observables  $A$  and  $B$ . This in turn allowed us to identify the set of zero-noise zero-disturbance states, for which a perfect measurement of  $A$  can be performed without disturbing the subsequent measurement of  $B$ . We thus showed that error-disturbance trade-off is inevitable only for state-independent measures of error and disturbance, which questions the physical meaning of recently developed state-dependent relations [113, 118, 124–128]. A natural extension to the considerations presented in this thesis is to study error-disturbance relations for measurements of more than two noncommuting observables. Since in general there does not exist a state that is unbiased with respect to more than two bases [103], it is possible that such non-trivial state-dependent relations can be formulated.

# Appendices

# Appendix A

## Alternative proof of Theorem 2.1

The Stinespring dilation theorem states that any CPTP map  $\mathcal{E}$  can be realised in the following way [148]:

$$\mathcal{E}(\rho) = \text{Tr}_A (U(\rho \otimes \tau_A)U^\dagger), \quad (\text{A.1})$$

where  $\tau_A$  is a state of an ancillary system and  $U$  is a unitary acting on the joint system. By denoting the final state by  $\sigma = \mathcal{E}(\rho)$  and expanding  $\tau_A$  in its own eigenbasis,  $\tau_A = \sum_i \lambda_i |i\rangle\langle i|$ , the above equation can be rewritten as [137]

$$\sigma = \sum_{k,l} W_{kl} \rho W_{kl}^\dagger, \quad (\text{A.2})$$

with  $W_{kl} = \sqrt{\lambda_k} \langle l|U|k\rangle$ . The final off-diagonal element (coherence between energy states)  $\sigma_{i'j'} = \langle \epsilon_{i'} | \mathcal{E}(\rho) | \epsilon_{j'} \rangle$  is thus given by

$$\sigma_{i'j'} = \sum_{i,j} \rho_{i,j} \sum_{k,l} \langle \epsilon_{i'} | W_{kl} | \epsilon_i \rangle \langle \epsilon_j | W_{kl}^\dagger | \epsilon_{j'} \rangle. \quad (\text{A.3})$$

We can now introduce a matrix  $X^{(xx')}$  with matrix elements given by  $X_{kl}^{(xx')} = \langle \epsilon_{x'} | W_{kl} | \epsilon_x \rangle$  that allows us to rewrite the above equation in the following way:

$$\sigma_{i'j'} = \sum_{i,j} \rho_{i,j} \text{Tr} \left( X^{(ii')} X^{(jj')\dagger} \right). \quad (\text{A.4})$$

The triangle inequality ensures that  $|\sigma_{i'j'}| \leq \sum_{i,j} |\rho_{i,j}| |\text{Tr} (X^{(ii')} X^{(jj')\dagger})|$ , so that using the Cauchy-Schwarz inequality we obtain:

$$|\sigma_{i'j'}| \leq \sum_{i,j} |\rho_{i,j}| \sqrt{\text{Tr} (X^{(ii')} X^{(ii')\dagger}) \text{Tr} (X^{(jj')} X^{(jj')\dagger})}. \quad (\text{A.5})$$

Note that the transition probability between diagonal states, as defined by Eq. (2.8), is given by

$$p_{i'|i} = \sum_{k,l} |\langle \epsilon_{i'} | W_{kl} | \epsilon_i \rangle|^2 = \text{Tr} \left( X^{(ii')} X^{(ii')\dagger} \right). \quad (\text{A.6})$$

Therefore, by substituting the above into Eq. (A.5), we arrive at a bound for processing coherence under a general CPTP map:

$$|\sigma_{i'j'}| \leq \sum_{i,j} \sqrt{p_{i'|i} p_{j'|j}} |\rho_{ij}|. \quad (\text{A.7})$$

Finally, the fact that  $\mathcal{E}$  is time-translation covariant greatly simplifies the obtained bound. From Eq. (2.4) we know that each mode of a quantum state transforms independently. This immediately implies that we can refine Eq. (A.7) to get Eq. (2.9).

# Appendix B

## Details of the proof of Theorem 3.1

We provide here the details of the repumping stage of the protocol. We start from a generic reference state  $\rho_R$  such that

$$\text{supp}(\rho_R) \cap \text{span}\{|\epsilon_0^R\rangle, \dots, |\epsilon_M^R\rangle\} = \emptyset,$$

where  $|\epsilon_i^R\rangle$  are eigenstates of the reference Hamiltonian. We impose the requirements  $Mp \gg 1$  and  $M(1-p) \gg 1$ , where  $p$  is fixed by Eq. (3.12).

We now compute the probability of the occurrence of the Kraus  $A_1$  on a generic reference state  $\sigma_R$ . From Eq. (3.23) and the fact that  $\Delta\Delta^\dagger = \mathbb{I}$  we obtain

$$p_1(\sigma_R) := \text{Tr}\left(A_1\sigma_R A_1^\dagger\right) = 1 - 2p(1-p)(1 - \text{Tr}(\bar{\Delta}\sigma_R)), \quad (\text{B.1})$$

where recall that  $\bar{\Delta} = (\Delta + \Delta^\dagger)/2$ . Define the state of the reference after performing work extraction on  $n \geq 1$  qubits through the following recurrence formula

$$\rho_R^{(n)} := A_0\rho_R^{(n-1)}A_0^\dagger + A_1\rho_R^{(n-1)}A_1^\dagger, \quad (\text{B.2})$$

where  $\rho_R^{(0)} = \rho_R$ . Because  $\rho_R$  has initially no support in the first  $M$  energy levels, we can extract work from  $M$  qubits before there is any overlap with the ground state. In other words,  $\langle \epsilon_0^R | \rho_R^{(n)} | \epsilon_0^R \rangle = 0$ ,  $\forall n \in \{1, \dots, M\}$ . The previous formula, together with Eq. (3.25), implies that  $\bar{\Delta}$  is conserved throughout the protocol. Hence we deduce that

$$\text{Tr}\left(A_1^M \rho_R A_1^{\dagger M}\right) = p_1^M(\rho_R) = (1 - 2p(1-p)(1 - \langle \bar{\Delta} \rangle))^M. \quad (\text{B.3})$$

For notational convenience, we will now drop the explicit dependence of  $p_1$  on  $\rho_R$  (initial state of the reference).

Using Eq. (B.2) we have

$$\rho_R^{(M)} = p_1^M \rho_{R,1} + (1 - p_1^M) \mathcal{E}_{\text{else}}^{(M)}(\rho_R), \quad (\text{B.4})$$

where  $\mathcal{E}_{\text{else}}^{(M)}$  contains all strings of  $A_0$ 's and  $A_1$ 's different from the string consisting only of  $A_1$ 's, and

$$\rho_{R,1} = \frac{A_1^M \rho_R A_1^{\dagger M}}{p_1^M}. \quad (\text{B.5})$$

We can now compute  $A_1^M$ :

$$A_1^M = \sum_{k=0}^M \binom{M}{k} (1-p)^{M-k} p^k \Delta^{\dagger k}. \quad (\text{B.6})$$

We see that  $A_1^M$  is binomially distributed in the number of lowering operations  $\Delta^\dagger$ . The average number of lowerings is  $\bar{M} = Mp$  and the standard deviation is  $\sigma_M = \sqrt{Mp(1-p)}$ . We can perform a number of repumpings as in Eq. (3.19). Let us denote this operation by  $\mathcal{P}$ . We have chosen  $M$  sufficiently large so that the confidence levels associated to  $\sigma_M$  are approximately gaussian. Hence, we can repump the reference  $\bar{M} + s\sigma_M^{4/3}$  times, which guarantees that the reference has arbitrarily small population in states  $|\epsilon_0^R\rangle \dots |\epsilon_M^R\rangle$  with a confidence level controlled by  $s > 0$  and increasing with  $M$ . More precisely, if  $P_M$  is the projector on the subspace spanned by  $\{|\epsilon_0^R\rangle, \dots, |\epsilon_M^R\rangle\}$  and  $P_M^\perp = \mathbb{I} - P_M$ ,

$$\text{Tr}(P_M^\perp \mathcal{P}(\rho_{R,1})) \geq \text{erf}(sM^{1/6}/\sqrt{2}) := \mathbb{E}_s(M^{1/6}), \quad (\text{B.7})$$

where  $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$  denotes the error function. Now, using Eq. (B.4) and Eq. (B.7)

$$p_{\text{succ}} := \text{Tr}(P_M^\perp \mathcal{P}(\rho_R^{(M)})) \geq p_1^M \mathbb{E}_s(M^{1/6}). \quad (\text{B.8})$$

This implies that in performing the two-outcome measurement  $\{P_M, P_M^\perp\}$  we would find the outcome  $P_M^\perp$  with probability given by Eq. (B.8).

