Author:
Lipka-Bartosik, Patryk J

Title:
Resource theories and their applications in quantum thermodynamics and nonlocality

University of Bristol

Doctoral Thesis

Resource theories and their applications in quantum thermodynamics and nonlocality

Author:
Patryk Lipka-Bartosik

Supervisor:
Dr. Paul Skrzypczyk

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy

in the

Quantum Information Theory Group
School of Physics

July 5, 2021

Declaration of Authorship

I, Patryk Lipka-Bartosik, declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award.

Except where indicated by specific reference in the text, the work is my own work.

Work done in collaboration with, or with the assistance of, others, is indicated as such.

Any views expressed in the dissertation are mine.

Signed:

Date:

UNIVERSITY OF BRISTOL

Abstract

Faculty of Science
School of Physics

Doctor of Philosophy

Resource theories and their applications in quantum thermodynamics and nonlocality

by Patryk Lipka-Bartosik

The laws of physics are often expressed as limitations on what physical systems can and cannot do. This fundamental idea is at the core of Quantum Resource Theories, a rapidly evolving framework that postulates certain physical limitations on the evolution of quantum systems, and formally describes their consequences.

The goal of this thesis is twofold. First, to use a resource-theoretic approach to better understand the role and significance of quantum effects in thermodynamic processes. Second, to utilise this approach in establishing a direct and operational interpretation of nonclassical effects encountered in quantum nonlocality.

In the first part we investigate the resource theory of quantum thermodynamics. First, we report a surprising property of quantum catalysis: We present numerical results indicating that in quantum thermodynamics, contrary to the intuitive understanding of catalysis, any non-equilibrium quantum state, given sufficient dimension, acts as a catalyst for all possible transformations. This result can be naturally extended to several other quantum resources theories, including the resource theories of entanglement and coherence. Second, we address the problem of defining work in quantum thermodynamics. We show that when the work reservoir is explicitly modelled as a quantum system, the effects associated with its bounded spectrum are emergent in the work distribution. These consequences are then shown have implications for the form of the celebrated Jarzynski equality and the free energy formulation of the second law of thermodynamics.

In the second part we provide an operational characterisation of certain nonlocal phenomena. First, we look at “nonclassical teleportation”, a recently introduced concept that can be though of as providing a complete description of the standard teleportation protocol. We show that “nonclassical teleportation” generalizes standard teleportation in the sense that it is the resource responsible for teleporting quantum correlations, instead of quantum states. Second, we discuss “Buscemi nonlocality”, a concept that can be viewed as a generalization of Bell nonlocality. Notably, we demonstrate that "Buscemi nonlocality" is a resource that allows two distant parties to measure distributed quantum states in a way that could never be achieved using only classical resources.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Paul Skrzypczyk. I want to particularly thank him for his amazing vigour and passion about learning new things, something he always brings to our meetings. For showing me how to think broadly, and in the same time for being an absolutely excellent teacher, willing to spend many hours to help me understand smaller problems. For giving me the freedom to explore different areas of quantum information, and for his patience when some of these explorations led to the shallows. For helping me develop my problem solving skills, and in the same time for always being there to help me with his time and invaluable guidance. Finally, I would like to thank him for being much more than just a supervisor, but a genuine "scientific father" who helped me make my first steps in the academic world.

In a very special way I would like to thank Sandu Popescu and Tony Short for always being there with a good piece of advice, and for their zealous willingness to explain problems clearly and pedagogically. Having these amazing researchers next door was an absolutely exceptional experience. I would also like to thank Michal Horodecki for his endless fervour in discussing new problems, for his ability to answer questions much faster than I could think about them and for teaching me how to use imagination as a crucial problem-solving skill. I would like to thank them all for sharing their excellent scientific experience, and in the same time for remaining humble and extremely friendly people.

No thesis is an island, and this thesis in particular is the product of multiple fruitful collaborations. I sincerely thank my co-authors: Andrés Ducuara, Michal Horodecki, Pawel Mazurek, Tom Purves and Paul Skrzypczyk. It has been a great pleasure and privilege to collaborate and co-author papers with these wonderful people, and I hope that there will be many more to follow. I profited a lot from our discussions, both related and those unrelated to physics. I want to further extend my gratitude to Edgar Aguilar, Anubhav Chaturvedi, Maté Farkas, Michal Horodecki, Pawel Mazurek, Debashis Saha, Roberto Salazar Vargas, and all people working in KCIK and ICTQT in Gdansk for their kindness and hospitality during my visits.

I would also like to thank all of my fellow collegues from the Bristol QIT group: Jan Lukas Bosse, João Doriguello, Andrés Ducuara, Ryan Mann, Lana Mineh, Alex Moylett, Stephen Piddock, Tom Purves, Changpeng Shao and Dominic Verdon. I would like to also thank Ashley Montanaro and Noah Linden, two excellent researchers whose ability to ask deep and meaningful questions during our QIT seminars always amazed me.

Almost finally, I would like to thank little Luke. During these couple of years I have learnt a lot about physics, quantum information and the theoretical life of a physicist. But this all seems so small compared to how much this little one taught me about myself, and about everything else outside myself. His genuine passion to learn new things, his stubbornness to do things by himself, his unceasing determination to make his first steps in everything, despite so many falls, taught me a lot, and definitely much more than I would ever expect.

Finally, my deepest thank you goes to Alicja. Her love and support, her constant willingness to follow me on the rocky academic path, her fantastic ability to lift me up in the most difficult moments has been, and always be, a great blessing for me. I am totally fortunate to have you by my side. Thank you deeply from my heart.

Contents

List of Figures

List of Symbols

General

$\mathbb{N}, \mathbb{R}, \mathbb{C}$	Natural, real and complex numbers
S	A quantum system
d_S	Dimension of system S

Acronyms

CPTP	Completely positive and trace-preserving
POVM	Positive operator-valued measurement
i.i.d	Independent and identically distributed
CQ	Classical-quantum
QRT	Quantum resource theory
ESD	Entanglement-assisted state discrimination
TSD	Teleportation-assisted state discrimination
DSD	Distributed state discrimination
TO	Thermal operation
TI	Translational invariance
ETI	Effective translational invariance
RoE	Robustness of entanglement
RoT	Robustness of teleportation
RoBN	Robustness of Buscemi nonlocality

Linear algebra

$\mathcal{H}(S)$	Finite-dimensional Hilbert space associated with system S
$\mathcal{L}(S)$	Set of bounded linear operators acting on $\mathcal{H}(S)$
$\mathcal{D}(S)$	Set of quantum states (density operators) acting on $\mathcal{H}(S)$
$X^\dagger$	Adjoint (or Hermitian conjugate) of an operator X
$\mathbb{1}^S$, $\mathbb{1}_S$	Identity operator on $\mathcal{H}_S$
$\mathcal{E}^\dagger : \mathcal{L}(B) \to \mathcal{L}(A)$	Adjoint of a linear map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$
$\text{supp}(X)$	Support of $X \in \mathcal{L}()$

Norms, distances and entropies

$\|\cdot\|_1$	Trace norm (1-norm)
$\mathcal{B}_\epsilon(\rho)$	Ball in trace norm of radius ϵ centered at ρ
$D(\cdot\|\|\cdot)$	Quantum relative entropy
$D_\alpha(\cdot\|\|\cdot)$	(sandwitched) Quantum Renyi divergence
$V(\cdot\|\|\cdot)$	Quantum information variance
$S(\cdot)$	von Neumann entropy
$H(\cdot)$	Shannon entropy

Dla Alicji

Chapter 1

Introduction

1.1 Some initial remarks

In this chapter we discuss the structure of this thesis and briefly summarize the content of each chapter. Then in Chapter 2 we review the mathematical framework relevant for our purposes. These chapters do not contain any original results; Where appropriate, they are accompanied with references to material that significantly expands the presented concepts. In Chapters 3 up to 6 we present original work that was published by the author and his co-authors in articles listed in Sec. 1.3. For better clarity of presentation these chapters have the following unified structure:

1. **Introduction**, where we briefly sketch the background, motivate our work, describe our particular contribution and display the chapters detailed structure,
2. **Details**, where we mention additional assumptions and specific details of the particular framework we use,
3. **Results**, where we present our approach towards solving the problems described in the Introduction,
4. **Discussion and open problems**, where we briefly discuss presented results, their implications, and highlight relevant open problems.
5. **Proofs**, where we give the more technical details of the results described in the chapter.

The most important notation, symbols and abbreviations are displayed in the table on page xv. Please refer to this table when notation becomes unclear at any point.

This main part of this thesis can be divided into two parts. Chapter 3 and Chapter 4 address problems related to the so-called resource theory of quantum thermodynamics. Chapters 5 and 6 address the theory of quantum nonlocality using a resource-theoretic approach. In Fig. 1.1 we display this structure in a graphical way.

To emphasize certain parts of this thesis we will use three types of color boxes. These are described below.

Main question

This will be used to highlight the most relevant questions that we will try to answer in this thesis.

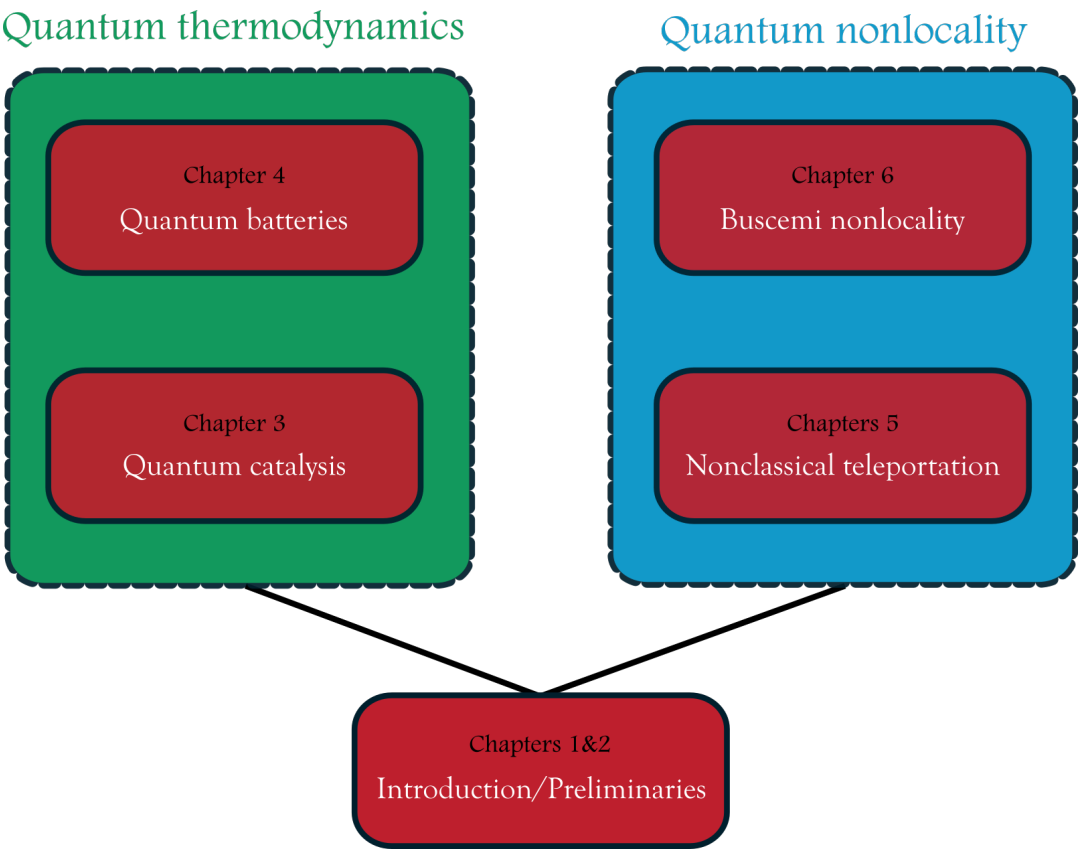

FIGURE 1.1: A schematic outline of this thesis.

Result

Result 1.1. This will be used to differentiate our contribution from the results already established in the literature.

Example

This will be used to highlight explicit examples that will (hopefully) help understand some of the concepts we discuss.

1.2 Outline

Below we briefly outline the content of each chapter.

Chapter 1: Introduction

In this chapter we discuss this thesis logistics. In particular, we explain its structure and list which parts constitute the original contribution of the author and his co-authors, and which parts contain a reviewed material. We also list the papers that constitute the backbone of this thesis.

Chapter 2: Preliminaries

This chapter introduces mathematical concepts necessary for addressing the more specific problems studied in the remaining part of the thesis. In particular, we remind some basic concepts of linear algebra that are widely used in quantum information. We then briefly review the basic toolkit of convex optimization and related concepts of convex analysis. After that we give a very short overview of entanglement and quantum nonlocality, restricting only to concepts which will be used thoughout the thesis. We finish by describing the framework of quantum resource theories.

Chapter 3: Catalytic universality

Quantum catalysts make previously impossible transformations possible. These special quantum resources such as ancillary entangled states open up new possibilities for manipulating objects without consuming or degrading the new resource. Moreover, quantum catalysts can be reused indefinitely, making them highly desirable resources.

Our interactions with the macroscopic world suggest that quantum catalysts must, at the very least, be finely tuned for a particular purpose. This intuition comes from our everyday experience: Since quantum catalysts model the behavior of thermal machines or ancillary experimental instruments, they need to be carefully tuned before they can aid in performing the desired transformation. However, the quantum world often behaves contrary to our intuitionand quantum catalysts are no different.

In this chapter we describe a numerical analysis that draw from quantum thermodynamics to conjecture that all states can be used as universal catalysts, given sufficient dimension. In this way, quantum catalysis is seen to be completely different from macroscopic catalysis, with its own rich structure, much of which is still to be understood.

Chapter 4: Quantum batteries with a ground state

In stochastic thermodynamics work is a random variable whose average is bounded by the change in the free energy of the system. In most standard treatments, however, the work reservoir that absorbs this change is either tacitly assumed or modelled using unphysical systems with unbounded Hamiltonians (i.e. the ideal weight) that assume a certain property called translational symmetry.

In this chapter we study a thermodynamic model with an explicit work reservoir (battery). In particular, we describe the consequences of introducing the battery's ground state and hence – of breaking its translational symmetry. The most striking consequence of this shift is the fact that the Jarzynski identity is replaced by a family of inequalities. Using these inequalities we obtain corrections to the second law of thermodynamics which vanish exponentially with the distance of the initial state of the battery to the bottom of its spectrum. Finally, we study an exemplary thermal operation which realizes the approximate Landauer erasure and demonstrate the consequences which arise when the ground state of the battery is explicitly introduced. In particular, we show that occupation of the ground state of any physical battery sets a lower bound on fluctuations of work. In the same time batteries with a Hamiltionian unbounded from below lead to certain unrealistic predictions, i.e. allow for a fluctuation-free erasure.

Chapter 5: Operational significance of nonclassical teleportation

Quantum teleportation is one of the most thought-provoking discoveries in the whole field of quantum information. The standard benchmark for assessing the quality of teleportation is the so-called average fidelity of teleportation. Importantly, according to this benchmark, some entangled states are not useful for teleportation. This means that they do not allow for teleporting quantum states better than some classical, i.e. separable, state. It was recently shown, however, that if one tries to describe the teleportation experiment more carefully, then all entangled states can lead to nonclassical effects appearing during the teleportation protocol. This means that there is no classical scheme able to reproduce the states teleported to Bob.

This chapter is devoted to studying the operational significance of this result. On the one hand, we demonstrate that every entangled state is useful for the task of teleporting quantum correlations, a protocol whose natural benchmark is the generalization of the average fidelity of teleportation. On the other hand, we show the strength of a particular entangled state and entangled measurement for teleportation - as quantified by the so-called robustness of teleportation - precisely characterizes their ability to offer an advantage in the task of subchannel discrimination with side information. This connection allows us to prove that every entangled state outperforms all separable states when acting as a quantum memory in this discrimination task. Finally, we describe quantum teleportation in the context of a resource theory of teleportation and show that these two operational tasks provide complete sets of monotones for two partial orders based on the notion of "teleportation simulation", one classical and one quantum.

Chapter 6: Operational significance of Buscemi nonlocality

Bell nonlocality is at the heart of the counterintuitive behavior of quantum mechanics. At the same time, it is also the building block of a large number of promising quantum technologies, like quantum cryptography or verification of quantum devices. Depending on the specific physical context, the statistics obtained in a Bell experiment can be viewed as a resource that can be utilized in various ways, providing genuine quantum advantages in a range of quantum information-processing tasks.