Performing such a measurement guarantees that the final state of the reference will have no support on a subspace spanned by  $\{|\epsilon_0^R\rangle \dots |\epsilon_M^R\rangle\}$ , similarly to the initial state. However, performing a selective measurement has a thermodynamic cost that we have to take into account. More precisely, such a measurement can be performed using an ancillary memory qubit system  $A$  described by trivial Hamiltonian  $H_A = 0$ . Then, taking the initial state of  $A$  to be a pure state  $|0\rangle$ , we can perform an operation on the joint reference-ancillary state described by the Kraus operators  $M_1 = P_M^\perp \otimes \mathbb{I}$  and  $M_2 = P_M \otimes \sigma_x$ . This operation is energy conserving, as the Kraus operators commute with the total

Hamiltonian  $H_R + H_A$ . Hence, it is free of thermodynamic cost. Now, the projective measurement on states  $|0\rangle$  and  $|1\rangle$  can be performed on the ancillary memory system. Observing the result 0 will project the reference on a subspace  $P_M^\perp$ , whereas observing the result 1 will project the reference on  $P_M$ . The thermodynamic cost associated with this projective measurement is the cost of erasing the memory system afterwards. This is given by

$$W_E = \frac{1}{\beta} h_2(p_{\text{succ}}), \quad (\text{B.9})$$

which can be made arbitrarily small as  $p_{\text{succ}} \rightarrow 1$ . Notice that we only needed to use a classical memory to record the measurement outcome. Also note that this cost has to be paid only after extracting work from  $M$  copies, hence the cost per copy scales as  $M^{-1}$ .

Define

$$\rho_R'' := \frac{P_M^\perp \mathcal{P}(\rho_R^{(M)}) P_M^\perp}{\left\| P_M^\perp \mathcal{P}(\rho_R^{(M)}) P_M^\perp \right\|}. \quad (\text{B.10})$$

Now, using the gentle measurement lemma [149, 150], Eq. (B.8) also implies

$$\left\| \rho_R'' - \mathcal{P}(\rho_R^{(M)}) \right\| \leq 2\sqrt{1 - p_{\text{succ}}}. \quad (\text{B.11})$$

From Eq. (B.11), and the following characterisation of the trace norm (see Ref. [148]),

$$\|\rho - \sigma\| = \max_{0 \leq A \leq \mathbb{I}} \text{Tr}(A(\rho - \sigma)),$$

we find that

$$\text{Tr}(\bar{\Delta} \rho_R'') \geq \text{Tr}(\bar{\Delta} \mathcal{P}(\rho_R^{(M)})) - 2\sqrt{1 - p_{\text{succ}}} = \text{Tr}(\bar{\Delta} \rho_R) - 2\sqrt{1 - p_{\text{succ}}}, \quad (\text{B.12})$$

where the last equality comes from the fact that  $\langle \bar{\Delta} \rangle$  is conserved in the protocol, up to the measurement. The last equation can be rewritten as

$$\text{Tr}(\bar{\Delta}(\rho_R - \rho_R'')) \leq 2\sqrt{1 - p_{\text{succ}}}. \quad (\text{B.13})$$

Exchanging the roles of  $\rho_R$  and  $\rho_R''$  and introducing

$$\delta \langle \bar{\Delta} \rangle = \text{Tr}(\bar{\Delta} \rho_R'') - \text{Tr}(\bar{\Delta} \rho_R), \quad (\text{B.14})$$

we conclude

$$|\delta \langle \bar{\Delta} \rangle| \leq 2\sqrt{1 - p_{\text{succ}}}. \quad (\text{B.15})$$

Using Eq. (B.8), this bounds the maximum allowed change of the quality parameter of the reference.

# Appendix C

## Proof of Theorem 4.1

The proof of Theorem 4.1 is based on the Alberti-Uhlmann theorem [84]. In the qubit case it yields necessary and sufficient conditions for the existence of a CPTP map  $\mathcal{E}$  such that  $\mathcal{E}(\rho) = \rho'$  and  $\mathcal{E}(\sigma) = \sigma'$ . These conditions are given by:

$$\|\lambda\rho - (1 - \lambda)\sigma\|_1 \geq \|\lambda\rho' - (1 - \lambda)\sigma'\|_1 \quad \forall \lambda \in [0, 1], \quad (\text{C.1})$$

where  $\|A\|_1 = \text{Tr}(\sqrt{AA^\dagger})$ . Thus, by simply setting  $\sigma = \sigma' = \gamma$ , we obtain necessary and sufficient conditions for the existence of a GP quantum map  $\mathcal{E}$  such that  $\mathcal{E}(\rho) = \rho'$ . Using the fact that norms are non-negative, these can be expressed by

$$\Delta_\lambda := D_\lambda(\rho) - D_\lambda(\rho') \geq 0 \quad \forall \lambda \in [0, 1], \quad (\text{C.2})$$

where we have defined  $D_\lambda(\rho) := \|\lambda\rho - (1 - \lambda)\gamma\|_1^2$ .

Before we find the necessary and sufficient conditions for Eq. (C.2) to hold, let us first simplify the problem. Namely, note that states connected via a unitary  $U(t) = e^{iHt}$  are reversibly interconvertible under GP operations, and thus belong to the same equivalence class. Hence, we can focus only on one representative of this class lying in the  $xz$  plane of the Bloch sphere with  $x \geq 0$ . This means that instead of considering general Bloch vectors of the form  $\mathbf{r}_\rho = (x, y, z)$  we can focus only on the ones given by  $\mathbf{r}_\rho = (x, 0, z)$ .

Now, using the parametrisation of qubit states introduced in Eqs. (4.10)-(4.11), we can write

$$D_\lambda^2(\rho) = a(\rho) + \sqrt{a(\rho)^2 - b(\rho)^2}, \quad (\text{C.3})$$

where

$$a(\rho) = \left(2 + \frac{x^2 + (z + \zeta)^2}{2}\right) \lambda^2 - (2 + \zeta(z + \zeta)) \lambda + \frac{1 + \zeta^2}{2}, \quad (\text{C.4a})$$

$$b(\rho) = (1 - 2\lambda)\sqrt{\zeta^2 - 2\zeta(z + \zeta)\lambda + (x^2 + (z + \zeta)^2)\lambda^2}. \quad (\text{C.4b})$$