In recent years, a broader notion of nonlocality has been proposed (Buscemi nonlocality), one that can be understood as arising from the quantization of the standard Bell scenario. It has remained unclear, however, whether this extended notion can also have a direct operational significance, in a manner similar to standard Bell nonlocality. Such an operational characterization is of both theoretical and practical significance, as not only would it allow for a better understanding of quantum nonlocality, but also show its suitability for addressing practical problems.

Our main result demonstrates that Buscemi nonlocality can also be viewed as a resource, one that allows two distant parties to measure distributed quantum states in a way that could never be achieved using only classical resources. We also show, in a quantitatie way, how Buscemi nonlocality relates with nonclassical teleportation and entanglement. Finally, we show that Buscemi nonlocality is directly related to the classical single-shot capacity of an arbitrary bipartite quantum-to-classical channel.

1.3 List of publications

During the time of my PhD, I had the great privilege to collaborate with many excellent researchers. The new results presented in this thesis are based on the material contained in the following papers.

1. Patryk Lipka-Bartosik and Paul Skrzypczyk. "All States are Universal Catalysts in Quantum Thermodynamics". In: *Phys. Rev. X* 11 (1). DOI: 10.1103/PhysRevX.11.011061 [Chapter 3]

2. Patryk Lipka-Bartosik, Paweł Mazurek, and Michał Horodecki. "Second law of thermodynamics for batteries with vacuum state". In: *Quantum* 5. DOI: 10.22331/q-2021-03-10-408 [Chapter 4]

3. Patryk Lipka-Bartosik et al. "Operational Significance of the Quantum Resource Theory of Buscemi Nonlocality". In: *PRX Quantum* 2 (2). DOI: 10.1103/PRXQuantum.2.020301 [Chapter 5]

4. Patryk Lipka-Bartosik and Paul Skrzypczyk. "Operational advantages provided by nonclassical teleportation". In: *Phys. Rev. Research* 2 (2). DOI: 10.1103/PhysRevResearch.2.023029 [Chapter 6]

Moreover, during this time I have also co-authored the following papers and preprints which are not part of this thesis:

1. Andrés F. Ducuara, Patryk Lipka-Bartosik, and Paul Skrzypczyk. "Multiobject operational tasks for convex quantum resource theories of state-measurement pairs". In: *Phys. Rev. Research* 2 (3). DOI: 10.1103/PhysRevResearch.2.033374

2. Patryk Lipka-Bartosik and Paul Skrzypczyk. *Catalytic quantum teleportation*. arXiv: 2102.11846 [quant-ph]

Chapter 2

Preliminaries

In this chapter we introduce the essential mathematical tools that will be used to prove most of the results described in this thesis. The goal of this short introduction is to remind about the most important technical concepts used in this thesis and to introduce the relevant notation.

In Section 2.1 we describe the elementary formalism of quantum information. In Section 2.2 we remind the toolbox of convex optimisation theory, including the basics of semidefinite programming and duality. In Sec. 2.3 and Sec. 2.4 we briefly mention the very basic concepts of entanglement theory and quantum nonlocality. Finally, in Section 2.5 we remind the basic ingredients of quantum resource theories, with a particular focus on the resource theory of quantum thermodynamics.

2.1 The formalism of quantum information

2.1.1 Basics of linear algebra

A Hilbert space $\mathcal{H}$ is a complex vector space equipped with an inner product. In this thesis we will deal exclusively with Hilbert spaces that are finite-dimensional. We will often associate Hilbert spaces with physical systems labelled with capital letter, i.e. $\mathcal{H}(S)$ corresponds to a Hilbert space associated with a quantum system S. Moreover, we will omit the label when the association is clear from the context.

We will make extensive use of Dirac's bra-ket notation. In this convention vectors in $\mathcal{H}$ are denoted with $|\cdot\rangle \in \mathcal{H}$, where the label inside the ket unambiguously identifies a given vector. With each Hilbert space $\mathcal{H}$ we associate a dual space of a linear functional $\mathcal{H}^*$ spanned by dual vectors $\langle\cdot|$. Because $\mathcal{H}$ is finite-dimensional, it is isomorphic to its dual space $\mathcal{H}^*$. Consequently, each vector $|u\rangle \in \mathcal{H}$ has a corresponding dual denoted with $\langle u| \in \mathcal{H}^*$.

The complex conjugate of a complex scalar $z \in \mathbb{C}$ is denoted with z^*. The inner product of the Hilbert space is denoted $\langle\cdot|\cdot\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$. The inner product is Hermitian, meaning that it satisfies $\langle u|v\rangle = \langle v|u\rangle^*$ for all $u, v \in \mathcal{H}$ and is linear in one of its arguments. A natural norm $\|\cdot\|$ can be defined for the Hilbert space $\mathcal{H}$ as $\|u\| := \sqrt{\langle u|u\rangle}$. An orthonormal basis $\{|i\rangle\}$ is a collection of linearly independent unit vectors $|i\rangle \in \mathcal{H}$ which span the whole Hilbert space.

Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ and corresponding vectors $\big|\psi\big\rangle \in \mathcal{H}$ and $\big|\psi'\big\rangle \in \mathcal{H}'$ we denote a linear operator that maps $\big|\psi\big\rangle$ into $\big|\psi'\big\rangle$ using the convenient notation $\big|\psi'\big\rangle\big\langle\psi\big|$.

Let $A : \mathcal{H} \to \mathcal{H}$ be a linear operator acting on $\mathcal{H}$. Given a basis $\{|i\rangle\}$ of $\mathcal{H}$ we can describe its action using its matrix representation, i.e. by constructing a matrix with elements $A_{ij} = \big\langle i\big|A\big|j\big\rangle$. Consequently, the operator A can be written as $A = \sum_{i,j} A_{ij} \big|i\big\rangle\big\langle j\big|$. The space of all linear operators on $\mathcal{H}(S)$ is denoted with $\mathcal{L}(S)$.

A linear operator A is positive semidefinite, denoted $A \geq 0$, if $\langle\psi|A|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$.

The adjoint of a linear operator A is the unique operator $A^\dagger$ satisfying $\langle i|A^\dagger|j\rangle = \langle j|A|i\rangle^*$ for all $|i\rangle$ and $|j\rangle$.

2.1.2 Quantum states

Quantum states are described by density operators, that is, positive semi-definite operators with a unit trace, i.e.

$$\rho \geq 0, \qquad \mathrm{tr}[\rho] = 1. \tag{2.1}$$

A density operator is necessarily Hermitian, and has eigenvalues between zero and one that sum to one. We will frequently denote density operators with greek letters, e.g. ρ, σ or γ.

The set of all density operators associated with a Hilbert space $\mathcal{H}(S)$ will be denoted by $\mathcal{D}(S)$. This is a convex set, meaning that any convex combination of density operators $\{\rho_i\}$, that is an object of the form $\rho = \sum_i p(i)\rho_i$ for some probability distribution $p(i)$, is also a valid density operator.

When its clear from context, we will omit writing explicitly that $\rho \in \mathcal{D}(S)$. When we want to emphasize that a quantum state is associated with a specific quantum system, say S, we will write ρ_S meaning $\rho \in \mathcal{D}(S)$.

Density operators form a convex set in the space of all positive-semidefinite operators. This means that any density operator ρ_S can be written as a convex combination of pure states $|i\rangle_S$ in $\mathcal{H}(S)$, that is

$$\rho_S = \sum_i p(i)\,|i\rangle\langle i|_S\,, \tag{2.2}$$

for some probability distribution $p(i)$. Some classes of quantum states have their own names due to their special properties. In what follows we display some of the widely used conventions.

Pure states. A quantum state ρ is called *pure* if its rank is equal to one. This means that there is a vector $|\psi\rangle$ in the system's Hilbert space, such that $\rho = |\psi\rangle\langle\psi|$.

Mixed states. Quantum states which are not pure are called *mixed states*.

Classical states. Classical probability distributions can be modeled as quantum density operators that are diagonal in a fixed orthonormal basis. Therefore, a quantum state ρ is said to be a *classical state* with respect to a basis $\{|i\rangle\}$ if it is diagonal in that basis.

2.1.3 Quantum dynamics

The evolution of an open quantum system from $\mathcal{L}(A)$ to $\mathcal{L}(B)$ can be described by a quantum channel, which is a linear map that is both completely positive (CP) and trace-preserving (TP). We will denote quantum channels with caligraphic letters, for example $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$. The complete positivity not only ensures that the channel maps positive operators into positive operators, but also that this remains true when

the evolving system is a part of some larger system. When we describe channel-state duality we will see a more compelling argument for why this should be the case.

Quantum channels can be represented in several equivalent ways. In what follows we highlight the most frequently used representations.

Theorem 2.1 (Quantum channel representations)**.** A linear map $\mathcal{E}$ from $\mathcal{L}(A)$ to $\mathcal{L}(B)$ is completely positive and trace-preserving if and only if any one of the following holds:

(i) (Kraus representation) There exists a set of linear operators $\{E_i\}$ called *Kraus operators* acting from $\mathcal{L}(A)$ to $\mathcal{L}(B)$ which for all $\rho \in \mathcal{D}(A)$ satisfy

$$\mathcal{E}[\rho] = \sum_i E_i \rho E_i^\dagger \qquad \text{and} \qquad \sum_i E_i^\dagger E_i = \mathbb{1}. \tag{2.3}$$

(ii) (Stinespring representation) There exists a system E with an associated Hilbert space $\mathcal{H}(E)$ and an isometry V_{AE} such that for all $\rho \in \mathcal{D}(A)$

$$\mathcal{E}(\rho) = \mathrm{tr}_E \left[V_{AE} \rho_A V_{AE}^\dagger \right]. \tag{2.4}$$

(iii) (Choi-Jamiołkowski representation) Given a matrix $J_\mathcal{E}$ (Choi matrix) defined as

$$J_\mathcal{E} = \sum_{i,j} |i\rangle\langle j|_A \otimes \mathcal{E}\left[|i\rangle\langle j|_{A'} \right], \tag{2.5}$$

where $\{|i\rangle_A\}$ and $\{|i\rangle_{A'}\}$ are two orthonormal bases for isomorphic Hilbert spaces $\mathcal{H}(A)$ and $\mathcal{H}(A')$, the following holds

$$J_\mathcal{E} \geq 0 \qquad \text{and} \qquad \mathrm{tr}_{A'} J_\mathcal{E} = \mathbb{1}_A. \tag{2.6}$$

Furthermore, the map $\mathcal{E}$ can be recovered from $J_\mathcal{E}$ in the following way

$$\mathcal{E}(\rho) = \mathrm{tr}_A[(\rho_A^T \otimes \mathbb{1}_{A'}) J_\mathcal{E}]. \tag{2.7}$$

We now move on to a special class of quantum channels that can be used to convert quantum information encoded in quantum states into classical information that can be described using only classical states.

2.1.4 Quantum measurements

Quantum measurements are a particular class of quantum channels that describe the most general way of extracting classical information from a quantum system. The measurement channel acts on a quantum system and outputs two types of information: a post-measurement quantum state that encodes the back-action of the measuring apparatus, and a classical state that encodes the outcome of the measurement.

More specifically, for measurements with discrete set of outcomes $a \in \{1, 2, \ldots, o_A\}$, a quantum state ρ is transformed into another quantum state σ_a with some probability $p(a)$. Quantum mechanics predicts that the post-measurement state of the measured system and the classical register can be described by the following state

$$\mathcal{I}[\rho] = \sum_a E_a \rho E_a^\dagger \otimes |a\rangle\langle a|, \tag{2.8}$$

where indices $a \in \{1, \ldots, o_A\}$ indicate the outcomes of the measurement appearing with probability $p(a) := \mathrm{tr}\left[E_a \rho E_a^\dagger\right]$ and $\{E_a\}$ are Kraus operators describing the action of the measuring apparatus on the quantum system. In general, we can even consider situations where each outcome a corresponds to a different set of Kraus operators; As here we will always associate a single Kraus operator with each measurement outcome, will not use this more general description. Formally, we say that the quantum channel of the form (2.8) is a *quantum instrument*. We will use this convention to emphasize when we are interested in the post-measurement state of the measured system. On the other hand, we will use the term *quantum measurement* when we are interested only in the measurement statistics $p(a)$. This process can be formally described using a completely positive and trace-preserving map of the form:

$$\mathcal{M}[\rho] = \mathrm{tr}_S\, \mathcal{I}[\rho] = \sum_a \mathrm{tr}\left[M_a \rho\right] |a\rangle\langle a| \,, \tag{2.9}$$

where $M_a := E_a^\dagger E_a$. Quantum measurement can be therefore interpreted as a quantum instrument followed by discarding the post-measurement state. A quantum measurement can be completely described using operators $\{M_a\}$. A collection of positive operators $\{M_a\}$ satisfying $\sum_a M_a = \mathbb{1}$ is called a POVM (positive operator-valued measurement) and provides a complete description of the measurement process when we are only interested in the measurement outcomes.

2.1.5 Distinguishing quantum states

Some pairs of quantum states can be distinguished quite easily. For instance, imagine that we have two quantum devices, one that outputs a quantum state $\omega_0 = |0\rangle\langle 0|$ and a second one that outputs $\omega_1 = |1\rangle\langle 1|$. The problem is that we do not a priori which device outputs which state. Clearly, a simple projective measurement in the basis $\{|0\rangle, |1\rangle\}$ applied on the state produced by any of these devices yields an outcome that can reveal which device we have at hand. In contrast, if the devices produce $\omega_0 = |0\rangle\langle 0|$ and $\omega_1 = (1-\epsilon)\,|0\rangle\langle 0| + \epsilon\,|1\rangle\langle 1|$ for some small and positive ϵ, the task becomes more difficult. This is because applying the same projective measurement as before will yield an outcome $|0\rangle$ with high probability, and thus will not be able to tell us which state was produced by the device. It turns out that in this simple case no measurement, however complex, will ever be able to give us a better answer. This provides an intuitive feeling that these two states are, in some sense, close to each other.

Let us now approach this more formally. Consider again two devices, one that outputs ω_0 and another one that outputs ω_1. Our goal is to determine which of the devices we have by applying a general quantum measurement $\mathbb{M} = \{M_x\}$ to the state produced by the device. Suppose that our prior probability of which device we have is given by some distribution $p(x)$, where outcome $x = 0$ is associated with ω_0 and $x = 1$ with ω_1. Then, given the results of the experiment we perform, we would like the probability of correctly guessing x to be as large as possible. This probability can be expressed as:

$$p_{\mathrm{guess}}(X) = \max_{\mathbb{M} \in \mathrm{POVM}} \sum_x p(x)\, \mathrm{tr}[M_x \omega_x]. \tag{2.10}$$

Since $x \in \{0,1\}$ and $p(0) = p(1) = 1/2$, we can further write $M_0 = X$ and $M_1 = \mathbb{1} - X$ for some operator $0 \leq X \leq \mathbb{1}$. With this Eq. (2.10) becomes:

$$p_{\text{guess}}(X) = \frac{1}{2} + \frac{1}{2} \max_{0 \leq X \leq \mathbb{1}} \text{tr}[X(\omega_0 - \omega_1)]. \tag{2.11}$$

We can regard the second term as a measure of distinguishability. If the two states are disjoint (have disjoint support), then we can guess perfectly, which is reflected in the fact that the second term becomes one. If they have an overlapping support then the second term is smaller, meaning that the probability of guessing is closer to the a priori distribution $p(x)$. This motivates introducing the following definition.

Definition 2.1 (Trace distance)
Let ρ and σ be quantum states in $\mathcal{D}$. The trace distance, $\|\rho - \sigma\|_1$, is defined as:

$$\|\rho - \sigma\|_1 = \max_X \quad 2\,\text{tr}\big[X(\rho - \sigma)\big] \tag{2.12}$$

$$\text{subject to} \quad 0 \leq X \leq \mathbb{1}. \tag{2.13}$$

A useful fact is that trace distance is a semi-definite optimisation problem, meaning that it can be efficiently computed numerically. We will discuss this type of problem in more detail In Sec. 2.2. Now as we have a well-motivated notion of a distance between quantum states, we can introduce the concept of a ball for density operators. In particular, we will denote

$$\mathcal{B}_\varepsilon(\rho) = \{\rho' : \|\rho' - \rho\|_1 \leq \varepsilon\}. \tag{2.14}$$

Therefore the set $\mathcal{B}_\varepsilon(\rho) \subseteq \mathcal{D}$ contains all quantum states that cannot be distinguished from ρ with a precision better than ϵ.