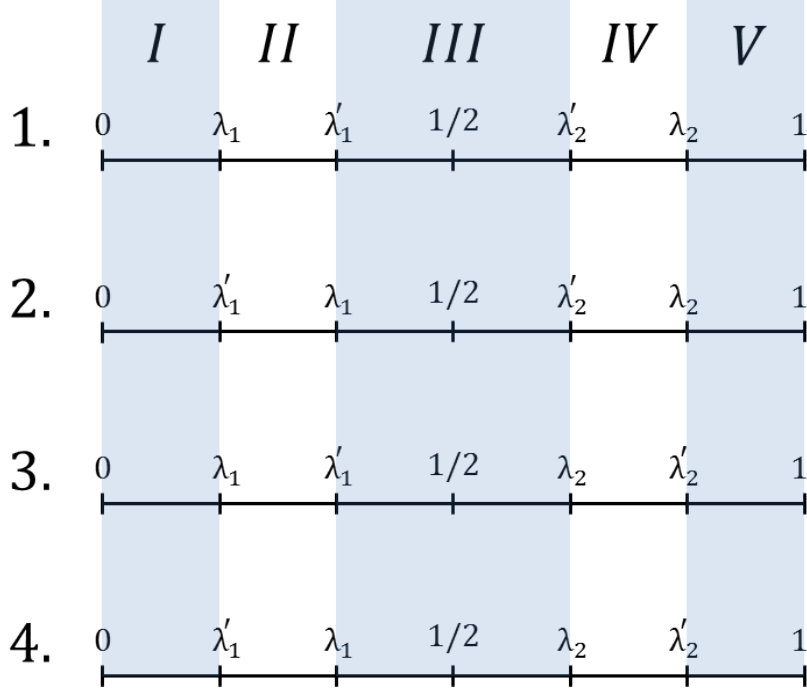


Figure C.1: *Possible orderings of eigenvalues.* The orderings 1–4 of eigenvalues  $\{\lambda_1, \lambda'_1, \lambda_2, \lambda'_2\}$ , together with labels I–V corresponding to different regions of the  $[0, 1]$  interval.

However, notice that

$$a(\rho)^2 - b(\rho)^2 = [a(\rho) - (1 - 2\lambda)^2]^2 := c(\rho)^2,$$

which means that

$$D_\lambda^2(\rho) = \begin{cases} (1 - 2\lambda)^2 & : c(\rho) \leq 0, \\ (1 - 2\lambda)^2 + 2c(\rho) & : c(\rho) > 0. \end{cases} \quad (\text{C.5})$$

Since  $c(\rho)$  is quadratic in  $\lambda$ , one can explicitly express the regions of  $\lambda$  with different solutions for  $D_\lambda^2(\rho)$  by finding the zeros of  $c(\rho) = A(\lambda - \lambda_1)(\lambda - \lambda_2)$ . Using elementary calculus one can show by direct calculation that  $0 \leq \lambda_1 \leq \frac{1}{2} \leq \lambda_2 \leq 1$  and  $A \leq 0$ . Hence, we can rewrite Eq. (C.3) as

$$D_\lambda^2(\rho) = \begin{cases} (1 - 2\lambda)^2 & : \lambda \in [0, \lambda_1] \text{ and } \lambda \in [\lambda_2, 1], \\ (1 - 2\lambda)^2 + 2c(\rho) & : \lambda \in (\lambda_1, \lambda_2). \end{cases} \quad (\text{C.6})$$

Thus, by analogously introducing  $c(\rho')$  with solutions given by  $\lambda'_1$  and  $\lambda'_2$ , the necessary and sufficient conditions for the existence of a GP map  $\mathcal{E}$  such that  $\mathcal{E}(\rho) = \rho'$  [specified by Eq. (C.2)] can be expressed as follows. There are four ways the eigenvalues  $\{\lambda_1, \lambda'_1, \lambda_2, \lambda'_2\}$  can be ordered and these are depicted in Fig. C.1. In regions I and V (see Fig. C.1),

independently of the ordering, we have  $\Delta_\lambda = 0$ . Now, in region II we have  $\Delta_\lambda = 2c(\rho) \geq 0$  for orderings 1. and 3., whereas for orderings 2. and 4. we have  $\Delta_\lambda = -2c(\rho') \leq 0$ . Similarly, in region IV one has that  $\Delta_\lambda$  is equal to  $2c(\rho) \geq 0$  for orderings 1. and 2., whereas for orderings 3. and 4. it is equal to  $-2c(\rho') \leq 0$ . Therefore, for  $\Delta_\lambda$  to be positive for all  $\lambda \in [0, 1]$ , the eigenvalues must be ordered according to ordering 1. The remaining condition to check is whether within region III this ordering also results in  $\Delta_\lambda \geq 0$ . We have  $\Delta_\lambda = 2[c(\rho) - c(\rho')]$  and since both quadratic functions,  $c(\rho)$  and  $c(\rho')$ , share the same constant term and  $\lambda \neq 0$ , the problem can actually be simplified to comparing two linear functions. This, in turn, can be done by simply comparing the functions at the edge of the region III, where  $c(\rho) \geq 0$  and  $c(\rho') = 0$ . Therefore, we conclude that Eq. (C.2) holds if and only if  $\lambda_1 \leq \lambda'_1$  and  $\lambda_2 \geq \lambda'_2$ .

In the final step of the proof we need to use explicit expressions for  $\lambda_1$  and  $\lambda_2$ ,

$$\lambda_1 = \frac{2 - \zeta(z + \zeta) - \delta}{4 - (z + \zeta)^2 - x^2}, \quad \lambda_2 = \frac{2 - \zeta(z + \zeta) + \delta}{4 - (z + \zeta)^2 - x^2}, \quad (\text{C.7})$$

where  $\delta$  is given by Eq. (4.12). By solving the above equations for  $z$  with fixed  $\lambda_m$  (with  $m \in \{1, 2\}$ ), one can find that the region of fixed  $\lambda_m$  is given by a circle centred at  $\mathbf{z}_m = [0, 0, \zeta(\lambda_m^{-1} - 1)]$  and of radius  $R_m = |\lambda_m^{-1} - 2|$ . It is a straightforward calculation to show that these centres and radii correspond exactly to the ones stated in Corollary 4.2. Moreover, since  $0 \leq \lambda_1 \leq \frac{1}{2} \leq \lambda_2 \leq 1$ , we have

$$\lambda_1 \leq \lambda'_1 \Leftrightarrow R_1(\rho) \geq R_1(\rho'), \quad (\text{C.8a})$$

$$\lambda_2 \geq \lambda'_2 \Leftrightarrow R_2(\rho) \geq R_2(\rho'). \quad (\text{C.8b})$$

Hence, the necessary and sufficient condition for the existence of a GP map between  $\rho$  and  $\rho'$ , specified by Eq. (C.2), is that  $R_1(\rho) \geq R_1(\rho')$  and  $R_2(\rho) \geq R_2(\rho')$ . This can be equivalently expressed with the use of simplified variables  $R_\pm(\rho)$  as in Theorem 4.1. Finally, we note that given two circles of radii  $R_1$  and  $R'_1$ , centred at  $(0, 0, \zeta(1 + R_1))$  and  $(0, 0, \zeta(1 + R'_1))$ , respectively, we have that the circle with smaller radius is contained within the circle of bigger radius. The same holds true for circles of radii  $R_2$  and  $R'_2$  centred at  $(0, 0, \zeta(1 - R_2))$  and  $(0, 0, \zeta(1 - R'_2))$ . We have thus finished the proof of Corollary 4.2.

# Appendix D

## Existence of a universal MUS for qubit systems

### Setting the scene

The Bloch sphere can be parametrised as in Chapter 5, so that  $\mathbf{a} = (0, 0, 1)$  and  $\mathbf{b} = (\sin \gamma, 0, \cos \gamma)$ , with  $\gamma \in (0, \pi/2)$ . Now, according to Lemma 6.4, if there exists an  $\epsilon$ -noisy universal MUS it will be described by

$$\rho_{\gamma/2}^\epsilon = \frac{\mathbb{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}, \quad \mathbf{r} = \pm|r| \left( \sin \frac{\gamma}{2}, 0, \cos \frac{\gamma}{2} \right), \quad (\text{D.1})$$

with  $|r| = (1 - \epsilon)$ , i.e., its Bloch vector  $\mathbf{r}$  will lie in the middle between  $\pm\mathbf{a}$  and  $\pm\mathbf{b}$ , and its length will differ from identity by the amount of noise  $\epsilon$ .