Importantly, the reasoning presented above provides an *operational characterisation* of the trace distance. By this we mean that trace distance not only can be viewed as a convenient mathematical concept, but also may be interpreted as a benchmark in a well-defined task. In Chapters 5 and 6 we will use analogous reasoning to establish an operational characterisation of various quantifiers that are relevant for quantum nonlocality.

2.1.6 Entropic quantifiers

In order to quantify information present in a quantum system we look at various information-processing tasks. These are particular protocols, or games, which process information in order to achieve certain goals. Performance in such tasks can be very often benchmarked using the so-called entropic quantifiers. These are functions that can be obtained from a single parent quantity called *quantum relative entropy*.

Definition 2.2 (Quantum relative entropy and quantum information variance)
Let $\rho \in \mathcal{D}$ be a quantum state and $\sigma \geq 0$ be a positive semidefinite operator such that $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$.

(i) The quantum relative entropy is defined as (Umegaki, 1962)

$$D(\rho||\sigma) = \text{tr}\,\rho(\log\rho - \log\sigma). \tag{2.15}$$

(ii) The quantum information variance is defined as (Li, 2014; Tomamichel and Hayashi, 2013)

$$V(\rho||\sigma) = \operatorname{tr}\rho(\log\rho - \log\sigma)^2 - D(\rho||\sigma)^2. \tag{2.16}$$

The quantum relative entropy can be seen as a parent for the following fundamental quantum information-theoretic quantities.

Definition 2.3 (Entropies derived from relative entropy)
Let ρ_{AB} be a quantum state in $\mathcal{D}(AB)$ with marginals ρ_A and ρ_B. Then we define

(i) the *von Neuman entropy*

$$S(A)_\rho \equiv S(\rho_A) = -D(\rho_A||\mathbb{1}_A) = -\operatorname{tr}\rho_A\log\rho_A. \tag{2.17}$$

If we write ρ_A in its eigenbasis, i.e. $\rho_A = \sum_a \lambda(a)\,|\lambda_a\rangle\langle\lambda_a|_A$, then the von Neuman entropy reduces to the *Shannon entropy* of a probability distribution $\lambda(a)$ given by

$$H(A)_\lambda = -\sum_a \lambda(a)\log\lambda(a). \tag{2.18}$$

When $\lambda(a)$ is a binary distribution, i.e. $\lambda(0) = \epsilon$ and $\lambda(1) = 1-\epsilon$, we call the corresponding quantity *binary entropy* and denote

$$h(\epsilon) = -\epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon). \tag{2.19}$$

(ii) the *quantum conditional entropy*

$$S(A|B)_\rho = -D(\rho_{AB}||\mathbb{1}_A\otimes\rho_B) = S(AB)_\rho - S(B)_\rho \tag{2.20}$$

(iii) the *quantum mutual information*

$$I(A|B)_\rho = D(\rho_{AB}||\rho_A\otimes\rho_B) = S(A)_\rho + S(B)_\rho - S(AB)_\rho \tag{2.21}$$

It is worth mentioning that the von Neuman entropy is continous with respect to trace distance, i.e. two states close in trace distance have a similar amount of entropy. This is captured by the following result from (Fannes, 1973; Audenaert, 2007).

Theorem 2.2 (Fannes-Audenaert inequality)**.** Given two quantum states $\rho, \sigma \in \mathcal{D}$ such that $\left\|\rho-\sigma\right\|_1 \leq \epsilon$ and d is the dimension of $\mathcal{H}$, it holds that

$$|S(\rho) - S(\sigma)| \leq \frac{1}{2}\epsilon\log(d-1) + h(\epsilon/2). \tag{2.22}$$

This inequality is sharp, meaning that there exist states ρ and σ for which it can be saturated.

The entropic quantities we have presented so far are useful measures of information in the i.i.d limit, i.e. when the task is carried out using infinitely many independent and identically distributed copies of a quantum system. Recent years have spurred a renewed interest in single instances of information-theoretic tasks, a scenario that is very often much more challenging to analyse. One of the reasons for this problem is that the quantifiers derived from quantum relative entropy are

no longer sufficient to characterise properties of quantum systems in this regime. Consequently, different measures of information have to be employed, and often such measures need to be tailored to the specific information-theoretic tasks. Such entropic quantities can also be derived from a single parent quantity called the *sandwiched Renyi divergence* that was introduced in (Müller-Lennert et al., 2013) and independently in (Wilde, Winter, and Yang, 2014). These single-shot measures are more precise, and also generalize quantum relative entropy in the sense that they all converge to this entropic quantity in the i.i.d. limit.

Definition 2.4 (Sandwiched Renyi divergences)
Let $\rho \in \mathcal{D}$ be a quantum state and $\sigma \geq 0$ be a positive semi-definite operator such that $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$. Then the *sandwiched Renyi divergences* is defined as

$$D_\alpha(\rho||\sigma) = \frac{\text{sgn}(\alpha)}{\alpha - 1} \log \text{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] \tag{2.23}$$

When the state ρ and operator σ from the above definition are classical, they can be diagonalised in the same eigenbasis, for example $\rho = \sum_i p(i) |i\rangle\langle i|$ and $\sigma = \sum_i q(i) |i\rangle\langle i|$. In this case we can write the sandwiched Renyi divergence in a much simpler form, that is

$$D_\alpha(\rho||\sigma) = \frac{\text{sgn}(\alpha)}{\alpha - 1} \log \sum_i p(i)^\alpha q(i)^{1-\alpha}. \tag{2.24}$$

2.2 Convex optimisation

Convex optimisation problems are far more general than the more common linear optimisation problems. In the same time they share the desirable properties of linear problems: in certain cases they can be solved quickly and reliably up to a relatively large size. It turns out that many important problems in quantum information can be formulated as convex optimisation problems. Several excellent textbooks, e.g. (Boyd and Vandenberghe, 2004) and (Watrous, 2018), cover all of the material relevant for us. Here we will briefly highlight the most important concepts that we will use throughout the thesis.

2.2.1 Convexity

We say that a subset of a real vector space $\mathcal{C}$ is convex if the line segment between any two points from the set lies in that set. More formally, we have the following definition.

Definition 2.5 (Convex set)
Let $\mathcal{C}$ be a subset of a real vector space and let x, y be two points inside that subset. We say $\mathcal{C}$ is convex if for all $0 \leq \lambda \leq 1$ we have

$$(1 - \lambda)x + \lambda y \in \mathcal{C}. \tag{2.25}$$

More generally, we can also consider convex combinations of points, that is points of the form $\sum_i \lambda(i) x_i$ where each $x_i \in \mathcal{C}$ and $\lambda(i) \geq 0$ for all i and $\sum_i \lambda(i) = 1$. Equivalently, we can say that a set is convex if it contains all of its convex combinations, that is $\sum_i \lambda(i) x_i \in \mathcal{C}$ for all positive $\lambda(i)$ such that $\sum_i \lambda(i) = 1$. We can think about such a combination as a mixture of different objects x_i from the set $\mathcal{C}$, appearing with respective probabilities $\lambda(i)$.

The notion of convexity can be also addressed as a property of functions defined over a convex set. In that case we have the following definition.

Definition 2.6 (Convex function)
A function $f : \mathcal{C} \to \mathbb{R}$, where $\mathcal{C}$ is a convex set, is *convex* if for all points $x, y \in \mathcal{C}$ and all $0 \leq \lambda \leq 1$, we have:

$$f[(1-\lambda)x + \lambda y] \leq (1-\lambda)f(x) + \lambda f(y). \tag{2.26}$$

Geometrically, this means that the line segment between points $[x, f(x)]$ and $[y, f(y)]$ lies everywhere above the graph of f. We say that a function f is *strictly convex* if strict inequality holds in (2.26) whenever $x \neq y$ and $0 < \lambda < 1$. We also say that f is *concave* if $-f$ is convex, and *strictly concave* if $-f$ is *strictly convex*.

Finally, let us remind the notion of a cone. In particular, for any set $\mathcal{C}$ we say that it is a *cone* if rescalling (multipliying by a scalar) its points does not move them outside of the set. More formally we have the following definition.

Definition 2.7 (Cone)
Let $\mathcal{C}$ be a subset of a real vector space. We say that $\mathcal{C}$ is a *cone* if for every point x inside the set and all $\lambda \geq 0$ we have

$$\lambda x \in \mathcal{C}. \tag{2.27}$$

Furthermore, we say that a subset of a real vector space is a *convex cone* if every linear combination of its points lies in that set, that is for all $\lambda(i) \geq 0$, not necessarily normalised, we have

$$\sum_i \lambda(i) x_i \in \mathcal{C}. \tag{2.28}$$

With these definitions we can now describe some relevant concepts from the theory of convex optimisation.

2.2.2 Semidefinite optimisation

General optimisation problems have the following form:

$$\begin{aligned} \inf \quad & f(x), && (2.29)\\ \text{subject to} \quad & x \in \mathcal{X}, && (2.30) \end{aligned}$$

We will refer to optimisation problems as *programs*. When the set $\mathcal{X}$ and the objective function $f(x)$ are both convex, the problem is said to be a *convex optimisation problem*. Solving such problems in general can be difficult, as without any further information about $\mathcal{X}$ not much can be done to make the problem more tractable.

Semidefinite programming is a subfield of convex optimization concerned with optimizing a linear objective over the intersection of a cone of positive semidefinite matrices with an affine space. In the last decades, semidefinite programs (or SDPs for short) have received a renewed interest and became a standard tool for solving engineering, combinatorial and mathematical problems. This is mainly due to the emergence of efficient solvers that allow to work out relatively complex SDP problems in a time that is polynomial in the size of the input. Moreover, many important problems can be relaxed and formulated as SDPs, therefore providing a useful analytical tool for giving insight into apparently complex problems.

A semidefinite optimisation problem, or semidefinite program, is specified by two Hermitian matrices A and B and a Hermiticity-preserving linear map $\mathcal{E}$. The optimisation task is to find a positive semidefinite Hermitian matrix $X \in \mathcal{C}$ which is a solution to the following convex optimisation problem:

$$\begin{aligned} \alpha = \quad & \inf_{X \in \mathcal{C}} \quad && \mathrm{tr}[AX] && (2.31) \\ & \text{subject to} \quad && \mathcal{E}(X) \leq B, && (2.32) \\ & && X \geq 0. && (2.33) \end{aligned}$$

Many relevant problems in quantum information can be formulated as semidefinite problems.

Duality

Every feasible X gives an upper bound on the value of the semidefinite program (2.31). In order to find upper bounds, we can add a non-negative term to the optimisation function in Eq. (2.31), which can only lead to a larger (or equal) value. Introducing positive semidefinite operators $Y \geq 0$ and $Z \geq 0$ we can infer that for all feasible X the function:

$$\begin{aligned} \mathcal{L}(X, Y, Z) &= \mathrm{tr}[AX] + \mathrm{tr}[Y(B - \mathcal{E}(X))] + \mathrm{tr}[XZ] && (2.34) \\ &= \mathrm{tr}[AX] + \mathrm{tr}[BY] - \mathrm{tr}\left[\mathcal{E}^\dagger X\right] + \mathrm{tr}[XZ] && (2.35) \\ &= \mathrm{tr}\left[X(A - \mathcal{E}^\dagger(Y) + Z)\right] + \mathrm{tr}[BY] && (2.36) \end{aligned}$$

is always larger or equal to the optimal value of the primal problem α. This can be easily seen by choosing a particular set of dual variables, i.e. $Y = Z = 0$. Defining $\beta := \max_{Y,Z \geq 0} \mathcal{L}(X, Y, Z)$ leads to the following *dual problem*:

$$\begin{aligned} \beta = \quad & \sup \quad && \mathrm{tr}[BY] && (2.37) \\ & \text{subject to} \quad && \mathcal{E}^\dagger(Y) \leq A, && (2.38) \\ & && Y \geq 0. && (2.39) \end{aligned}$$

The significance of the duality of convex optimisation problems in analytical calculations is twofold. First, it is a convenient (analytical!) tool that allows for computing useful bounds for various information-theoretic tasks. Second, it can be viewed as a simple method of reformulating complex problems. In particular, it is very often the case that problems which may look difficult in their primal form become much more intuitive when written in their dual forms. In Chapter 5 and Chapter 6 we will see that the dual formulation of several information-theoretic measures often reveals their operational character, i.e. can be interpreted as a benchmark of the performance in some information-processing task.

Weak and strong duality

The primal (2.31) and the dual problem (2.37) are both closely related. It follows directly from the definition of the dual that:

$$\alpha \geq \beta. \tag{2.40}$$

This relationship is known as *weak duality* and is valid for any pair of problems defined as in Eq. (2.31) and Eq. (2.37). This simple observation gives us a way of

finding simple bounds for the optimal value attained by one of the two problems, by choosing a feasible candidate in the other.

For some problems we can find an even stronger relationship between the solutions of the primal and the dual problems. More specifically, under certain mild conditions the optimal values of both problems coincide, leading to *strong duality*, that is when

$$\alpha = \beta. \tag{2.41}$$

A sufficient set of conditions for strong duality to hold is know as Slater's conditions (Watrous, 2018) and can be summarised as follows.

Definition 2.8 (Slater's conditions)
For any semidefinite program written as in Eq. (2.31) with a dual written as in Eq. (2.37) strong duality holds when either of the following is true:

(i) The primal problem is feasible and the dual problem is strictly feasible. Then there exists a choice for X such that $\mathrm{tr}[AX] = \alpha$.

(ii) The dual problem is feasible and the primal problem is strictly feasible. Then there exists a choice for Y such that $\mathrm{tr}[BY] = \beta$.

(iii) Both problems are strictly feasible. Then there exist choices for X and Y such that $\alpha = \beta = \mathrm{tr}[AX] = \mathrm{tr}[BY]$.

Even though we have described the notion of duality from the perspective of semidefinite programs, an analogous reasoning is also valid for general convex optimisation problems. In the following chapters we will make extensive use of this fact and use duality as a convenient tool of characterising problems that cannot be written in a semidefinite form.

2.3 Entanglement theory

The existence of quantum entanglement is one of the most striking consequences implied by the laws of quantum mechanics (Einstein, Podolsky, and Rosen, 1935; Bell, 1964). It is manifested when correlations between different particles are strong enough so that the action on one of them affects the other, in a way so subtle that it cannot be explained by any classical mechanism. Although entanglement was initially recognized as a bizarre property separating quantum from classical physics, it is nowadays viewed as an indispensable resource with an enormous number of modern applications.

One of the first results showing how entanglement can be processed and consumed to provide a useful advantage was the cryptographic scheme proposed by Ekert in (Ekert, 1991). In the next years it was realized that this is not a singular case, and in fact entanglement can be used to fuel other quantum information processing tasks as well. Several years later Shor presented a factorisation algorithm that outperformed the best classical algorithm by using entanglement (Shor, 1997). Also novel communication protocols such as teleportation of quantum states (Bennett et al., 1993) and superdense coding (Bennett and Wiesner, 1992) were proposed, all of them showing how entangled states can be seen as valuable resources.

Entangled quantum states are the basic fuel for all of the above tasks. To make this statement precise, and to understand how entanglement is utilised in these tasks, it is necessary to develop a mathematical theory of entanglement, a theory that could

answer the questions: when is a quantum system in an entangled state, and how much entanglement is required for a specific information-processing task?

The simplest scenario allowing for studying entanglement consists of two subsystems, often denoted Alice (A) and Bob (B) that are spatially separated. Such a system can be described using a composite Hilbert space $\mathcal{H}(AB)$ that can be defined in the following way.

Definition 2.9 (Composite Hilbert space)
The Hilbert space $\mathcal{H}(AB)$ of a quantum system composed of two subsystems A and B, each of them described using individual Hilbert spaces $\mathcal{H}(A)$ and $\mathcal{H}(B)$, is given by

$$\mathcal{H}(AB) = \mathcal{H}(A) \otimes \mathcal{H}(B) = \text{span}\left\{ |i\rangle_A \otimes \left|j\right\rangle_B \,\middle|\, i, j \right\}, \tag{2.42}$$

where i and j range over all elements from the basis of $\mathcal{H}(A)$ and $\mathcal{H}(B)$ respectively.

2.3.1 Bipartite quantum states

States on the composite Hilbert space $\mathcal{H}(AB)$ are called *bipartite states* and exhibit a surprisingly rich structure. In what follows we will summarise the widely used terminology due to (Schrödinger, 1936) (in the case of pure states) and (Werner, 1989) (in the case of mixed states).

Definition 2.10
A quantum state ρ_{AB} is called a *separable state* if it can be written in the form:

$$\rho_{AB} = \sum_i p(i) \rho_A^{(i)} \otimes \rho_B^{(i)}, \tag{2.43}$$

for some ensembles of states $\{\rho_A^{(i)}\}$ and $\{\rho_B^{(i)}\}$ and a probability distribution $p(i)$. Separable states can only share classical correlations and can be further subdivided into:

(i) **product states**, when the mixture contains only one element, meaning the system is a tensor-product of two independent states. Such states share no correlations.

(ii) **classical states**, when both ensembles of states share a common eigenbasis. In other words, $\rho_A^{(i)} = |i\rangle\langle i|_A$ and $\rho_B^{(i)} = |i\rangle\langle i|_B$ for some orthonormal bases $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$.

(iii) **classical-quantum** (CQ) states, when one of the systems is classical with respect to some basis, i.e. $\rho_A^{(i)} = |i\rangle\langle i|_A$ while $\rho_B^{(i)}$ are arbitrary.

Importantly, separable states form a convex subset of all density operators.

A quantum state is said to be *entangled* when it is not separable, i.e. cannot be written as in Eq. (2.43).

2.3.2 Bipartite quantum operations

Let us now discuss different types of bipartite quantum operations, that is quantum channels that can be performed in the bipartite scenario. In general we would like to study quantum channels that obey certain realistic constraints, e.g. have unlimited access only to classical communication. In what follows we will assume that two

distant observers, Alice and Bob, are in possession of a system prepared in a bipartite quantum state ρ_{AB}.

Local operations (LO)

A bipartite operation $\mathcal{E}_{AB}$ is called a *local operation* (LO) if it can be written as a pair of quantum channels applied separately by Alice and Bob, i.e. $\mathcal{E}_{AB} = \mathcal{E}_A \otimes \mathcal{E}_B$ for some local quantum channels $\mathcal{E}_A$ and $\mathcal{E}_B$. Local operations can be mathematically described as follows

$$\mathcal{E}_{AB}[\rho] = (\mathcal{E}_A \otimes \mathcal{E}_B)[\rho] = \sum_{i,j} \left(E_i^A \otimes F_j^B\right) \rho_{AB} \left(E_i^A \otimes F_j^B\right)^\dagger, \tag{2.44}$$

where $\{E_i^A\}$ and $\{F_j^B\}$ are Kraus operators associated with local channels $\mathcal{E}_A$ and $\mathcal{E}_B$ respectively.

Local operations and shared randomness (LOSR)

When Alice and Bob apply local operations and are additionally allowed to share classical randomness, the resulting class of bipartite operations is called *local operations and shared randomness*. This type of operations can be mathematically described in the following way

$$\mathcal{E}_{AB}[\rho] = \sum_{\lambda} p(\lambda) \left(\mathcal{E}_\lambda^A \otimes \mathcal{E}_\lambda^B\right)[\rho], \tag{2.45}$$

where $\{\mathcal{E}_\lambda^A\}$ and $\{\mathcal{E}_\lambda^B\}$ are two ensembles of local quantum channels (LO).

Local operations and classical communication (LOCC)

When Alice and Bob apply local operations and, additionally, Alice is allowed to send classical information to Bob, the corresponding bipartite operation is called 1-LOCC($A \to B$). Such operation can be mathematically modelled as

$$\mathcal{E}_{AB}[\rho] = \sum_i \left(\mathcal{I}_i^A \otimes \mathcal{E}_i^B\right)[\rho] \tag{2.46}$$

where $\mathcal{I}^A = \{\mathcal{I}_i^A\}$ is a collection of subchannels, so that $\mathcal{E}^A := \sum_i \mathcal{I}_i^A$ is a quantum channel. Equivalently, the whole protocol can be seen as Alice applying a quantum instrument locally to A with outcome i, sending this outcome to Bob via a classical channel and then having Bob apply a corresponding local quantum channel $\mathcal{E}_i^B$.

If both parties are allowed to communicate with each other for an unlimited number of rounds, then the corresponding bipartite operation is called an LOCC. Every operation of this form can be understood as a sequence of 1-LOCC($A \to B$) and 1-LOCC($B \to A$) operations. It can be readily verified that the following chain of relations hold

$$\text{LO} \subset \text{LOSR} \subset \text{1-LOCC} \subset \text{LOCC}. \tag{2.47}$$

The mathematical structure of LOCC's is complicated and more details can be found in (Chitambar et al., 2014).

2.4 Quantum nonlocality

In the early years of quantum mechanics entanglement was recognized mainly as a bizarre property that distinguished it from classical physics. It was mainly due to the discovery of Bell nonlocality (Bell, 1964) and subsequent development of Bell inequalities that allowed this distinction to be formulated quantitatively and to verify the predictions of quantum theory in an experimentally feasible setting.

In simple terms Bell nonlocality refers to the situation when correlations shared between spatially separated parties cannot be explained as arising from a shared classical source (so called local hidden variable). We refer to such correlations as *nonlocal correlations*. These correlations are one of the most intriguing aspects of nature. Importantly, besides their foundational interest, nonlocal correlations find applications in cryptographic and information-processing tasks such as, for example, device-independent quantum key distribution (Ekert, 1991; Pironio et al., 2009) or reduction of communication complexity (Dam, 2013; Brassard et al., 2006).

Bell nonlocality is today perceived as a phenomenon in its own right and can be defined and tested irrespectively of the underlying theory. In this thesis we will be exclusively interested in quantum correlations, and so we will assume that they are produced according to the laws of quantum mechanics. Under this assumption the concept of Bell nonlocality is perhaps best understood in terms of a quantum Bell experiment, which is sometimes also called a "no-signalling game". In the next section we will formulate Bell nonlocality using this type of tasks.

2.4.1 Nonlocality from the perspective of no-signalling games

Bell nonlocality can be understood from the perspective of no-signalling games, which also provides an intuitive understanding of Bell inequalities. Such games have been extensively studied in computer science for a long time, where they form a special instance of *interactive proof systems* (Cleve et al., 2004).

The standard scenario of a no-signalling game involves two cooperating players (Alice and Bob) who play the game against a third party, the referee. The referee chooses a question $x \in \mathcal{X}$ for Alice and $y \in \mathcal{Y}$ for Bob according to some probability distribution $p(x, y) : \mathcal{X} \times \mathcal{Y} \to [0, 1]$, where $\mathcal{X}$ and $\mathcal{Y}$ denote finite sets of questions. Without communicating, and therefore, without knowing what question the other player was asked, Alice (Bob) returns an answer $a \in \mathcal{A}$ ($b \in \mathcal{B}$) from a finite set of possible answers $\mathcal{A}$ ($\mathcal{B}$). Based on the questions asked and the received answers, the referee determines whether the players win or lose the game, according to a pre-arranged set of rules. Such rules are typically expressed using a function $V : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to [0, 1]$, where $V(a, b, x, y) = 1$ if and only if Alice and Bob win the game by answering a and b for questions x and y.

Alice and Bob know the rules of the game, that is, they know the function V and the distribution of questions $p(x, y)$. Before the game starts they can agree on any strategy which provides them with the best chances of winning. However, once the game starts, they are not allowed to communicate any more. In the classical setting any strategy they can possibly devise can be encoded in a classical memory system, represented by a shared random variable λ and a probability distribution $p(\lambda)$. In the more general quantum case, any possible strategy can be described by a shared quantum state ρ and a choice of local measurements.

In order to relate the above game setting with Bell inequalities note that the referee's questions x and y can be thought of as labels for different measurement settings. Similarly, the answers correspond to the outcomes of local measurements.

Any measurement strategy (be it classical or quantum) leads to a conditional probability $p(a,b|x,y)$ which describes when Alice and Bob give answers a and b for questions x and y, respectively. In the language of Bell inequalities $p(a,b|x,y)$ determine the probability that Alice and Bob obtain measurement outcomes a and b when performing the measurements labelled by x and y. The average probability that Alice and Bob win can be written as:

$$p^{V}_{\text{guess}}(\mathcal{G},\mathbb{M}_A,\mathbb{M}_B) = \sum_{a,b,x,y} p(x,y)p(a,b|x,y)V(a,b,x,y), \tag{2.48}$$

where $\mathcal{G} = \{p(x,y), V\}$ defines the game and the conditional probabilities $p(a,b|x,y)$ are related to the local measurements $\mathbb{M}_A = \{M^A_{a|x}\}$ for Alice and $\mathbb{M}_B = \{M^B_{b|y}\}$ for Bob, via the Born rule:

$$p(a,b|x,y) = \text{tr}\left[\left(M^A_{a|x} \otimes M^B_{b|y}\right)\rho^{AB}\right]. \tag{2.49}$$

With this in mind, Bell inequalities can be thought of as upper bounds on the average guessing probability $p_{\text{guess}}(\mathcal{G},\mathbb{M})$ with which Alice and Bob can win a nonlocal game $\mathcal{G}$ using a classical strategy (i.e. when ρ^{AB} is a separable state), optimized over all local measurements $\{M^A_{a|x}\}$ and $\{M^B_{b|y}\}$. A violation of a Bell inequality corresponds to the situation when there is a quantum strategy which uses an entangled shared state and outperforms the best classical strategy in a particular game $\mathcal{G}$.

Importantly, there are entangled states which can never violate any Bell inequality (Werner, 1989; Barrett, 2002; Augusiak, Demianowicz, and Acín, 2014). In the language of no-signalling games this means that there are states ρ^{AB} which, although entangled, can never outperform the best classical strategy. Interestingly, in (Buscemi, 2012) Buscemi showed that when we modify the rules of the no-signalling game and allow the referee to ask *quantum* instead of classical questions, then all entangled states can outperform the best classical strategy in some nonlocal game, or equivalently, violate the corresponding Bell inequality. This realisation leads to a new notion of nonlocality, one that can be refered to as *Buscemi nonlocality*. We will address this type of nonlocality in more detail in Chapter 6. In particular, we will show that, similar to the more standard Bell nonlocality, it also admits a direct operational characterization.

2.5 The formalism of quantum resource theories

2.5.1 The general structure of quantum resource theories

Ever since quantum information theory was conceived, it was clear that it is a theory centered around the problem of converting different resources. The formalism of quantum resource theories is a set of mathematical tools and methods that aims to mimic this reasoning to solve more general problems, i.e., those that can be associated with resources other than information or entanglement. It has been developed with the goal to systematically quantify different properties of quantum systems. Quantum resource theories (QRTs) can be nowadays classified in terms of *objects* and *resources* studied in a given theory. Classification of QRTs with respect to the object lead to the resource theories of states (Brandão and Gour, 2015), measurements (Skrzypczyk and Linden, 2019; Designolle, Farkas, and Kaniewski, 2019a; Oszmaniec et al., 2017), channels (Devetak, Harrow, and Winter, 2008; Liu

and Winter, 2019; Liu and Yuan, 2019; Wilde, 2013), and boxes (Skrzypczyk, Brunner, and Popescu, 2009; Schmid, Rosset, and Buscemi, 2020). On the other hand, classifying QRTs with respect to the type of the studied resource leads to the resource theories of pure (Bennett and Wiesner, 1992; Bennett et al., 1993; Nielsen, 1999) and mixed-state entanglement (Bennett et al., 1996b; Vidal and Tarrach, 1999), coherence (Baumgratz, Cramer, and Plenio, 2014; Winter and Yang, 2016; Napoli et al., 2016), purity (Horodecki, Horodecki, and Oppenheim, 2003; Streltsov et al., 2018), athermality (Janzing et al., 2000; Horodecki and Oppenheim, 2013), nonlocality (Cavalcanti and Skrzypczyk, 2016), asymmetry (Marvian and Spekkens, 2013), measurement incompatibility (Buscemi, Chitambar, and Zhou, 2020), teleportation (Cavalcanti, Skrzypczyk, and Supic, 2017), amongst many more. Its worth mentioning that although many QRTs use essentially the same mathematical formalism, their physical implications can be genuinely different. Hence the wide applicability of the framework to otherwise unrelated problems is a truly surprising aspect of nature.

Irrespective of the above classification, all QRTs share a common set of core postulates. We now present the axioms that abstractly define the notion of a Quantum Resource Theory (Chitambar and Gour, 2019). We will then discuss some of its implications and present a particular example of a resource theory that fits in this framework.

Consider a subset of quantum channels $\mathcal{F} \in \text{CPTP}$ and denote by $\mathcal{F}(A \to B)$ a set obtained by assigning to any pair of input/output quantum systems A and B, a corresponding set of quantum channels from $\mathcal{F}$ that act from A to B. Let us also denote with "1" a trivial quantum system. With this we can now present the following definition.

Definition 2.11 (Quantum Resource Theory)
A set of quantum channels $\mathcal{F}$ is called a Quantum Resource Theory (QRT) if it includes

1. For any system A the identity map $\text{id}_A \in \mathcal{F}(A \to A)$.
2. For any three physical systems A, B and C, if the channels $\mathcal{E} \in \mathcal{F}(A \to B)$ and $\mathcal{N} \in \mathcal{F}(B \to C)$ then their concatenation $\mathcal{N} \circ \mathcal{E} \in \mathcal{F}(A \to C)$.
3. For any system A, the set $\mathcal{F}(A \to 1)$ is not empty, i.e $\mathcal{F}(A \to 1) = \{\text{Tr}\}$.
4. For any three systems A, B and C and any channel $\mathcal{E} \in \mathcal{F}(A \to B)$, we have

$$\mathcal{E}_A \otimes \text{id}_C \in \mathcal{F}(AC \to BC). \tag{2.50}$$

Let us briefly discuss these conditions and some of their consequences. The first condition means that any QRT must necessarily have the identity map as a free operation. This is intuitive, since if we don't do anything with our physical system then we should not expect that it suddenly turns it into a resource. The second condition means that any QRT is closed under channel compositions. In other words, we cannot generate resources by successively applying only free operations. The third condition means that discarding physical systems can be always done for free, i.e. the trace is always a free operation. The final condition means that free operations are "completely free". This means that we cannot turn a free operation into a resourceful one by doing something trivial, i.e appending another system and doing nothing to it.

Definition 2.12 (Free states)
The set $\mathcal{F}(A) \equiv \mathcal{F}(1 \to A)$ is called the *set of free states*.

Maps which take the trivial input 1 to quantum systems can be thought of as preparations of quantum states. This identification is a convention that allows to formally infer an important principle valid for all QRT's satisfying conditions (1-3).

Fact 2.1 (The Golden Rule)**.** If $\rho \in \mathcal{F}(A)$ and $\mathcal{E} \in \mathcal{F}(A \to B)$ then $\mathcal{E}(\rho) \in \mathcal{F}(B)$.

This means that it is impossible to generate new resources using only free operations. We can view this as the most important property of all QRT's, as violating it will make the theory (asymptotically) trivial. More specifically, given a sufficient number of free states, one would be able to create arbitrary non-free states. As a result, every state becomes free and the theory becomes trivial.

Proof. The fact follows simply from the identification of the set of free states $\mathcal{F}(A)$ with channels $\mathcal{F}(1 \to A)$ and condition (2). □

Let us now consider two free channels $\mathcal{E} \in \mathcal{F}(A \to 1)$ and $\mathcal{N} \in \mathcal{F}(1 \to B)$ defined as $\mathcal{E}[\rho_A] := \mathrm{tr}\big[\rho_A\big]$ and $\mathcal{N}[1] := \sigma_B$. Due to the axiom (2) their concatenation $\mathcal{N} \circ \mathcal{E} \in \mathcal{F}(A \to B)$ is also free. Therefore we have the following

Fact 2.2. In any QRT the replacement channel:

$$\mathcal{N}_\sigma^{A\to B}(\rho_A) := \mathrm{tr}\big[\rho_A\big]\,\sigma_B \qquad \forall\, \rho_A \in \mathcal{D}(A) \tag{2.51}$$

with σ_B being a free state, is free, i.e. $\mathcal{N}_\sigma^{A\to B} \in \mathcal{F}(A \to B)$.

This simple fact has powerful consequences, i.e. every state can be converted to an arbitrary free state by some free operation. In other words, all free states are equivalent.

The last condition from Definition (2.11) allows to infer a few interesting and general properties of QRT's. In particular, using properties (2) and (4) we may conclude that a tensor product of two free maps is also free, i.e.

Fact 2.3. For $\mathcal{E}_1 \in \mathcal{F}(A \to B)$ and $\mathcal{E}_2 \in \mathcal{F}(A' \to B')$ we have:

$$(\mathcal{E}_1 \otimes \mathcal{E}_2) = (\mathcal{E}_1 \otimes \mathrm{id}) \circ (\mathrm{id} \otimes \mathcal{E}_2) \in \mathcal{F}(AA' \to BB'). \tag{2.52}$$

Furthermore, using this fact and identifying the set of free states with free maps that have a trivial input from Definition (2.12), we can infer that the product of two (or more) free states is also free, i.e.