We will first prove that there is no universal MUS for qubit systems when we choose  $\mathcal{F} = \mathcal{U}_0$ , i.e., within the framework of majorisation uncertainty relations. Let  $\rho_0^\epsilon$  be a state described by a Bloch vector  $\mathbf{r}' = (0, 0, |r|)$ . Then, by direct calculation, one can check that for any given  $\epsilon$  the distribution  $\mathbf{p}^A(\rho_0^\epsilon) \otimes \mathbf{p}^B(\rho_0^\epsilon)$  is not majorised by  $\mathbf{p}^A(\rho_{\gamma/2}^\epsilon) \otimes \mathbf{p}^B(\rho_{\gamma/2}^\epsilon)$ .

Now, in order to prove that there exists  $0 < \epsilon < 1$  such that  $\rho_{\gamma/2}^\epsilon$  is a universal MUS within  $\mathcal{S}_2^\epsilon$  for  $\mathcal{F} = \mathcal{U}$  and  $\mathcal{F} = \mathcal{U}_+$ , we must show that all Rényi entropies of  $\mathbf{p}^A(\rho_{\gamma/2}^\epsilon) \otimes \mathbf{p}^B(\rho_{\gamma/2}^\epsilon)$  are smaller than for any other state  $\rho^\epsilon \in \mathcal{S}_2^\epsilon$  that can generally be described by a Bloch vector  $\mathbf{q}$ ,

$$\rho^\epsilon = \frac{\mathbb{I} + \mathbf{q} \cdot \boldsymbol{\sigma}}{2}, \quad \mathbf{q} = |q|(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (\text{D.2})$$

with  $|q| \leq (1 - \epsilon)$ . As we will show in the next subsection it is actually sufficient to restrict

the comparison to states described by Bloch vectors

$$\mathbf{q}_\theta = (1 - \epsilon)(\sin \theta, 0, \cos \theta), \quad \theta \in [0, \gamma], \quad (\text{D.3})$$

i.e., one may assume  $|q| = (1 - \epsilon)$ ,  $\theta \in [0, \gamma]$  and  $\phi = 0$  in Eq. (D.2) (note that this corresponds to Bloch vectors lying between  $\pm \mathbf{a}$  and  $\pm \mathbf{b}$ ). Thus, the existence of a universal MUS in  $\mathcal{S}_2^\epsilon$  can be proven by showing that among the distributions

$$\mathbf{p}^{AB} := (p_1^A, 1 - p_1^A) \otimes (p_1^B, 1 - p_1^B), \quad (\text{D.4})$$

with

$$p_1^A = \frac{1 + (1 - \epsilon) \cos \theta}{2}, \quad p_1^B = \frac{1 + (1 - \epsilon) \cos(\gamma - \theta)}{2}, \quad (\text{D.5})$$

the one with  $\theta = \gamma/2$  minimises all Rényi entropies. As our main goal is to prove that a universal MUS can exist for  $\mathcal{F} = \mathcal{U}$  and  $\mathcal{F} = \mathcal{U}_+$ , we will just focus on the particular choice of  $\gamma = \pi/4$ . In the last subsection of this appendix we will prove that for a noise level  $\epsilon = 1/2$  (Bloch vector length  $1/2$ ) the above condition holds, so that a state specified by Eq. (D.1) is in fact the  $\epsilon$ -noisy universal MUS.

### Simplifying the set of states

The way to prove that we can restrict the comparison to states  $\rho_\theta^\epsilon$ , described by Bloch vectors  $\mathbf{q}_\theta$  specified by Eq. (D.3), is to show that for every state  $\rho^\epsilon \in \mathcal{S}_2^\epsilon$  there exists an  $\epsilon$ -pseudo-pure state  $\rho_\theta^\epsilon$  such that

$$\mathbf{p}^A(\rho_\theta^\epsilon) \otimes \mathbf{p}^B(\rho_\theta^\epsilon) \succ \mathbf{p}^A(\rho^\epsilon) \otimes \mathbf{p}^B(\rho^\epsilon). \quad (\text{D.6})$$

Since majorisation implies trumping relation, it means that for any state in  $\mathcal{S}_2^\epsilon$  there exists a state  $\rho_\theta^\epsilon$  for which all Rényi entropies  $H_\alpha$  are lower. Hence, if  $\rho_{\gamma/2}^\epsilon$  minimises all  $H_\alpha$  among  $\rho_\theta^\epsilon$  states, all the remaining states in  $\mathcal{S}_2^\epsilon$  must necessarily have higher  $H_\alpha$  for all  $\alpha$ .

Using the transitivity of majorisation we will prove our claim by restricting the subset  $\mathcal{S}_2^\epsilon$  in a few steps, each time removing states that are majorised by states in the remaining subset. First, note that for any state  $\rho^\epsilon$  with  $|q| < (1 - \epsilon)$  there exists a state  $\tau^\epsilon$  with  $|q| = (1 - \epsilon)$  and the same  $(\theta, \phi)$ , such that  $\mathbf{p}^A(\tau^\epsilon) \succ \mathbf{p}^A(\rho^\epsilon)$  and  $\mathbf{p}^B(\tau^\epsilon) \succ \mathbf{p}^B(\rho^\epsilon)$ , so that  $\mathbf{p}^A(\tau^\epsilon) \otimes \mathbf{p}^B(\tau^\epsilon) \succ \mathbf{p}^A(\rho^\epsilon) \otimes \mathbf{p}^B(\rho^\epsilon)$ . Hence, we can restrict the states that we need to compare  $\rho_{\gamma/2}^\epsilon$  with to states having a Bloch vector of length  $|q| = (1 - \epsilon)$ . Next, we note that states with fixed  $|q|$  and  $\theta$  have a fixed  $\mathbf{p}^A$ , but  $\mathbf{p}^B$  that depends on  $\phi$ . Moreover,

$\mathbf{p}^B$  for  $\phi = 0$  or  $\phi = \pi$  majorises all other  $\mathbf{p}^B$  with different  $\phi$ . Hence we can restrict the considered set of states to the ones that lie in the plane spanned by  $\mathbf{a}$  and  $\mathbf{b}$ . Additionally, due to the symmetry of the problem, we only need to look at  $\theta \in [0, \pi]$  and  $\gamma \in (0, \pi/2)$  (we exclude  $\gamma = 0$  as it is a trivial case, and  $\gamma = \pi/2$  because Theorem 6.6 holds).

Finally, we need to show that for any state  $\rho^\epsilon$  that lies in this plane, has  $|q| = (1 - \epsilon)$  and  $\theta \in [0, \pi]$  there exists  $\rho_\theta^\epsilon$  also in that plane and with the same length of the Bloch vector, but with  $\theta \in [0, \gamma]$ , such that Eq. (D.6) holds. It is straightforward to check that for a state with  $\theta = \gamma$  both  $\mathbf{p}^A$  and  $\mathbf{p}^B$  majorise the corresponding distributions obtained for a state described by  $\theta \in (\gamma, \pi/2]$ . Similarly the probability distribution with  $\theta = 0$  majorises the ones described by  $\theta \in [\gamma + \pi/2, \pi]$ . The only thing left is to show that for every state described by  $\theta \in [\pi/2, \pi/2 + \gamma]$  there is a state with  $\theta \in [0, \gamma]$  such that Eq. (D.6) holds. One can achieve this by mapping  $\theta$  of every state from the first set to  $\theta - \pi/2$  in the second set. This ends the proof.