Fact 2.4. For $\rho \in \mathcal{F}(A)$ and $\sigma \in \mathcal{F}(B)$ we have $\rho \otimes \sigma \in \mathcal{F}(AB)$

Similarily we find that appending a free state is a free operation, i.e.

Fact 2.5. For $\rho \in \mathcal{D}(A)$ and $\sigma \in \mathcal{F}(B)$ the channel $\mathcal{N}_\sigma^{A\to AB}(\rho_A) := \rho_A \otimes \sigma_B \in \mathcal{F}(A \to AB)$.

Many relevant resource theories found in the literature are *convex*, meaning that the set of free states $\mathcal{F}(A)$ is convex. This is an important property that assures certain important mathematical properties, as well as allows for quantifying the resource content in a very illustrative and natural way using the notion of a distance. More physically, convexity of resource theory means that when we toss a coin to obtain a probabilistic mixture of free states or free channels, this mixture will also be a free object of this resource theory. From now on we will only consider convex QRTs. Moreover, we will be interested in theories whose set of free states $\mathcal{F}(A)$ is closed.

This effectively means that when we take an infimum or supremum over that set, we can replace it with taking simply the minimal or maximal element of that set.

Having established the very basic formal structure of general QRTs, let us now move on to the problem of quantifying the resource content of a given quantum state. For simplicity in what follows we will focus exclusively on quantifying properties of states. However, the measures we are going to define can be naturally extended to other objects like, for example, measurements or quantum channels.

2.5.2 Quantifying resources

One of the benefits of formulating problems using the formalism of resource theories is that it readily provides precise and operationally well-motivated methods to quantify the resource content. This problem has been studied in multiple ways, ranging from more axiomatic (Regula, 2017), to more operational approaches (Liu, Hu, and Lloyd, 2017). Here we will only review the measures that are relevant for our purposes. Before we do that, however, let us first state some desirable properties of resource measures.

Definition 2.13 (Resource measure)
A function $f : \mathcal{F}(A) \to \mathbb{R}$ is a *resource measure* if it is non-increasing under free operations, that is

$$f(\rho) \geq f[\mathcal{F}(\rho)], \tag{2.53}$$

for all $\mathcal{F} \in \mathcal{F}(A \to A)$

In this thesis we will also consider resource measures that satisfy additional properties, that is

(i) **Faithfulness**, meaning that it vanishes if and only if the state is free, i.e

$$f(\rho) = 0 \iff \rho \in \mathcal{F}(A). \tag{2.54}$$

(ii) **Convexity**, meaning that by mixing two quantum states ρ_1 and ρ_2 one cannot obtain a more resourceful one, i.e for

$$\rho_{\text{mix}} = (1-p)\rho_1 + p\rho_2, \tag{2.55}$$

and for any $0 \leq p \leq 1$ we necessarily have

$$f(\rho_{\text{mix}}) \leq (1-p)f(\rho_1) + pf(\rho_2). \tag{2.56}$$

Naturally, demanding both of these properties at once may not lead to useful resource measures; One can easily imagine situations where faithfulness does not necessarily hold (e.g. when measuring entanglement using a fixed entanglement witnesses in the case of the resource theory of bipartite entanglement). In general, depending on the specific situation, we will be interested in resource measures that satisfy one of (or both) properties $(i - ii)$.

For convex QRTs we can quite easily define a natural resource measure, one that in the same time admits an appealing geometric interpretation. The main idea is to associate the resource content of a state with its distance from the set of free states. Naturally, this depends on the accepted notion of distance, i.e. a

function that is contractive under quantum channels. Let us recall that a function $d : \mathcal{D}(A) \otimes \mathcal{D}(A) \to \mathbb{R}_{\geq 0}$ is a *contraction* under quantum channels if and only if

$$d(\rho, \sigma) \geq d[\mathcal{E}(\rho), \mathcal{E}(\sigma)] \tag{2.57}$$

for all $\rho, \sigma \in \mathcal{D}(A)$. If d is such a function then a natural resource measure can be defined by taking

$$f_d(\rho) = \min_{\sigma \in \mathcal{F}(A)} d(\rho, \sigma). \tag{2.58}$$

It can be easily verified that any measure defined in this way satisfies the conditions from Def. 2.13. Notice also that the distance measure in Eq. (2.58) is only assumed to be monotonic under CPTP maps, hence it is not necessarily a metric and does not need to be symmetric in its arguments. Now, depending on the specific function d we can define very general and wide classes of resource measures. We will focus here on two such classes: the so-called entropic measures and robustness-based measures.

Entropic measures

Many tasks in quantum information theory can be quantified using quantum Renyi divergences. In fact these general functions can be also used to obtain a family of useful measures of resources that have quite interesting properties. In particular, we say that a function is an *entropic resource measure* if it is of the form

$$D_\alpha(\rho||\mathcal{F}(A)) = \min_{\sigma \in \mathcal{F}(A)} D_\alpha(\rho||\sigma), \tag{2.59}$$

where $\alpha \in \mathbb{R}$ is an arbitrary parameter that specifies a particular measure from this family. Let us now mention a few important examples of such measures that can be obtained by looking at the limiting cases of quantum Renyi divergences. For quantum states ρ and σ such that $\rho \subseteq \sigma$ we have

$$D_0(\rho||\sigma) \equiv \lim_{\alpha \to 0} D_\alpha(\rho||\sigma) = -\log \operatorname{tr}\left[\Pi_\rho \sigma\right], \tag{2.60}$$

$$D_1(\rho||\sigma) \equiv \lim_{\alpha \to 1} D_\alpha(\rho||\sigma) = D(\rho||\sigma), \tag{2.61}$$

$$D_\infty(\rho||\sigma) \equiv \lim_{\alpha \to \infty} D_\alpha(\rho||\sigma) = \min\left\{\lambda \mid \rho \leq 2^\lambda \sigma\right\}. \tag{2.62}$$

In the above we used Π_ρ to denote the projector onto the support of ρ. Sometimes in the literature the quantity in Eq. (2.60) is referred to as the *min-relative entropy*. Similarly we see that Eq. (2.61) is just our ordinary quantum relative entropy. The final quantity in Eq. (2.62) is also sometimes called *max-relative entropy*. These quantities can be then supplied into Eq. (2.59) to construct resource measures. Of particular importance for our purposes is the so-called max-relative entropy of a resource that corresponds to

$$D_\infty(\rho||\mathcal{F}(A)) = \min_{\sigma \in \mathcal{F}(A)} \left\{\lambda \mid \rho \leq 2^\lambda \sigma\right\}. \tag{2.63}$$

The relative entropies are not proper metrics in the mathematical sense, as they do not satisfy the triangle inequality. This fact, however, does not diminish their significance for resource theories in general. In the following chapters we will study a particular example of a resource theory, the so-called resource theory of quantum thermodynamics. We will later see that in that theory the quantities $D_\alpha(\rho||\mathcal{F}(A))$ form

an important (and in some cases also complete) family of monotones that generalise the standard second law of thermodynamics.

Robustness-based measures

An alternative way of quantifying the resource content of a quantum state can be obtained by specifying its robustness against noise. Let ρ be a quantum state whose resource content we want to quantify and let $\mathcal{F}(A)$ be the set of free states of our QRT. Consider the mixture

$$\rho_{\text{mix}} = (1-\lambda)\rho + \lambda\gamma, \tag{2.64}$$

where $0 \leq \lambda \leq 1$ and γ is an arbitrary quantum state that encodes the structure of the noise. Our goal is to choose the noise γ, as well as its minimal magnitude λ so that the mixture in Eq. (2.64) becomes a free state, i.e. $\rho_{\text{mix}} \in \mathcal{F}(A)$. When this can be achieved for $\lambda = 0$, then it follows that the state ρ must be a free state. Using this reasoning we can obtain different resource measures by allowing for different types of noise in Eq. (2.64). For the purpose of this thesis we will resort with the most general choice, that is we will allow γ to be any quantum state, i.e. $\gamma \in \mathcal{D}(A)$. Furthermore, for some mathematical reasons that will become clear later on, we will parametrise the noise with a parameter $r = \lambda/(1-\lambda)$. We are now ready to introduce a general robustness-based measure that will be used throughout this thesis.

Definition 2.14 (Generalised robustness of a quantum state)
For a quantum state $\rho \in \mathcal{D}(A)$ the *generalised robustness* with respect to the free set $\mathcal{F}(A)$ is given by

$$\begin{aligned} \mathbf{R}_{\mathcal{F}(A)}(\rho) \;=\; & \min \quad r && (2.65)\\ & \text{subject to} \quad \frac{1}{1+r}\rho + \frac{r}{1+r}\gamma = \sigma && (2.66)\\ & \qquad\qquad \sigma \in \mathcal{F}(A), \quad \gamma \in \mathcal{D}(A). && (2.67) \end{aligned}$$

Different robustness-based measures can be also obtained by choosing different free sets $\mathcal{F}(A)$. Moreover, one can even define an analogous quantity for more general objects like ensembles of quantum states, generalised measurements or even quantum channels. Beyond its study in entanglement theory (Takagi and Zhuang, 2018), the notion of robustness has recently been investigated in the QRTs of coherence and asymmetry (Napoli et al., 2016), steering (Piani and Watrous, 2015) and many more. Interestingly, generalised robustness is closely related with the max-relative entropy of a resource defined in Eq. (2.63). More specifically, we have the following relationship between resource measures

$$D_\infty(\rho||\mathcal{F}(A)) = \log\left[1 + \mathbf{R}_{\mathcal{F}(A)}(\rho)\right] \tag{2.68}$$

This connection had been noticed for the first time in (Datta, 2009).

2.6 Majorization in quantum resource theories

The set of free channels induces a natural preorder between quantum states. However, characterising this preorder in the general case is impossible without exploiting additional properties of a specific resource theory. In general, determining whether there exists a free operation $\mathcal{F} \in \mathcal{F}(A \to B)$ such that $\mathcal{F}(\rho_A) = \sigma_B$ is difficult

and in most cases strongly depends on the structure of a particular resource theory. However, some of the well-known resource theories exhibit a certain degree of universality when it comes to answering this question. More specifically, in such theories the question of whether one state can be converted into another state can be answered using a mathematical concept of majorization. This allows for a much simpler characterisation of such theories, and very often leads to a computationally efficient method of determining the free channel that performs the desired conversion. Quantum resource theories that can be characterised using majorization are called *majorization-based* and are the main topic of this section.

2.6.1 Majorization

Majorization is a mathematical concept from matrix analysis (Marshall, Olkin, and Arnold, 2011). It has found applications in various areas of science, ranging from mathematics (Hardy et al., 1952), economics (Kleiner, Moldovanu, and Strack, 2020), social sciences (Boland, 1989) up to quantum information (Nielsen, 1999) and even quantum thermodynamics (Horodecki and Oppenheim, 2013).

Before we define majorization, let us first discuss what we mean by "mixing" two probability distributions. Formally, we say that $\boldsymbol{q}$ is more mixed than $\boldsymbol{p}$ when there exists a doubly-stochastic matrix M that can take $\boldsymbol{p}$ into $\boldsymbol{q}$, i.e

$$\boldsymbol{q} = M\boldsymbol{p} \quad \text{for some} \quad M \quad \text{doubly-stochastic} \tag{2.69}$$

Doubly-stochastic matrices can be viewed as classical channels, i.e. they are linear maps that map probability vectors into probability vectors (since their columns sum to one) and whose fixed point is the uniform distribution.

In simple term majorization is a quantitative way of answering the following question: given two probability vectors $\boldsymbol{p}$ and $\boldsymbol{q}$, which one of them is more mixed? For certain simple cases this can be easily solved, e.g when $\boldsymbol{p} = [1, 0, 0]$ and $\boldsymbol{q} = [1/3, 1/3, 1/3]$ we can deduce that $\boldsymbol{q}$ is more mixed than $\boldsymbol{p}$ since $\boldsymbol{q}$ can be obtained by uniformly mixing all elements of $\boldsymbol{p}$. However, for generic probability distributions it may not be so easy to give a straightforward answer here. Moreover, for some pairs of vectors determining which one of them is more mixed may not be even possible. In this sense majorization is a binary relationship, denoted here with " $\succ$", that captures when one state is more mixed than another. More formally, we have the following definition.

Definition 2.15 (Majorization)
Let $\boldsymbol{p} \in \mathbb{R}^d$ and $\boldsymbol{q} \in \mathbb{R}^d$ be two probability vectors. Then $\boldsymbol{p}$ majorizes $\boldsymbol{q}$, denoted $\boldsymbol{p} \succ \boldsymbol{q}$, if and only if

$$\sum_{i=1}^{k} p(i) \leq \sum_{i=1}^{k} q(i) \quad \text{for all} \;\; k = 1, \ldots, d. \tag{2.70}$$

Technically, the relationship between majorization as defined in Def. 2.15 and the notion of mixedness (defined using doubly-stochastic matrices) is not trivial, and is the main content of the celebrated theorem by Hardy, Littlewood and Polya (Hardy, 1929). For our purposes it is enough to remember that these two notions are equivalent.

2.6.2 Relative majorization

The statement that $\boldsymbol{p} \succ \boldsymbol{q}$ can be regarded as saying that the elements of $\boldsymbol{p}$ are more equal than those of $\boldsymbol{q}$. In other words, $\boldsymbol{p}$ is closer than $\boldsymbol{q}$ to the uniform vector $\boldsymbol{u} = (1/d, \ldots, 1/d)$. A simple generalization of majorization can be obtained by substituting a more general probability vector $\boldsymbol{g} \in \mathbb{R}^d$ in place of $\boldsymbol{u}$. The resulting generalization is termed *relative majorization*.

Definition 2.16 (Relative majorization)
Let $\boldsymbol{p} \in \mathbb{R}^d$ and $\boldsymbol{q} \in \mathbb{R}^d$ be two probability vectors and let $\boldsymbol{g} \in \mathbb{R}^d$ be a reference vector. Then $\boldsymbol{p}$ majorizes $\boldsymbol{q}$ relative to $\boldsymbol{g}$, denoted $\boldsymbol{p} \succ_{\boldsymbol{g}} \boldsymbol{q}$, if and only if there exists a matrix R such that

(i) $R\boldsymbol{p} = \boldsymbol{q}$,

(ii) $R\boldsymbol{g} = \boldsymbol{g}$,

(iii) $\mathbf{1}^T R = \mathbf{1}^T$.

To be more precise, what in the literature is widely known as "relative majorization" corresponds to the case when the relative state may also change, i.e. the condition (ii) states instead $R\boldsymbol{g} = \boldsymbol{g}'$. Since in this thesis we will exclusively use the simpler case when $\boldsymbol{g}' = \boldsymbol{g}$, we will refer to this less general notion of majorization also using the term "relative majorization".

As we can see, the generalisation is obtained by considering linear maps that preserve a fixed (rather than uniform) vector. Intuitively, any such classical channel can only map states closer to $\boldsymbol{g}$, therefore making them less distinguishable from $\boldsymbol{g}$.

Let us now describe a technical tool that was introduced in (Brandão and Gour, 2015). In practice, it makes the computation much easier as it allows to formulate problems initially related to relative majorization using the ordinary notion of majorization.

Definition 2.17 (Embedding map)
The embedding map $\Gamma_{\boldsymbol{d}} : \mathbb{R}^n \to \mathbb{R}^D$ is a transformation between vectors such that

$$\Gamma_{\boldsymbol{d}}(\boldsymbol{x}) := \left(\underbrace{\frac{x(1)}{d_1}, \ldots, \frac{x(1)}{d_1}}_{d_1 \text{ terms}}, \underbrace{\frac{x(2)}{d_2}, \ldots, \frac{x(2)}{d_2}}_{d_2 \text{ terms}}, \ldots, \underbrace{\frac{x(n)}{d_n}, \ldots, \frac{x(n)}{d_n}}_{d_n \text{ terms}} \right) \tag{2.71}$$

$$= \bigoplus_{k=1}^{n} x(k)\boldsymbol{u}_k, \tag{2.72}$$

where $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ is a vector of natural numbers which sum to $D = \sum_{i=1}^n d_i$ and $\boldsymbol{u}_k := \frac{1}{k}(1, 1, \ldots, 1)$ is a $k-$dimensional uniform vector.

Intuitively, the embedding map is is a transformation that allows for translating between different (equivalent) descriptions of a physical system. For example, if $\boldsymbol{p}$ and $\boldsymbol{q}$ are statistical descriptions of a system in the canonical ensemble, then $\Gamma_{\boldsymbol{d}}(\boldsymbol{p})$ and $\Gamma_{\boldsymbol{d}}(\boldsymbol{q})$ for a suitably chosen $\boldsymbol{d}$ describe these two states in the microcanonical ensemble (see e.g. Egloff et al., 2015).

Majorization admits several equivalent characterizations. So far we have introduced two of them: the standard formulation using partial sums, and a stochastic formulation using doubly-stochastic matrices. We have also seen that this second

formulation allows to extend it to relative majorization. In what follows we will explore a third formulation using a graphical characterization in terms of the so-called Lorenz curves.

2.6.3 Lorenz curve formulation

Majorization relation, as well as its generalization called relative majorization, can be described graphically using the concept of a Lorenz curve.

Definition 2.18 (Lorenz curve)
Let $\mathcal{L}(\boldsymbol{p}||\boldsymbol{g})$ be a piece-wise linear curve on the plane constructed by joining points of the form $[x(i), y(i)]$ where

$$x(0) = 0, \qquad x(i) = g(\pi(i)) \quad \text{for} \quad i = 1, \ldots, d, \tag{2.73}$$
$$y(0) = 0, \qquad y(i) = p(\pi(i)) \quad \text{for} \quad i = 1, \ldots, d, \tag{2.74}$$

and $\pi(i)$ is a permutation that orders $p(i)/g(i)$ non-increasingly.

Lorenz curves give an equivalent characterization of the (relative) majorization relation. This is captured by the following lemma from (Horodecki and Oppenheim, 2013).

Lemma 2.1 (Lorenz curve characterisation of majorization)**.** A probability vector $\boldsymbol{p}$ majorizes $\boldsymbol{q}$ relative to $\boldsymbol{g}$, that is

$$\boldsymbol{p} \succ_{\boldsymbol{g}} \boldsymbol{q} \tag{2.75}$$

if and only if $\mathcal{L}(\boldsymbol{p}||\boldsymbol{g})$ lies everywhere above $\mathcal{L}(\boldsymbol{q}||\boldsymbol{g})$.

Having established some basic facts about majorization let us discuss its relevance for quantum resource theories.

2.6.4 Majorization-based quantum resource theories

In Sec. 2.5.1 we have seen that free operations in a natural way induce a preorder on the set of all quantum states. More specifically, we say that ρ can be transformed into σ using free operations if we can find a free operation $\mathcal{F} \in \mathfrak{F}(A)$ such that $\mathcal{F}(\rho) = \sigma$. In general checking if there exist a free operation that can take one state into another is a difficult problem. Moreover, it may be the case that the natural preorder induced by free operations does not have a simple characterisation. Hopefully, some of the most important resource theories share a certain property that allows this preorder to be simply characterised using the concept of majorization. This class of resource theories is known as *majorization-based* QRTs.

More specifically, within majorization-based QRTs quantum states can be represented using appropriately chosen probability distributions. These distributions encode state's affiliation to the specific resource, and decompose the set of all quantum states into equivalence classes depending on the resource contained within a given quantum state. Such an encoding is specific to a given resource theory. For example, in the case of quantum thermodynamics we associate:

$$\boldsymbol{p} = [p(1), p(2), \ldots, p(d)], \qquad \boldsymbol{q} = [q(1), q(2), \ldots, q(d)], \tag{2.76}$$

where $p(i) = \langle E_i|\rho|E_i\rangle$ and $q_i = \langle E_i|\sigma|E_i\rangle$ are the state's occupations in the energy eigenbasis $\{|E_i\rangle\}$. Each majorization-based QRT is equipped with a function that maps quantum states into a corresponding resource encoding.

Given such an encoding, in all majorization-based QRTs a free operation transforming the input into the output exists if and only if the resource encoding of the input relatively majorizes the resource encoding of the output.

Some examples of majorization-based QRTs are presented in Table 2.1.

QRT	Free ops.	Free states	Reference
Entanglement (pure)	LOCC	**SEP**$(A:B)$	(Nielsen, 1999; Bennett et al., 1996a)
Coherence (pure)	IO	$\sum_i \alpha_i \lvert i\rangle$ s.t. $\alpha_i \alpha_j^* = \delta_{ij}$	(Du, Bai, and Guo, 2015)
Purity	NO	$\frac{1}{d}\mathbb{1}$	(Gour et al., 2015)
Thermodynamics	TO	$\frac{1}{Z}e^{-\beta H}$	(Horodecki and Oppenheim, 2013)

TABLE 2.1: Quantum resource theories whose transformation laws are governed by majorization. The corresponding free operations are: local operations and classical communication (LOCC), incoherent operations (IO) and noisy operations (NO) and thermal operations (TO).

In the next section we will carefully examine a specific majorization-based QRT to show how all these concepts can be used to study a particular problem.

2.6.5 Resource theory of quantum thermodynamics

In this section we will address a specific majorization-based QRT called the resource theory of quantum thermodynamics, also sometimes referred to as the resource theory of thermal operations. A more complete description of this framework can be found in several excellent reviews, i.e (Goold et al., 2016; Vinjanampathy and Anders, 2016; Lostaglio, 2019). Our presentation is also influenced by the very illustrative review (Landi and Paternostro, 2020), which provides a complimentary description of how the information-theoretic approach fits within the surroundings of more standard approaches to quantum thermodynamics.

Basic assumptions

The setting of thermal operations consists of a system S and a heat bath B with respective Hamiltonians H_S and H_B. We prepare these systems in states ρ_S and ρ_B and model their interaction by means of a global unitary U. The final state of the system after the interaction becomes

$$\mathcal{T}(\rho_S) = \mathrm{tr}_B[U(\rho_S \otimes \rho_B)U^\dagger]. \tag{2.77}$$

Before describing the basic assumptions of the theory, let us emphasize that Eq. (2.77) describes a very general family of maps. All interactions between the two sybsystems are contained in the unitary U, which may describe either a weak or strong coupling interaction, cyclic and time-dependent processes, as well as protocols involving external driving forces. Moreover, at this point we make no assumptions about the initial state of the bath (other than that it starts in a product state with respect to the system). Therefore it may very well describe a macroscopic reference frame, a genuine thermodynamic heat bath, or even an ancilla with dimension comparable to S. In fact, if we also accept that the choice of the Hamiltonian and initial state of the B subsystem is arbitrary, then (2.77) describes the most general type of evolution allowed by quantum theory. Since we are interested in thermodynamic transformations, we need to make some additional assumptions.

Assumption 1 (Zeroth Law). Our first assumption is the quantum mechanical analog of the zeroth law of thermodynamics. In particular, we assume that the heat bath starts in a state that is in thermal equilibrium with respect to an inverse temperature $\beta = 1/kT$. There are multiple ways of defining what thermal equilibrium means in this context; For our purposes it is enough to note that all of these imply that the initial state of the heat bath is described by the Gibbs state, that is

$$\rho_B = \tau_B \equiv \frac{1}{Z_B} e^{-\beta H_B}, \tag{2.78}$$

where $Z_B := e^{-\beta H_B}$ is the partition function of the heat bath. This also allows us to associate the change in the baths internal energy with heat flowing from the heat bath to the system, i.e.

$$Q_B = -\beta \operatorname{tr}\left[H_B(\rho'_B - \tau_B)\right], \tag{2.79}$$

where $\rho'_B = \operatorname{tr}_S[U(\rho_S \otimes \rho_B)U^\dagger]$ is the final reduced state of the heat bath.

Assumption 2 (First Law). We further assume that the energy of the system and the bath is conserved at all times during the interaction. Formally this means that the unitary U commutes with the total Hamiltoninan of the composite system, i.e.

$$[U, H_S + H_B] = 0, \tag{2.80}$$

where $H_S + H_B \equiv H_S \otimes \mathbb{1}_B + \mathbb{1}_S \otimes H_B$. We will refer to any map of the form (2.77) satisfying the above two assumptions as a *thermal operation*. Therefore the set of all thermal operations acting on system S, denoted $\mathcal{F}(S \to S)$, is the set of free channels of this particular resource theory. In the same time, the free states associated with system S are given by:

$$\mathcal{F}(S) \equiv \mathcal{F}(1 \to S) = \tau_S, \tag{2.81}$$

that is, all Gibbs states of the corresponding Hamiltonian H_S.

Thermomajorization

In general, a complete characterization of transformations that are possible under thermal operations is far from trivial and, despite some success in the low-dimensional case (Cwiklinski et al., 2015), the general problem is still open. Still, for states block-diagonal in the energy eigenbasis, there is a simple criterion determining when there exists a thermal operation such that $\mathcal{T}[\rho_S] = \sigma_S$. This criterion can be formulated using relative majorization.

Let us start by constructing a resource representation of the input and output states. To do so, we first write the Hamiltonian of the system as $H_S = \sum_{i=1}^{d_S} E_i \, |i\rangle\langle i|_S$ and define

$$\boldsymbol{p} = [p(1), p(2), \ldots, p(d_S)], \qquad \boldsymbol{q} = [q(1), q(2), \ldots, q(d_S)], \tag{2.82}$$

where $p(i) = \langle i|\rho_S|i\rangle$ and $q(i) = \langle i|\sigma_S|i\rangle$ are occupations in the energy eigenbasis. Similarly we denote the thermal state with $\tau_S = \operatorname{diag}[\boldsymbol{g}] = [g(1), g(2), \ldots, g(d_S)]$ where $g(i) = e^{-\beta E_i}/Z_S$ and $Z_S = \sum_i e^{-\beta E_i}$. With this we can now recall the following theorem proven in (Horodecki and Oppenheim, 2013).

Theorem 2.3 (Thermal operations for block-diagonal states)**.** A thermal operation $\mathcal{T}$ such that $\mathcal{T}[\rho] = \sigma$ exists if and only if

$$\boldsymbol{p} \succ_{\boldsymbol{g}} \boldsymbol{q}. \tag{2.83}$$

Notice that also recovers, as a special case, the relation between the resource theory of purity and standard majorization by either considering the limit of infinite temperature ($\beta \to 0$) or fully degenerate system's Hamiltonian ($H_S \propto \mathbb{1}_S$). In these limits thermal operations turn into noisy operations, that is the most general maps that can be implemented by unitary dynamics and access to a source of randomness.

In practice, to check this condition we can use two different approaches. First, we can use the Lorenz curve construction described in Sec. 2.6.3 to construct respective Lorenz curves and compare their elbows, which amounts to $d_S - 1$ comparisons in total. Alternatively, we can write an linear program and compute the stochastic matrix associated with the transformation. Both of these routines can be viewed as a special case of relative majorization with $\boldsymbol{g}$ corresponding to the Gibbs state.

Although the Lorenz curve in practice gives the easiest way of checking the existance of the desired thermal operation, it does not tell us anything about the very transformation. Therefore in some cases it may be more useful to use a more computationally expensive method based on finding a stochastic matrix. Let us briefly discuss this approach.

Stochastic description

For states block-diagonal in the energy eigenbasis the action of any thermal operation can be fully encoded in a stochastic matrix R with elements $R_{ij} := r(j|i)$ defined via:

$$r(j|i) := \left\langle j \middle| \mathcal{T}[|i\rangle\langle i|] \middle| j \right\rangle . \tag{2.84}$$

Due to the Assumptions 1 and 2, matrix elements $r(j|i)$ must satisfy certain conditions in order to describe a valid thermal operation that takes the state $\rho_S = \mathrm{diag}[\boldsymbol{p}] = \sum_i p(i)\ |i\rangle\langle i|$ to another state $\sigma_S = \mathrm{diag}[\boldsymbol{q}] = \sum_j q(j)\ |j\rangle\langle j|$. These constraints take the following form.

Definition 2.19 (Stochastic characterization of thermomajorization)
A resource encoding $\boldsymbol{p} \succ_{\boldsymbol{g}} \boldsymbol{q}$ if and only if there is a matrix R with elements given by $R_{ij} = r(j|i)$ such that:

$$\sum_i p(i)\, r(j|i) = q(j) \qquad \text{for all } j, \tag{2.85}$$

$$\sum_i r(j|i) e^{-\beta(E_i - E_j)} = 1 \qquad \text{for all } j, \tag{2.86}$$

$$\sum_j r(j|i) = 1 \qquad \text{for all } i, \tag{2.87}$$

$$r(j|i) \geq 0 \qquad \text{for all } i, j, \tag{2.88}$$

Let us briefly discuss the meaning of these constraints. Condition (2.85) implies that the map $\mathcal{T}$ is able to transform ρ_S into σ_S, and condition (2.86) ensures that the fixed point of the map is the Gibbs state τ_S. Conditions (2.87) and (2.88) ensure that R is a stochastic matrix and therefore $\mathcal{T}$ is a completely positive and trace-preserving channel. All of these conditions are linear in $r(j|i)$ and therefore finding a matrix R that transforms $\boldsymbol{p}$ into $\boldsymbol{q}$ can be cast in the form of a linear program.

The themormajorization criteria can be viewed as a refined version of the second law of thermodynamics, valid for arbitrary (hence, neither macroscopic nor average) thermodynamic processes. Still, we would expect that they correctly reproduce the standard second law of thermodynamics in an appropriate (e.g. the i.i.d) limit. However, this is not the case for the resource theory of quantum thermodynamics as defined so far. In Chapters 3 and 4 we will see, however, that natural extensions of the framework allow to easily reproduce this desired macroscopic behavior.

Thermodynamic work

An important thermodynamic concept which we have not discussed yet is the notion of thermodynamic work. From a purely thermodynamic perspective, work is a controlled flow of energy that can be defined in various equivalent ways in classical thermodynamics. Strikingly, this is not the case in quantum thermodynamics where the very definition of work is difficult to formulate. One of the reasons for this is that, in order to define work in a meaningful way, one has to use a measurement to determine the system's initial and final energy. [1] This comes at a cost, as the measurement inevitably disturbs the state of the system, so that some part of the energy change has to be associated with the measurement back-action, rather than work performed on the system. Moreover, a quantum measurement in most cases destroys coherences between energy levels, therefore also affecting potential quantum effects.

Standard approach. The standard and widely accepted approach to define work in the quantum regime involves the so-called *two-projective measurement scheme* (Campisi, Hänggi, and Talkner, 2011; Talkner, Lutz, and Hänggi, 2007). For simplicity we will assume that the initial and final Hamiltonian of the system are the same. The protocol of measuring work can be then described as follows.

Definition 2.20 (Two-projective measurement scheme)
The standard approach of measuring work in quantum systems relies on two energy measurements, before and after the evolution U, that is

1. Projective energy measurement on ρ_S, yielding outcome E_i with probability $p(i) = \langle i|\rho_S|i\rangle$, that is
$$\rho \to \sum_i p(i)\,|i\rangle\langle i|\,. \tag{2.89}$$
2. Unitary evolution U of the post-measurement state:
$$|i\rangle \to U\,|i\rangle \tag{2.90}$$
3. Projective energy measurement on the evolved state, resulting in E_j with probability $p(j|i) := |\langle j|U|i\rangle|^2$ and such that
$$\sum_i p(i) U\,|i\rangle\langle i|\,U^\dagger \to \sum_{i,j} p(i)p(j|i)\,|j\rangle\langle j|\,. \tag{2.91}$$

[1] There are ways to define thermodynamic work that do not make reference to explicit measurements, or minimize their backaction on the analysed thermodynamic system. Most of them, however, are still subject to a debate and therefore we will not be addressing them here.

The above procedure allows to collect statistics corresponding to a random variable $w := E_i - E_j$ that is distributed according to

$$p(w) = \sum_{i,j} p(i)p(j|i)\delta[w - (E_j - E_i)]. \tag{2.92}$$

It can be found that the average work obtained in this protocol is given by:

$$\langle w \rangle := \sum_w p(w)w = \mathrm{tr}\big[H\mathcal{D}(\rho)\big] - \mathrm{tr}\big[H'U\mathcal{D}(\rho)U^\dagger\big], \tag{2.93}$$

where $\mathcal{D}[X] := \sum_i \langle i|X|i\rangle\, |i\rangle\langle i|$ is a channel that is completely dephasing in the energy eigenbasis of H. For general quantum states ρ_S, the average value in Eq. (2.93) can be different than the corresponding average energy change on the system, that is

$$\langle w \rangle \leq \Delta E_S, \tag{2.94}$$

where $\Delta E_S := \mathrm{tr}\big[H\rho\big] - \mathrm{tr}\big[H'U\rho U^\dagger\big]$.