### Proving the existence of a universal MUS

As already announced, here we will prove that among distributions  $\mathbf{p}^{AB} := (p_1^A, 1 - p_1^A) \otimes (p_1^B, 1 - p_1^B)$ , specified by Eq. (D.5) with  $\gamma = \pi/4$  and  $\epsilon = 1/2$ , the one with  $\theta = \pi/8$  minimises Rényi entropies for all  $\alpha$ . In order to simplify the calculations we use a slightly different parametrisation. Namely, we perform a substitution  $\theta \rightarrow \theta - \pi/8$  and, due to symmetry, we only consider  $\theta \in [0, \pi/8]$  (hence  $\theta$ , instead of measuring the angle from the  $z$  axis, measures the angle from the state we want to prove is a universal MUS). To shorten the notation let us also introduce

$$t_{\pm}^{\pm} = \left[ 2 \pm \cos\left(\frac{\pi}{8} \pm \theta\right) \right]^{\alpha-1}, \quad (\text{D.7})$$

where the subscript refers to the sign  $\pm$  in front of  $\theta$ .

It is straightforward to show that independently of  $\alpha$  the Rényi entropy  $H_\alpha$  of the distribution  $\mathbf{p}^{AB}$  has an extremum for  $\theta = 0$ . However, we need to show that this is the only extremum and that it is actually a minimum. Once we prove the former, the latter can be easily verified by checking that  $H_\alpha$  of the distribution at the extremum  $\theta = 0$  is smaller than at the edge of the region  $\theta = \pi/8$ . To prove the uniqueness of the extremum we will show that  $\frac{\partial}{\partial \theta} H_\alpha(\mathbf{p}^{AB}) = 0$  has only a single solution for  $\theta = 0$ . Unless  $\alpha = 0$  or

$\alpha = 1$  (which will be handled separately) vanishing of this derivative is equivalent to

$$\zeta := A \sin \theta + B \cos \theta = 0, \quad (\text{D.8})$$

where

$$A = \left(2 \cos \frac{\pi}{8} - \cos \theta\right) t_-^- t_+^- - \left(2 \cos \frac{\pi}{8} + \cos \theta\right) t_+^+ t_+^+, \quad (\text{D.9a})$$

$$B = \left(2 \sin \frac{\pi}{8} + \sin \theta\right) t_-^+ t_+^- - \left(2 \sin \frac{\pi}{8} - \sin \theta\right) t_-^- t_+^+. \quad (\text{D.9b})$$

The proof consists of two main parts. First we will show that for all  $\alpha \geq 2$  and  $\theta \in (0, \pi/8]$  we have  $\zeta < 0$ . We achieve this by finding a function  $\zeta'$  that upper-bounds  $\zeta$  in the considered parameter region and proving that it is negative. Next, we will prove that for all  $\alpha \leq -3$  and  $\theta \in (0, \pi/8]$  we have  $\zeta > 0$ , this time by finding a function  $\zeta'$  that lower-bounds  $\zeta$  and showing that it is always positive. Finally, in the remaining region  $\alpha \in [-3, 2]$  the non-vanishing of  $\zeta$  can be easily verified numerically (with  $\alpha = 0$  and  $\alpha = 1$  considered separately).

Let us start with  $\alpha \geq 2$ . For  $\theta \in (0, \pi/8]$  we have that a function  $\zeta'$  obtained by exchanging  $A$  with  $-2t_-^+ t_+^+ \cos \theta$  upper-bounds  $\zeta$  (this is because  $t_+^+ t_+^+ \geq t_-^- t_+^-$ ). To show that  $\zeta'$  is always negative we divide it by the positive quantity  $t_-^+ t_+^+ \cos \theta$ , and show that the obtained expression  $\zeta_1 + \zeta_2$  is always negative, where

$$\zeta_1 = -\sin \theta + s_+ \sin \theta, \quad (\text{D.10a})$$

$$\zeta_2 = -\sin \theta + 2s_- \sin \frac{\pi}{8}, \quad (\text{D.10b})$$

and

$$s_{\pm} = \frac{t_{\pm}^-}{t_{\pm}^+} \pm \frac{t_{\pm}^-}{t_{\pm}^+}. \quad (\text{D.11})$$

Now, using the fact that  $(a+b)^x \geq a^x + b^x$  for  $a, b > 0$  and  $x \geq 1$  one can easily show that  $s_+ < 1$ , which results in  $\zeta_1 < 0$ . In order to show that also  $\zeta_2$  is negative it is sufficient to prove that  $\zeta_2$  is a monotonically decreasing function with  $\theta$  (since for  $\theta = 0$  it vanishes). This can be shown by upper-bounding terms dependent on  $\theta \in (0, \pi/8]$  in the expression for  $\frac{\partial \zeta_2}{\partial \theta}$  that leads to

$$\frac{\partial \zeta_2}{\partial \theta} \leq -\cos \frac{\pi}{8} + \frac{8\sqrt{2} \sin \frac{\pi}{8}}{(\cos \frac{\pi}{8} - 2)^4} \frac{\alpha - 1}{2^\alpha}. \quad (\text{D.12})$$

The above expression is maximised for  $\alpha = (1 + \ln 2)/\ln 2$  and it is then negative, so that  $\zeta_2$  is negative for  $\theta \in (0, \pi/8]$ . This ends the first part of the proof.

We now turn to the case when  $\alpha \leq -3$ . For  $\theta \in (0, \pi/8]$  we then have that a function  $\zeta'$  obtained by exchanging the second term in the expression for  $A$  by  $-(2 \cos \frac{\pi}{8} + \cos \theta) t_-^+ t_+^-$  lower-bounds  $\zeta$  (this is because  $t_+^+ \leq t_+^-$ ). Moreover, substituting 1 for all  $\cos \theta$  will also lower-bound the expression for  $\zeta$ , as for  $\alpha \leq -3$  we have  $B \leq 0$ . We further lower-bound the expression by performing a sequence of divisions and subtractions of positive numbers: first dividing by  $t_-^+ t_+^-$ , then subtracting  $t_+^+ \sin \theta / t_+^-$  and finally dividing again by  $2 \sin \pi/8$ . This leaves us with the lower-bound of the form:

$$\sin \theta \frac{2 \cos \frac{\pi}{8} \left(1 - \frac{t_-^+}{t_+^-}\right) - 1}{2 \sin \frac{\pi}{8}} + \left(\frac{t_-^+}{t_+^-} - \frac{t_+^+}{t_+^-}\right). \quad (\text{D.13})$$

As the term standing by  $\sin \theta$  is a monotonically decreasing function of  $\alpha$  it achieves minimum at the edge of the considered parameter space, i.e., for  $\alpha = -3$ . It is then straightforward to verify that it is always bigger than 1, so we can actually lower-bound  $\zeta$  with

$$\sin \theta + \left(\frac{t_-^+}{t_+^-} - \frac{t_+^+}{t_+^-}\right). \quad (\text{D.14})$$

To show that the above equation is always positive for  $\alpha \leq -3$  one can equivalently show that  $\sin \theta - s_-$  is positive for  $\alpha \geq 3$ . This can be achieved using a method analogous to the one used to show that  $\zeta_2 < 0$ . Thus, the function lower-bounding  $\zeta$  for  $\alpha \leq -3$  is always positive unless  $\theta = 0$  and so is  $\zeta$  itself. This completes the second part of the proof.

# Bibliography

- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.* **81**, 865–942 (2009).
- [2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels,” *Phys. Rev. Lett.* **70**, 1895 (1993).
- [3] C. H. Bennett and S. J. Wiesner, “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states,” *Phys. Rev. Lett.* **69**, 2881–2884 (1992).
- [4] C. Carathéodory, “Untersuchungen über die Grundlagen der Thermodynamik,” *Math. Ann.* **67**, 355–386 (1909).
- [5] R. Giles, *Mathematical Foundations of Thermodynamics* (Pergamon Press, 1964).
- [6] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, “Resource theory of quantum states out of thermal equilibrium,” *Phys. Rev. Lett.* **111**, 250404 (2013).
- [7] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, “A measure of majorization emerging from single-shot statistical mechanics,” *New J. Phys.* **17**, 073001 (2015).
- [8] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, “The second laws of quantum thermodynamics,” *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3275 (2015).
- [9] M. Lostaglio, D. Jennings, and T. Rudolph, “Description of quantum coherence in thermodynamic processes requires constraints beyond free energy,” *Nat. Commun.* **6**, 6383 (2015).

- 
- [10] M. Horodecki and J. Oppenheim, “Fundamental limitations for quantum and nanoscale thermodynamics,” *Nat. Commun.* **4**, 2059 (2013).
- [11] J. Åberg, “Truly work-like work extraction via a single-shot analysis,” *Nat. Commun.* **4**, 1925 (2013).
- [12] C. A. Rodríguez-Rosario, T. Frauenheim, and A. Aspuru-Guzik, “Thermodynamics of quantum coherence,” *arXiv:1308.1245* (2013).
- [13] P. Skrzypczyk, A. J. Short, and S. Popescu, “Extracting work from quantum systems,” *arXiv:1302.2811* (2013).
- [14] J. Åberg, “Catalytic coherence,” *Phys. Rev. Lett.* **113**, 150402 (2014).
- [15] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, “Limitations on the evolution of quantum coherences: Towards fully quantum second laws of thermodynamics,” *Phys. Rev. Lett.* **115**, 210403 (2015).
- [16] V. Narasimhachar and G. Gour, “Low-temperature thermodynamics with quantum coherence,” *Nat. Commun.* **6**, 7689 (2015).
- [17] O. Karlström, H. Linke, G. Karlström, and A. Wacker, “Increasing thermoelectric performance using coherent transport,” *Phys. Rev. B* **84**, 113415 (2011).
- [18] H. Vazquez, R. Skouta, S. Schneebeli, M. Kamenetska, R. Breslow, L. Venkataraman, and M. Hybertsen, “Probing the conductance superposition law in single-molecule circuits with parallel paths,” *Nat. Nanotechnol.* **7**, 663–667 (2012).
- [19] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, “Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies,” *Nature* **437**, 231–234 (2005).
- [20] V. Serreli, C. Lee, E. Kay, and D. Leigh, “A molecular information ratchet,” *Nature* **445** (2007).
- [21] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, “Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality,” *Nat. Phys.* **6**, 988–992 (2010).

- [22] Alemany, A. and Ritort, F., “Fluctuation theorems in small systems: extending thermodynamics to the nanoscale,” *Europhys. News* **41**, 27–30 (2010).
- [23] J. Cheng, S. Sreelatha, R. Hou, A. Efremov, R. Liu, J. R. C. van der Maarel, and Z. Wang, “Bipedal nanowalker by pure physical mechanisms,” *Phys. Rev. Lett.* **109**, 238104 (2012).
- [24] S. Lloyd, “Quantum coherence in biological systems,” *J. Phys.: Conf. Ser.* **302**, 012037 (2011).
- [25] N. Lambert, Y.-N. Chen, Y.-C. Cheng, C.-M. Li, G.-Y. Chen, and F. Nori, “Quantum biology,” *Nat. Phys.* **9** (2013).
- [26] E. M. Gauger, E. Rieper, J. J. Morton, S. C. Benjamin, and V. Vedral, “Sustained quantum coherence and entanglement in the avian compass,” *Phys. Rev. Lett.* **106**, 040503 (2011).
- [27] M. Planck, “Ueber das Gesetz der Energieverteilung im Normalspectrum,” *Ann. Phys. (Berlin)* **309**, 553–563 (1901).
- [28] H. P. Robertson, “The uncertainty principle,” *Phys. Rev.* **34**, 163 (1929).
- [29] W. Heisenberg, *The physical principles of the quantum theory* (Courier Dover Publications, Mineola, 1949).
- [30] E. Kennard, “Zur Quantenmechanik einfacher Bewegungstypen,” *Z. Phys.* **44**, 326–352 (1927).
- [31] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.* **27**, 379–423 (1948).
- [32] I. Białyński-Birula and J. Mycielski, “Uncertainty relations for information entropy in wave mechanics,” *Commun. Math. Phys.* **44**, 129–132 (1975).
- [33] D. Deutsch, “Uncertainty in quantum measurements,” *Phys. Rev. Lett.* **50**, 631–633 (1983).
- [34] H. Maassen and J. B. Uffink, “Generalized entropic uncertainty relations,” *Phys. Rev. Lett.* **60**, 1103–1106 (1988).

- [35] S. Wehner and A. Winter, “Entropic uncertainty relations – a survey,” *New J. Phys.* **12**, 025009 (2010).
- [36] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, “Entropic uncertainty relations and their applications,” *arXiv:1511.04857* (2015).
- [37] M. Koashi, “Unconditional security of quantum key distribution and the uncertainty principle,” *J. Phys.: Conf. Ser.* **36**, 98 (2006).
- [38] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, “The uncertainty principle in the presence of quantum memory,” *Nat. Phys.* **6**, 659–662 (2010).
- [39] P. J. Coles, J. Kaniewski, and S. Wehner, “Equivalence of wave–particle duality to entropic uncertainty,” *Nat. Commun.* **5** (2014).
- [40] C.-F. Li, J.-S. Xu, X.-Y. Xu, K. Li, and G.-C. Guo, “Experimental investigation of the entanglement-assisted entropic uncertainty principle,” *Nat. Phys.* **7**, 752–756 (2011).
- [41] R. Prevedel, D. R. Hamel, R. Colbeck, K. Fisher, and K. J. Resch, “Experimental investigation of the uncertainty principle in the presence of quantum memory and its application to witnessing entanglement,” *Nat. Phys.* **7**, 757–761 (2011).
- [42] A. Renyi, “On measures of entropy and information,” in *Fourth Berkeley Symposium on Mathematical Statistics and Probability* (1961) pp. 547–561.
- [43] M. P. Müller and M. Pastena, “A generalization of majorization that characterizes Shannon entropy,” *IEEE Trans. Inf. Theory* **62**, 1711–1720 (2016).
- [44] R. Bhatia, *Matrix analysis*, Vol. 169 (Springer Science & Business Media, 1997).
- [45] S. Daftuar and M. Klimesh, “Mathematical structure of entanglement catalysis,” *Phys. Rev. A* **64**, 042314 (2001).
- [46] M. Klimesh, “Inequalities that collectively completely characterize the catalytic majorization relation,” *arXiv:0709.3680* (2007).
- [47] S. Turgut, “Catalytic transformations for bipartite pure states,” *J. Phys. A* **40**, 12185 (2007).

- [48] I. Marvian and R. W. Spekkens, “Modes of asymmetry: the application of harmonic analysis to symmetric quantum dynamics and quantum reference frames,” *Phys. Rev. A* **90**, 062110 (2014).
- [49] I. Marvian and R. W. Spekkens, “Extending Noether’s theorem by quantifying the asymmetry of quantum states,” *Nat. Commun.* **5**, 3821 (2014).
- [50] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, “Thermodynamic cost of reliability and low temperatures: Tightening Landauer’s principle and the second law,” *Int. J. Theor. Phys.* **39**, 2717–2753 (2000).
- [51] I. Marvian and R. W. Spekkens, “The theory of manipulations of pure state asymmetry: I. basic tools, equivalence classes and single copy transformations,” *New J. Phys.* **15**, 033001 (2013).
- [52] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, “Reference frames, superselection rules, and quantum information,” *Rev. Mod. Phys.* **79**, 555 (2007).
- [53] M. Keyl and R. F. Werner, “Optimal cloning of pure states, testing single clones,” *J. Math. Phys.* **40**, 3283–3299 (1999).
- [54] G. Gour and R. W. Spekkens, “The resource theory of quantum reference frames: manipulations and monotones,” *New J. Phys.* **10**, 033023 (2008).
- [55] I. Marvian, *Symmetry, Asymmetry and Quantum Information*, Ph.D. thesis, University of Waterloo (2012).
- [56] F. Reif, *Fundamentals of Statistical and Thermal Physics*, Fundamentals of Physics Series (McGraw-Hill, 1965).
- [57] E. Ruch, R. Schraner, and T. H. Seligman, “The mixing distance,” *J. Chem. Phys.* **69** (1978).
- [58] D. Jonathan and M. B. Plenio, “Entanglement-assisted local manipulation of pure quantum states,” *Phys. Rev. Lett.* **83**, 3566–3569 (1999).
- [59] C. Jarzynski, “Nonequilibrium equality for free energy differences,” *Phys. Rev. Lett.* **78**, 2690–2693 (1997).