Explicit battery models. A different way of defining work can be obtained by explicitly modelling the behavior of an ancillary work register, responsible for storing and supplying work. This clearly follows the resource theoretic spirit of explicitly accounting for all systems that are involved in the process. In what follows we will describe the two most common battery models that appear in the literature.

Qubit battery. The first model of a battery in the framework of thermal operations was introduced by Horodecki and Oppenheim in (Horodecki and Oppenheim, 2013), who considered the battery to be a two-level system (called by them a *wit*) with Hamiltonian $H_W = \delta\,|1\rangle\langle 1|_W$. This allowed them to define a notion of deterministic work as the energy difference between the ground state and the excited state of the wit [see also (Åberg, 2013)]. In this way thermal operations assisted with a wit battery take the form:

$$\mathcal{T}_{\mathrm{wit}}\left[\rho_S \otimes |i\rangle\langle i|_W\right] = \sigma_S \otimes \big|j\big\rangle\big\langle j\big|_W, \tag{2.95}$$

where $(i, j) = (0, 1)$ when the transformation stores work in the battery (distillation) or $(i, j) = (1, 0)$ when the transformation consumes work (formation). The deterministic work of transition is defined to be the maximal (distillation) or minimal (formation) value of energy separation δ for which the input state thermomajorizes the output state, i.e.:

$$\rho_S \otimes |i\rangle\langle i|_W \succ_\beta \sigma_S \otimes \big|j\big\rangle\big\langle j\big|_W. \tag{2.96}$$

For distillation this optimization yields the maximal amount of work that we are guaranteed to extract from the state ρ_S by converting it into another state σ_S. For formation it gives the least amount of work that has to be supplied to guarantee transition from ρ_S to σ_S. This can be further generalized to cases where one allows the transformation to fail with some error probability ϵ. This is equivalent to transforming the input into a state σ_S^ϵ which is at most ϵ-close to the desired output state σ_S according to the trace distance. When we take the input state to be a thermal state $\rho_S = \tau_S$, then the corresponding work is called *work of formation* and is defined as

$$w_F^\varepsilon(\rho) = \min_{\rho' \in \mathcal{B}^\varepsilon(\rho)} \big[F_{\max}(\rho') - F(\tau)\big], \tag{2.97}$$

where $F(\rho) = \mathrm{tr}\left[H_S \rho_S\right] - kTS(\rho_S)$ is the non-equilibrium free energy, $F_{\max}(\rho) = kTD_\infty(\rho||\tau)$ and the minimum is taken over all states ρ' that are ε-close to ρ in trace distance. For the case when the output state is a thermal state, one defines the *work of distillation* as:

$$w_D^\varepsilon(\rho) = \max_{\rho' \in \mathcal{B}(\rho)} \left[F(\tau) - F_{\min}(\rho')\right] \tag{2.98}$$

where $F_{\min}(\rho) = kTD_0(\rho||\tau)$.

Ideal weight battery. Some recent works (Skrzypczyk, Short, and Popescu, 2014; Richens and Masanes, 2016; Alhambra et al., 2016) have proposed a different model of a thermodynamic battery, an *ideal weight* with Hamiltonian $H_W = \int x \, |x\rangle\langle x|_W \, \mathrm{d}x$, where the basis is formed from orthonormal states $\{|x\rangle_W, x \in \mathbb{R}\}$ representing position of the weight. Thermal operations acting on the system S and the weight are then given by a map defined as in Eq. (2.77), but now with an additional assumption that the global unitary U commutes with translations on the weight. This ensures that the weight cannot be used as an entropy dump, i.e. the joint entropy of the system S and heat bath B can never decrease by applying this type of transformation. More formally, translational invariance (TI) implies that any thermal operation reduced to the system and bath, that is $\mathcal{T}_{SB}[\cdot] := \mathrm{tr}_W[U((\cdot)_{SW} \otimes \rho_W)U^\dagger]$, for any initial state of the weight $\rho_W = \int_{-\infty}^{\infty} p_W(x) \, |x\rangle\langle x|_W \, \mathrm{d}x$, can be written as a mixture of unitaries (Masanes and Oppenheim, 2017), that is

$$\mathcal{T}_{SB}\left[\cdot\right] = \sum_x p(x) U_x(\cdot)U_x^\dagger, \tag{2.99}$$

for some ensemble $\{p(x), U_x\}$ where $\{U_x\}$ depend only on the global unitary U and not on the state ρ_W. Such a mixture cannot decrease the entropy of the system and the bath (but can increase it). In this way the energy difference of the battery may be associated solely with the work exerted by (or extracted from) system S. This can be viewed as a way of "proving" one of the statements of the second law of thermodynamics using purely the resource theoretic paradigm.

Chapter 3

Catalytic universality

3.1 Introduction

3.1.1 Background

It is a folklore knowledge that quantum resources must be consumed in order to perform useful tasks. A paradigmatic example is that of entanglement used in the ideal teleportation experiment, where a single maximally-entangled pair fuels the process of transferring an unknown quantum state between remote recipients. What is less known, however, is that in certain cases the very presence of quantum resources can be useful, without the resources being consumed or degraded. This surprising and yet not clearly understood phenomenon is called *quantum catalysis* and was introduced in (Jonathan and Plenio, 1999) and further analysed in a series of works (Turgut, 2007; Daftuar and Klimesh, 2001; Dam and Hayden, 2003; Duarte, Drumond, and Cunha, 2015; Aubrun and Nechita, 2007; Aubrun and Nechita, 2009). In simple terms, quantum catalysis demonstrates that access to a special resourceful state (the catalyst) can sometimes allow two distant parties to use their primary resource in a way that would otherwise be impossible. Importantly, the catalytic ancillary state is not consumed during the process, so that the agent can repeat their task again, or use the catalyst for a different task. This makes catalysis a particularly interesting extension of standard resource theories.

One of the most promising directions for catalysis is in the resource theory of quantum thermodynamics where catalysts are natural models of thermal machines, clocks or ancillary apparatuses that facilitate thermodynamic transformations.

3.1.2 Motivation

The laws of physics are often expressed as limitations on what physical systems can and cannot do. The second law of thermodynamics is a cardinal example of this approach: it says which thermodynamic transformations can be performed under given conditions. Specifically, at a constant background temperature and volume the transition between two equilibrium states can occur if and only if the Helmholtz free energy decreases during the process. The second law describes a relationship between average quantities (energy and entropy) and hence specifies the *typical* thermodynamic behavior, i.e. justified in the limit of a large number of identically distributed and weakly interacting systems.

One of the most striking differences between standard thermodynamics and its microscopic counterpart is that transformations between states can become significantly more demanding. More specifically, there are paradigms where they are no longer described by a single second law, but an entire family of conditions, the so-called "second laws of quantum thermodynamics" (Brandão et al., 2015). In this

way the free energy loses its meaning as the unique indicator of which state transitions are possible — its role is replaced by a family of generalized free energies, a collection of information-theoretic quantities closely related to the Renyi entropies (Rényi, 1961). This captures the idea that for microscopic systems more structure of the energy distribution must be specified in order to determine their thermodynamic properties. Importantly, by invoking typicality arguments it can be shown that in the limit of identically distributed and weakly interacting systems all members of this family of quantities approach the Helmholtz free energy, thus recovering the standard second law as a special case.

However, these results rely on a specific assumption: that there exists some thermal machine or 'catalyst' which is not consumed by the protocol but nonetheless makes the transformation possible. More specifically, if the second laws are satisfied for a pair of states ρ and σ then there is a quantum state ω which is unchanged by the protocol but still enables the joint transformation $\rho \otimes \omega \to \sigma \otimes \omega$. This becomes more natural once we realize that standard treatments implicitly adopt an analogous assumption; to perform a thermodynamic transformation one always needs to supply additional devices which can be cyclically reused (e.g. engines, refrigerators or heat pumps). In this way the ancillary state ω models the behavior of a thermal machine or an experimental apparatus which facilitates or even enables the transformation. This phenomenon of "lifting restrictions without being consumed" is called *quantum catalysis*.

Arguably, one of the most important problems within this approach to thermodynamics is how to find a catalyst which can be useful for a given transformation. Many of the existing results are based on constructing a very specific catalyst. This, however, may be obscuring the true physical mechanism behind catalysis. Furthermore, it is still not well understood which properties of quantum states are relevant for catalysis. The second laws only guarantee the existence of the catalyst; even if they are satisfied by a pair of states, it may still be difficult to find which state catalyzes a particular process. This intuition comes from our macroscopic experience: chemical reactions can be catalyzed only by appropriately chosen chemical compounds; similarly thermal machines need to be carefully tuned so that the desired transformation may happen. In this way a natural question appears: how can we find a state which catalyzes a given transformation and how special are these states?

3.1.3 Contribution

In this chapter we push forward our understanding of catalysis by numerically reporting a surprising property of multi-copy catalysts which we term *catalytic universality*. More specifically, we present a numerical analysis that demonstrates that generic large-dimensional states can act as universal catalysts. Moreover, we conjecture that any state, as long as enough copies of it are available, can serve as a catalyst for all allowable transformations.

3.1.4 Structure

This chapter is structured as follows. In Sec. 3.2 we introduce the relevant framework that we use to study catalysis in quantum thermodynamics. In Sec. 3.3 we describe catalytic universality and present numerical results indicating that generic large-dimensional states act as universal catalysts. Finally, we conjecture that all multi-copy states can be universal catalysts, given sufficiently many copies. To conclude,

in Sec. 3.4 we briefly summarise our results, discuss their implications and highlight some open problems that follow naturally from these findings.

3.2 Catalysis in quantum thermodynamics

3.2.1 Second laws of thermodynamics

One of the most striking differences between standard thermodynamics and its microscopic counterpart is that transformations between states can become significantly more demanding. More specifically, in the resource theory of quantum thermodynamics transformations are no longer described by a single second law, but an entire family of conditions, the so-called "second laws of quantum thermodynamics" (Brandão et al., 2015). In this way the free energy loses its meaning as the unique indicator of which state transitions are possible — its role is replaced by a family of generalized free energies, a collection of information-theoretic quantities closely related to quantum Renyi divergences (Rényi, 1961):

$$F_\alpha(\rho) := \frac{1}{\beta}\left[D_\alpha(\rho||\tau) - \log Z\right], \tag{3.1}$$

where $\tau = e^{-\beta H}/Z$ is a Gibbs state, $Z = \operatorname{tr} e^{-\beta H}$ is the partition function and $D_\alpha(\rho||\tau)$ are the quantum (sandwiched) Renyi divergences (Wilde, Winter, and Yang, 2014). These entropic quantities provide necessary conditions for the existence of a transformation between states. In particular, an important result of (Brandão et al., 2015) states that there exists a catalyst ω that enables the transformation $\rho \to \sigma$ only if:

$$F_\alpha(\rho) > F_\alpha(\sigma) \qquad \forall\, \alpha \geq 0. \tag{3.2}$$

Clearly, these relations are much stricter than the standard second law that can be seen as Eq. (3.2) for the case $\alpha = 1$. One way to understand this intuitively is that transforming microscopic systems requires a complete knowledge about the structure of the system's energy distribution. In other words, all moments of the system's energy distribution must be specified to determine its thermodynamic properties.

To be more precise, the statement of the second laws which we give here requires two additional technical assumptions to be formally correct. First, it assumes an arbitrarily small but nonzero error in the transformation. Second, it requires borrowing a qubit in a pure state that is given back with an arbitrarily small, but again, nonzero error [1].

Unfortunately, for general quantum states the conditions (3.2) are necessary, but not sufficient. This means that they do not guarantee the existence of a suitable catalyst ω and a thermal operation mapping $\rho \otimes \omega \to \sigma \otimes \omega$. However, these conditions do become sufficient if the states ρ and σ are block-diagonal in the basis determined by H_S (energy eigenbasis). This implies that they commute with the operator τ and hence the sandwiched Renyi divergence $D_\alpha(\rho||\tau)$ for $\alpha \geq 0$ simplifies to:

$$D_\alpha(\boldsymbol{p}||\boldsymbol{g}) = \frac{1}{\alpha - 1}\log\left[\sum_i p_i^\alpha g_i^{1-\alpha}\right]. \tag{3.3}$$

[1] Without these two technical assumptions the precise form of the second laws should read: $\forall\, \alpha \in \mathbb{R} \quad F_\alpha(\rho_S) > F_\alpha(\sigma_S)$

This also allows the second laws to be written in a much simpler form, and to see more clearly the connection between F_α and the non-equilibrium Helmholtz free energy, which is given by $F_1 = -k_B T \log Z_S$. It is important to note that the second laws (3.2) are *strictly* looser than the thermo-majorization criteria. This means that there are transformations which cannot be realized via thermal operations, i.e. without a catalyst, but can be performed when given access to a one. This is precisely due to this realization that catalysis is an important and highly non-trivial phenomenon in the resource theory of thermodynamics.

3.2.2 Thermodynamic limit

The standard second law of thermodynamics not only can be seen as a special case of (3.2), but also as their macroscopic limit. To see this consider transforming *collectively* a large number n of copies of ρ into n copies of σ, where both states are block-diagonal in the energy eigenbasis [therefore the relative entropy term can be simply written as in (3.3)]. In particular, it was shown in (Brandão et al., 2015) that for any $\epsilon > 0$ we have

$$\lim_{n\to\infty} \max_{\widetilde{\rho}\in\mathcal{B}_\epsilon(\rho)} \frac{F_\alpha(\widetilde{\rho}^{\otimes n})}{n} = F_1(\rho). \tag{3.4}$$

In other words, in the limit of identically distributed and weakly interacting systems all members of the family of generalised free energies approach nonequilibrium free energy $F_1(\rho) \equiv F(\rho)$, thus recovering the standard second law as a special case.

3.2.3 Disturbing the catalyst

So far we have discussed the case of *exact catalysis*, corresponding to an idealised scenario where the catalyst is returned perfectly undisturbed. In reality, however, any physical map will inevitably modify the catalyst. It is then important that our definition of catalysis is robust against such implementation imperfections.

Catalysis can be very naturally generalized to more physical scenarios which allow for small perturbation in the final state of the catalyst. This relaxation leads to *inexact catalysis*, where the error on the catalyst ϵ_C is defined in the most general form as:

$$\epsilon_C := \left\| \mathrm{Tr}_S \, \mathcal{T}_{SC}[\rho_S \otimes \omega_C] - \omega_C \right\|_1, \tag{3.5}$$

where $\mathcal{T}_{SC}$ is a CPTP map acting on the system and the catalyst. The case of exact catalysis can be recovered by (i) setting the error on the catalyst to zero, i.e. $\epsilon_C = 0$ and (ii) allowing no correlations between the system and the catalyst, i.e. demanding that the two subsystems end up in a product form. This will assure that not only the state of the catalyst remains the same, but also that it remains uncorrelated with the system. This is also the regime in which all of the second laws must be satisfied in order to transform one state into another. The case when $\epsilon_C = 0$ but arbitrary correlations between S and C are allowed to build up has been thoroughly studied in (Müller, 2018). There it was found that using a finely-tuned catalyst one can transform ρ_S into σ_S, as long as the free energy of ρ_S is higher than the free energy of σ_S. This leads to the conclusion that only one of the family of second laws remains, namely the non-equilibrium Helmholtz free energy F_1. Moreover, the authors of (Brandão et al., 2015) showed that when the error on the catalyst scales *linearly* with the number of particles (up to a constant factor) $n = \log d_C$, that is when $\epsilon_C \sim 1/n$,

then the non-equilibrium Helmholtz free energy F_1 again completely describes all possible transformations. Finally, in (Ng et al., 2015) it was found that the second laws completely vanish (meaning that all state transitions become possible) when the error on the catalyst surpasses a certain threshold. This threshold was determined to be:

$$\epsilon_C^{\text{bnd}} = \frac{d_S - 1}{1 + (d_S - 1)\log d_C} \sim \frac{1}{n}. \tag{3.6}$$

In other words, this is the minimal error which can be achieved under the assumption that all states can be converted between each other. Intuitively, one might view this as a process where the resource required to carrry out the transformation is extracted from the catalyst and used to facilitate the transformation. This sets a boundary between genuine catalysis and exploiting the catalyst; whenever error on the catalyst scales with its dimension better than (3.6) then there are state transitions which are not allowed and so the partial order induced by the second laws is recovered to a certain degree. Whenever the scaling of ϵ_C is worse or equal to (3.6) we will refer to the corresponding regime of catalysis as *embezzlement* regime. On the contrary, we will use the term *genuine catalysis* to indicate that transformations are still governed by the second laws (or some non-empty subset of them).