- [60] L. Szilard, “Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen,” *Z. Phys.* **53**, 840–856 (1929).
- [61] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, “Quantum coherence, time-translation symmetry, and thermodynamics,” *Phys. Rev. X* **5**, 021001 (2015).
- [62] A. Jamiołkowski, “Linear transformations which preserve trace and positive semidefiniteness of operators,” *Rep. Math. Phys.* **3**, 275–278 (1972).
- [63] G. M. D’ariano and P. Lo Presti, “Optimal nonuniversally covariant cloning,” *Phys. Rev. A* **64**, 042308 (2001).
- [64] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.* **10**, 285–290 (1975).
- [65] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, New York, NY, USA, 2006).
- [66] J. C. Maxwell, *Theory of Heat* (Longmans, Green and Co., 1873).
- [67] P. Skrzypczyk, A. J. Short, and S. Popescu, “Work extraction and thermodynamics for individual quantum systems,” *Nat. Commun.* **5**, 4185 (2014).
- [68] P. Kammerlander and J. Anders, “Coherence and measurement in quantum thermodynamics,” *Sci. Rep.* **6**, 22174 (2016).
- [69] M. F. Frenzel, D. Jennings, and T. Rudolph, “Reexamination of pure qubit work extraction,” *Phys. Rev. E* **90**, 052136 (2014).
- [70] R. Gallego, J. Eisert, and H. Wilming, “Thermodynamic work from operational principles,” [arXiv:1504.05056](https://arxiv.org/abs/1504.05056) (2015).
- [71] O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, “Inadequacy of von neumann entropy for characterizing extractable work,” *New J. Phys.* **13**, 053015 (2011).
- [72] N. Y. Halpern, A. J. Garner, O. C. Dahlsten, and V. Vedral, “Introducing one-shot work into fluctuation relations,” *New J. Phys.* **17**, 095003 (2015).
- [73] K. Korzekwa, M. Lostaglio, D. Jennings, and T. Rudolph, “Quantum and classical entropic uncertainty relations,” *Phys. Rev. A* **89**, 042122 (2014).

- [74] Y. Aharonov and L. Susskind, “Charge superselection rule,” *Phys. Rev.* **155**, 1428–1431 (1967).
- [75] D. Gelbwaser-Klimovsky, R. Alicki, and G. Kurizki, “Work and energy gain of heat-pumped quantized amplifiers,” *Europhys. Lett.* **103**, 60005 (2013).
- [76] D. Gelbwaser-Klimovsky and G. Kurizki, “Heat-machine control by quantum-state preparation: From quantum engines to refrigerators,” *Phys. Rev. E* **90**, 022102 (2014).
- [77] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, “Degradation of a quantum reference frame,” *New J. Phys.* **8**, 58 (2006).
- [78] G. A. White, J. A. Vaccaro, and H. M. Wiseman, “The consumption of reference resources,” *AIP Conf. Ser.* **1110**, 79–82 (2009).
- [79] M. Ahmadi, D. Jennings, and T. Rudolph, “Dynamics of a quantum reference frame undergoing selective measurements and coherent interactions,” *Phys. Rev. A* **82**, 032320 (2010).
- [80] G. Gour, I. Marvian, and R. W. Spekkens, “Measuring the quality of a quantum reference frame: the relative entropy of frameness,” *Phys. Rev. A* **80**, 012307 (2009).
- [81] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, “Quantum communication using a bounded-size quantum reference frame,” *New J. Phys.* **11**, 063013 (2009).
- [82] G. Birkhoff and J. Von Neumann, “The logic of quantum mechanics,” *Ann. Math.*, 823–843 (1936).
- [83] F. Cicalese and U. Vaccaro, “Supermodularity and subadditivity properties of the entropy on the majorization lattice,” *IEEE Trans. Inf. Theory* **48**, 933–938 (2002).
- [84] P. Alberti and A. Uhlmann, “A problem relating to positive linear maps on matrix algebras,” *Rep. Math. Phys.* **18**, 163–176 (1980).
- [85] Z. Puchała, Ł. Rudnicki, and K. Życzkowski, “Majorization entropic uncertainty relations,” *J. Phys. A* **46**, 272002 (2013).

- [86] Ł. Rudnicki, Z. Puchała, and K. Życzkowski, “Strong majorization entropic uncertainty relations,” *Phys. Rev. A* **89**, 052115 (2014).
- [87] M. H. Partovi, “Majorization formulation of uncertainty in quantum mechanics,” *Phys. Rev. A* **84**, 052117 (2011).
- [88] S. Friedland, V. Gheorghiu, and G. Gour, “Universal uncertainty relations,” *Phys. Rev. Lett.* **111**, 230401 (2013).
- [89] J. I. de Vicente and J. Sánchez-Ruiz, “Improved bounds on entropic uncertainty relations,” *Phys. Rev. A* **77**, 042110 (2008).
- [90] P. J. Coles and M. Piani, “Improved entropic uncertainty relations and information exclusion relations,” *Phys. Rev. A* **89**, 022112 (2014).
- [91] S. Luo, “Quantum versus classical uncertainty,” *Theor. Math. Phys.* **143**, 681–688 (2005).
- [92] J. Åberg, “Quantifying superposition,” *arXiv:0612146* (2006).
- [93] T. Baumgratz, M. Cramer, and M. Plenio, “Quantifying coherence,” *Phys. Rev. Lett.* **113**, 140401 (2014).
- [94] P. J. Coles, R. Colbeck, L. Yu, and M. Żwolak, “Uncertainty relations from simple entropic properties,” *Phys. Rev. Lett.* **108**, 210405 (2012).
- [95] T. Sagawa and M. Ueda, “Second law of thermodynamics with discrete quantum feedback control,” *Phys. Rev. Lett.* **100**, 080403 (2008).
- [96] H. Groenewold, “A problem of information gain by quantal measurements,” *Int. J. Theor. Phys.* **4**, 327–338 (1971).
- [97] M. Ozawa, “On information gain by quantum measurements of continuous observables,” *J. Math. Phys.* **27**, 759–763 (1986).
- [98] T. Sagawa, “Thermodynamics of information processing in small systems,” *Progr. Theor. Phys.* **127**, 1–56 (2012).

- [99] K. Korzekwa, D. Jennings, and T. Rudolph, “Operational constraints on state-dependent formulations of quantum error-disturbance trade-off relations,” *Phys. Rev. A* **89**, 052108 (2014).
- [100] P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths, “Information-theoretic treatment of tripartite systems and quantum channels,” *Phys. Rev. A* **83**, 062338 (2011).
- [101] J. Sánchez-Ruiz, “Optimal entropic uncertainty relation in two-dimensional Hilbert space,” *Phys. Lett. A* **244**, 189–195 (1998).
- [102] G. Ghirardi, L. Marinatto, and R. Romano, “An optimal entropic uncertainty relation in a two-dimensional Hilbert space,” *Phys. Lett. A* **317**, 32–36 (2003).
- [103] Z. Puchała, Ł. Rudnicki, K. Chabuda, M. Paraniak, and K. Życzkowski, “Certainty relations, mutual entanglement, and nondisplaceable manifolds,” *Phys. Rev. A* **92**, 032109 (2015).
- [104] G. Bosyk, M. Portesi, and A. Plastino, “Collision entropy and optimal uncertainty,” *Phys. Rev. A* **85**, 012108 (2012).
- [105] S. Luo, “Heisenberg uncertainty relation for mixed states,” *Phys. Rev. A* **72**, 042110 (2005).
- [106] Y. M. Park, “Improvement of uncertainty relations for mixed states,” *J. Math. Phys.* **46**, 042109 (2005).
- [107] H. J. Landau and H. O. Pollak, “Prolate spheroidal wave functions, Fourier analysis and uncertainty – II,” *Bell Syst. Tech. J.* **40**, 65–84 (1961).
- [108] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, “The resource theory of informational nonequilibrium in thermodynamics,” *Phys. Rep.* **583**, 1 (2015).
- [109] D. G. Cory, A. F. Fahmy, and T. F. Havel, “Ensemble quantum computing by NMR spectroscopy,” *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1634–1639 (1997).
- [110] P. Faist, J. Oppenheim, and R. Renner, “Gibbs-preserving maps outperform thermal operations in the quantum regime,” *New J. Phys.* **17**, 043003 (2015).

- [111] V. Scarani, M. Ziman, P. Stelmachovic, N. Gisin, and V. Buzek, “Thermalizing quantum machines: Dissipation and entanglement,” *Phys. Rev. Lett.* **88**, 097905 (2002).
- [112] M. Lostaglio, M. P. Müller, and M. Pastena, “Stochastic independence as a resource in small-scale thermodynamics,” *Phys. Rev. Lett.* **115**, 150402 (2015).
- [113] M. Ozawa, “Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement,” *Phys. Rev. A* **67**, 042105 (2003).
- [114] P. Busch, P. Lahti, and R. F. Werner, “Proof of Heisenberg’s error-disturbance relation,” *Phys. Rev. Lett.* **111**, 160405 (2013).
- [115] M. Ozawa, “Disproving Heisenberg’s error-disturbance relation,” [arXiv:1308.3540](https://arxiv.org/abs/1308.3540) (2013).
- [116] J. Dressel and F. Nori, “Certainty in Heisenberg’s uncertainty principle: Revisiting definitions for estimation errors and disturbance,” *Phys. Rev. A* **89**, 022106 (2014).
- [117] A. C. Ipsen, “Error-disturbance relations for finite dimensional systems,” [arXiv:1311.0259](https://arxiv.org/abs/1311.0259) (2013).
- [118] C. Branciard, “Error-tradeoff and error-disturbance relations for incompatible quantum measurements,” *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6742–6747 (2013).
- [119] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, “Experimental joint quantum measurements with minimum uncertainty,” *Phys. Rev. Lett.* **112**, 020401 (2014).
- [120] P. Busch, P. Lahti, and R. F. Werner, “Heisenberg uncertainty for qubit measurements,” *Phys. Rev. A* **89**, 012129 (2014).
- [121] P. Busch, P. Lahti, and R. F. Werner, “Colloquium: Quantum root-mean-square error and measurement uncertainty relations,” *Rev. Mod. Phys.* **86**, 1261 (2014).
- [122] L. A. Rozema, D. H. Mahler, A. Hayat, and A. M. Steinberg, “A note on different definitions of momentum disturbance,” *Quantum Stud.: Math. Found.* **2**, 17–22 (2015).

- [123] C. Bastos, A. E. Bernardini, O. Bertolami, N. C. Dias, and J. N. Prata, “Robertson-Schrödinger formulation of Ozawa’s uncertainty principle,” *J. Phys.: Conf. Ser.* **626**, 012050 (2015).
- [124] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, “Experimental demonstration of a universally valid error-disturbance uncertainty relation in spin measurements,” *Nat. Phys.* **8**, 185–189 (2012).
- [125] K. Fujikawa, “Universally valid heisenberg uncertainty relation,” *Phys. Rev. A* **85**, 062117 (2012).
- [126] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, “Experimental violation and reformulation of the Heisenberg’s error-disturbance uncertainty relation,” *Sci. Rep.* **3** (2013).
- [127] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, “Experimental test of error-disturbance uncertainty relations by weak measurement,” *Phys. Rev. Lett.* **112**, 020402 (2014).
- [128] X.-M. Lu, S. Yu, K. Fujikawa, and C. Oh, “Improved error-tradeoff and error-disturbance relations in terms of measurement error components,” *Phys. Rev. A* **90**, 042113 (2014).
- [129] C. A. Fuchs, “Information gain vs. state disturbance in quantum theory,” [arXiv:quant-ph/9611010](https://arxiv.org/abs/quant-ph/9611010) (1996).
- [130] C. A. Fuchs and A. Peres, “Quantum-state disturbance versus information gain: uncertainty relations for quantum information,” *Phys. Rev. A* **53**, 2038 (1996).
- [131] M. Idel and M. M. Wolf, “Sinkhorn normal form for unitary matrices,” *Linear Algebra Appl.* **471**, 76–84 (2015).
- [132] M. De Gosson, *Symplectic geometry and quantum mechanics* (Birkhäuser Verlag, Basel, 2006).
- [133] C.-H. Cho, “Holomorphic discs, spin structures, and Floer cohomology of the Clifford torus,” *Int. Math. Res. Notices* **2004**, 1803–1843 (2004).

- 
- [134] S. Lisi, “Given two basis sets for a finite Hilbert space, does an unbiased vector exist?” <http://math.stackexchange.com/q/29819> (2011).
- [135] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, “On mutually unbiased bases,” *Int. J. Quantum Inf.* **8**, 535–640 (2010).
- [136] M. Ozawa, “Universal uncertainty principle in the measurement operator formalism,” *J. Opt. B-Quantum S. O.* **7**, S672 (2005).
- [137] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, 2002).
- [138] G. Lindblad, “On the generators of quantum dynamical semigroups,” *Commun. Math. Phys.* **48**, 119–130 (1976).
- [139] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, “Completely positive dynamical semigroups of n-level systems,” *J. Math. Phys.* **17**, 821–825 (1976).
- [140] A. Berut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, “Experimental verification of Landauer’s principle linking information and thermodynamics,” *Nature* **483** (2012).
- [141] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, “Experimental realization of a Szilard engine with a single electron,” *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13786–13789 (2014).
- [142] M. A. Nielsen, “Conditions for a class of entanglement transformations,” *Phys. Rev. Lett.* **83**, 436 (1999).
- [143] S. Du, Z. Bai, and Y. Guo, “Conditions for coherence transformations under incoherent operations,” *Phys. Rev. A* **91**, 052120 (2015).
- [144] G. Bosyk, G. Sergioli, H. Freytes, F. Holik, and G. Bellomo, “Approximate transformations of bipartite pure-state entanglement from the majorization lattice,” [arXiv:1608.04818](https://arxiv.org/abs/1608.04818) (2016).
- [145] R. Uzdin, A. Levy, and R. Kosloff, “Quantum equivalence and quantum signatures in heat engines,” *Phys. Rev. X* **5**, 031044 (2015).

- [146] N. Killoran, S. F. Huelga, and M. B. Plenio, “Enhancing light-harvesting power with coherent vibrational interactions: a quantum heat engine picture,” *J. Chem. Phys.* **143**, 155102 (2015).
- [147] J. Åberg, “Fully quantum fluctuation theorems,” [arXiv:1601.01302](#) (2016).
- [148] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
- [149] A. Winter, “Coding theorem and strong converse for quantum channels,” *IEEE Trans. Inf. Theory* **45**, 2481–2485 (1999).
- [150] T. Ogawa and H. Nagaoka, “A new proof of the channel coding theorem via hypothesis testing in quantum information theory,” [arXiv:quant-ph/0208139](#) (2002).