3.2.4 Embezzlement

In the last section we saw that expressing error on the catalyst using trace distance leads to certain difficulties, e.g. thermodynamic work can be extracted from a thermal state with arbitrarily small disturbance on the catalyst. This phenomenon is generally known as *thermal embezzlement* and clearly contradicts the first two laws of thermodynamics.

Clearly, embezzlement is a form of cheating and as such, our bookkeeping should make it clear when we use the catalyst properly, or when work — or other resources — are being embezzled during the transformation. It turns out that there are two natural ways in which we can eliminate embezzlement from thermodynamics. The first approach is to change the distance measure quantifying the error. For example, Ref. (Brandão et al., 2015) proposed to consider as the thermodynamic distance measure the so-called *work distance* which they defined as:

$$W_{\text{dist}}(\rho, \sigma) := kT \inf_{\alpha} \left[F_\alpha(\rho) - F_\alpha(\sigma) \right] \tag{3.7}$$

Using work distance in (3.5) allows to solve the embezzlement problem, i.e. small disturbance as quantified by work distance implies small work cost in recovering the initial catalyst state. However, this modification in the same time leads to new difficulties; First, it is not clear if work distance has any direct operational meaning and secondly, computing it requires solving a relatively difficult optimisation problem.

An alternative approach involves keeping track of how the error on the catalyst scales with its dimension. This has been a fruitful approach, as it allows to precisely determine when the error is large enough for the embezzlement to occur.

Even though embezzlement is not a proper form of catalysis, it is still an interesting phenomenon which has found several important applications mostly in the resource theory of pure-state entanglement. The power of embezzling has been exploited in several areas of quantum information, such as coherent state exchange protocols (Leung, Toner, and Watrous, 2008) or entangled projection games (Dinur,

Steurer, and Vidick, 2013). Moreover, embezzlement can also be viewed as a protocol for hiding quantum states from external observers. With this interpretation it was used to prove the quantum version of the reverse Shannon theorem (Berta, Christandl, and Renner, 2011; Bennett et al., 2014).

Today embezzlement is still a mysterious concept and its full significance in a thermodynamic context still constitutes an important open problem. Even though its role is not well understood, quantifying which states can be used as embezzlers is a very relevant problem. Recent studies revealed a few families of universal embezzling states, both in the context of the resource theory of entanglement (Leung, Toner, and Watrous, 2008) and thermodynamics (Ng et al., 2015). Such universal embezzlers have the power to "catalyze" any state transformation. In the case of the resource theory of entanglement, a further study exposed another family of universal embezzling states and some of their properties were examined in (Leung and Wang, 2014). In general, however, very few such families of universal embezzlers are known, and the effects related to the dimension, entropy or energy of the embezzler have been hardly studied.

3.2.5 The state of the catalyst

One of the most ubiquitous aspects of the results on catalysis is their silence about the precise form of the state of the catalyst. In particular, the very result of (Brandão et al., 2015) only guarantees existence of a catalyst; it does not specify which state of the catalyst ω might be used to carry out the desired transformation. This theme is present in most results on catalysis; they are very often based on existence proofs, or more rarely - constructing a very specific and hardly feasible catalyst state. This, however, may be obscuring the true physical mechanism behind catalysis. Furthermore, it is still not well understood which properties of quantum states are relevant for catalysis. The second laws only guarantee the existence of the catalyst; even if they are satisfied by a pair of states, it may still be difficult to find which state catalyzes a particular process.

This intuition comes from our macroscopic experience: chemical reactions can be catalyzed only by appropriately chosen chemical compounds; similarly thermal machines need to be carefully tuned so that the desired transformation may happen. Therefore, both macroscopic intuition, as well as previous results hint that we should expect that this fine-tuning of catalysts is necessary to exploit the power of catalysis in thermodynamic protocols.

This leads us to one of the most important problems in the subfield of catalysis.

Main question

How can we find a catalyst that will be useful for a given transformation?

Consequently, we can further ask which properties of states make them good catalysts and how special are these states? Finally, not much is also known about thermodynamic properties of catalysts, like their average energy, entropy or dimension. As such, there is still a lot to be understood about catalysis.

In this chapter we approach these questions in detail. In particular, we extend the understanding of catalysis by reporting a surprising property of multi-copy states which we term *catalytic universality*. We describe numerical results indicating that generic large-dimensional states act as universal catalysts, given sufficient dimension.

3.3 Results

In this section we will present the main result of this Chapter, i.e. we describe a numerical analysis showing that generic large-dimensional states act as universal catalysts, given sufficient dimension.

3.3.1 Catalytic universality for generic catalysts

In this section we present numerical evidence which indicates that the catalytic universality phenomenon concerns almost all large dimensional states. While this analysis is entirely based on numerical results, we finish with a plausible conjecture that catalytic universality holds for any state, as long as the catalyst is composed of sufficiently many particles. The proof of this conjecture was published in (Lipka-Bartosik and Skrzypczyk, 2021a), though later on it was realised that the proof has a serious flaw. We leave the correction of the proof for the near future.

Throughout this section the Hamiltonians of the system S and the catalyst C are fully degenerate, meaning that $H_S \propto H_C \propto \mathbb{1}$. This will allow us to simplify both the presentation and numerical computation, since checking majorization is computationally easier than checking thermo-majorization. It should be noted that this does not reduce generality of our findings, as thermo-majorization criteria can always be expressed in terms of standard majorization using the embedding map (Brandão et al., 2015). Moreover, for the purpose of visualisation we focus here on the case when $d_S = 3$. This will allow us to describe the numerical findings in a more natural and visually appealing way. In the Appendix we report further numerical evidence which indicates that these conclusions naturally extend to larger dimensional systems.

Fixed initial and final state

Let us consider two states ρ_S and σ_S such that $\boldsymbol{p} = \mathrm{diag}[\rho_S]$ and $\boldsymbol{q} = \mathrm{diag}[\sigma_S]$ and chosen such that (i) they satisfy the corresponding second laws (3.2), meaning that $H_\alpha(\boldsymbol{p}) \leq H_\alpha(\boldsymbol{q})$ for all $\alpha \geq 0$ and such that (ii) the probability vector $\boldsymbol{p}$ does *not* majorize $\boldsymbol{q}$ and vice versa. In this way we know that neither of ρ_S and σ_S can be transformed into each other, but there exists a catalyst ω_C which can be used to facilitate the transformation from ρ_S to σ_S. For illustrative purposes let us choose the following two representative states:

$$\boldsymbol{p}^\star = (0.65,\, 0.2,\, 0.15), \qquad \boldsymbol{q}^\star = (0.5,\, 0.4,\, 0.1). \tag{3.8}$$

It is easy to check that $\boldsymbol{p}^\star$ and $\boldsymbol{q}^\star$ are incomparable using e.g. the concept of Lorenz curves (Horodecki and Oppenheim, 2013). In this part of the section we focus exclusively on these particular states.

Suppose now that we choose a probability distribution $\mathcal{P}_{\text{dist}}$ and draw d_C positive numbers according to this distribution. We organize them in a vector and then normalize, obtaining a valid probability vector. In this way we have a simple method of sampling random catalysts, which for a large dimension d_C well approximates drawing from the probability simplex. Let us denote with $\boldsymbol{c}^{(i)} = [c_1^{(i)}, c_2^{(i)}, \ldots, c_{d_C}^{(i)}]$ an i-th probability vector obtained via this method. Each $\boldsymbol{c}^{(i)}$ will model a random catalyst drawn according to a respective probability distribution. In this way the

elements of each such random catalyst are given by:

$$c_k^{(i)} = \frac{X_{\text{dist}}^k}{\sum_{k=1}^{d_C} X_{\text{dist}}^k}, \tag{3.9}$$

where X_{dist}^k is a random variable drawn according to the probability distribution $\mathcal{P}_{\text{dist}}$. In what follows we will consider three different distributions:

$$\mathcal{P}_{\text{ray}} \longrightarrow \text{Prob}(X_{\text{ray}}^k = x) \sim x e^{-x^2/2}, \tag{3.10}$$

$$\mathcal{P}_{\text{uni}} \longrightarrow \text{Prob}(X_{\text{uni}}^k = x) \sim const, \tag{3.11}$$

$$\mathcal{P}_{\text{exp}} \longrightarrow \text{Prob}(X_{\text{exp}}^k = x) \sim e^{-x}. \tag{3.12}$$

Next we fix the error which we can tolerate on the catalyst ϵ_C and repeat the process of sampling catalysts many times. Having done so we can now ask: how frequently does a randomly chosen catalyst catalyzes a given state transformation for a fixed error? Counting the frequency of cases in which the following transformation is possible:

$$\boldsymbol{p}^\star \otimes \boldsymbol{c}^{(i)} \longrightarrow \boldsymbol{q}^\star \otimes \widetilde{\boldsymbol{c}}^{(i)}, \tag{3.13}$$

with $\widetilde{\boldsymbol{c}}^{(i)}$ chosen such that $\left\|\widetilde{\boldsymbol{c}}^{(i)} - \boldsymbol{c}^{(i)}\right\|_1 \leq \epsilon_C$, leads to the success probability denoted with $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$. This is an estimate of the probability that a randomly chosen catalyst can help in facilitating a given state transformation, with the disturbance on the catalyst at most ϵ_C. Moreover, using the results of Ref. (Horodecki, Oppenheim, and Sparaciari, 2018) we can readily determine the final state of the catalyst $\widetilde{\boldsymbol{c}}^{(i)}$ to be the so-called ϵ-flattest state [see (ibid.) for the general method of constructing these states]. To summarise, the success probability $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$ is computed using the following set of steps.

Algorithm 1: Estimating $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C)$ by sampling

Input: $\boldsymbol{p}$, $\boldsymbol{q}$, d_C, ϵ_C, $\mathcal{P}_{dist}$
Output: Estimate of $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C)$
Parameters: N_C `// precision of estimation`
$pos = 0$
foreach $i \in \{1, \ldots, N_C\}$ **do**
  $\boldsymbol{c}^{(i)} \leftarrow$ random catalyst sampled using $\mathcal{P}_{dist}$
  if there exists $\widetilde{\boldsymbol{c}}^{(i)}$ s.t. $\boldsymbol{p} \otimes \boldsymbol{c}^{(i)} \longrightarrow \boldsymbol{q} \otimes \widetilde{\boldsymbol{c}}^{(i)}$ and $\left\|\widetilde{\boldsymbol{c}}^{(i)} - \boldsymbol{c}^{(i)}\right\|_1 \leq \epsilon_C$ **then**
    $pos \leftarrow pos + 1$
$p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C) \leftarrow pos/N_C$

The results of this numerical experiment are summarized in Fig. 3.1. As we can see, when we increase the dimension d_C, the probability that a randomly chosen catalyst can catalyze a given transformation increases and very rapidly approaches a fixed value. This value, as well as the rate at which it is approached, depends on the specific distribution $\mathcal{P}_{\text{dist}}$ we choose. This indicates that the success probability $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$ depends both on the dimension of the catalyst and the distribution of its eigenvalues. This numerical experiment allows us to conclude that:

Random states are universal catalysts

There are state transformations for which random states act as catalysts with high probability.

Fixed initial state and arbitrary final state

In the previous section we studied how useful random catalysts are for a fixed state transformation. We now go one step further and generalise this investigation to arbitrary final states, while still keeping the initial state fixed. Let us then consider again the state $\boldsymbol{p}^\star$ given by Eq. (3.8) as the input state and let $\boldsymbol{q}$ be an arbitrary state. Since $\boldsymbol{p}^\star$ and $\boldsymbol{q}$ are d_S-dimensional probability vectors, it is useful to think of them as points in the space of all d_S-dimensional probability vectors, the so-called probability simplex Δ_{d_S} defined as:

$$\Delta_N := \left\{ \boldsymbol{x} = (x_1, \ldots, x_N) \,|\, x_i \geq 0 \text{ and } \sum_{i=1}^{N} x_i = 1 \right\} \tag{3.14}$$

Let us now define two sets of states inside Δ_{d_S}:

$$\mathsf{S}(\boldsymbol{p}) = \{\boldsymbol{q} \,|\, \boldsymbol{p} \to \boldsymbol{q} \text{ and } \boldsymbol{q} \in \Delta_{d_S}\}, \tag{3.15}$$

$$\mathsf{T}(\boldsymbol{p}) = \{\boldsymbol{q} \,|\, \boldsymbol{p} \otimes \boldsymbol{c} \to \boldsymbol{q} \otimes \boldsymbol{c},\ \boldsymbol{q} \in \Delta_{d_S} \text{ and } \boldsymbol{c} \in \Delta\}, \tag{3.16}$$

where Δ is the set of all N-dimensional probability simplices Δ_N for all natural N. The set $\mathsf{S}(\boldsymbol{p})$ contains all states $\boldsymbol{q}$ which are (thermo-) majorised by $\boldsymbol{p}$, that is all states which can be reached via thermal operations when starting from a state described by $\boldsymbol{p}$. Mathematically, such a set is often referred to as a *downward closure* at $\boldsymbol{p}$. The set $\mathsf{T}(\boldsymbol{p})$ contains all states $\boldsymbol{q}$ which can be reached via thermal operations with the help of some (unspecified) catalyst (i.e. a *catalytic downward closure*). We will refer to these sets as the thermal and the catalytic-thermal set, respectively. It can be readily verified that $\mathsf{S}(\boldsymbol{p}) \subseteq \mathsf{T}(\boldsymbol{p})$ for all $\boldsymbol{p}$. In this language, the main result of Ref. (Jonathan and Plenio, 1999) shows that $\mathsf{S}(\boldsymbol{p}) \subset \mathsf{T}(\boldsymbol{p})$ for some $\boldsymbol{p}$. Moreover, due to the results of Refs. (Daftuar and Klimesh, 2001; Turgut, 2007; Brandão et al., 2015) a complete characterisation of the set $\mathsf{T}(\boldsymbol{p})$ is known and whether $\boldsymbol{q} \in \mathsf{T}(\boldsymbol{p})$ is determined by the second laws (3.2).

Let us now perform our second numerical experiment. In the previous section we saw that for a fixed transformation, catalysts sampled from the exponential distribution (3.12) achieve a high probability of success $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$, even for moderate dimensions of catalysts. Let us now use the exponential distribution to sample random catalysts and compute the associated success probability. Furthermore, to assure that we do not work in the embezzlement regime, we also fix the allowable error on the catalyst to be $\epsilon_C = \mu \epsilon_C^{bnd}$, where ϵ_C^{bnd} is the embezzlement bound from Eq. (3.6) and $0 \leq \mu < 1$. For arbitrary points $\boldsymbol{q} \in \Delta_{d_S}$ we then estimate $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C)$, the probability that a random catalyst can be used to transform $\boldsymbol{p}^\star$ into $\boldsymbol{q}$ using the method described in Alg. 1.

The results of this numerical experiment are summarized in Fig. 3.2. For the purpose of illustration we also draw the sets $\mathsf{S}(\boldsymbol{p}^\star)$ and $\mathsf{T}(\boldsymbol{p}^\star)$. The numerics demonstrate that the probability $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C)$ is large for most $\boldsymbol{q}$ in the region $\mathsf{D}(\boldsymbol{p}^\star) := \mathsf{T}(\boldsymbol{p}^\star) \setminus \mathsf{S}(\boldsymbol{p}^\star)$ even for small dimension of the catalyst (e.g. when $d_C = 2^4$). Interestingly, the success probability increases significantly with the dimension of

the catalyst, so that for $d_C = 2^8$ the value of $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C) \approx 1$ for almost all points $\boldsymbol{q}$ inside $\mathsf{D}(\boldsymbol{p}^\star)$. As a consequence, we can infer the following.

Random states are universal catalysts

There are (input) states for which random states act as catalysts with high probability *for all* possible output states.

Arbitrary initial and final states

In the previous section we saw that a random catalyst can catalyze most of the possible state transformations for a fixed state $\boldsymbol{p}$, even for catalysts with a moderate dimension. In this section we extend our analysis and show that this behavior is a generic feature valid for arbitrary initial states.

Before going into the details, let us emphasize that not all initial states $\boldsymbol{p}$ lead to an interesting catalytic advantage. For example, when the system starts in a thermal state, $\boldsymbol{p} = \boldsymbol{g}$, there is no catalyst which can enhance the system's transformation potential. In that case the thermal and catalytic-thermal sets coincide, i.e. $\mathsf{S}(\boldsymbol{g}) = \mathsf{T}(\boldsymbol{g})$. Similar situation happens when the initial state of the system is a pure state, i.e. when it is already the maximally-resourceful state. One natural way to quantify the potential for a catalytic improvement is to estimate the volume of set defined as the difference between the thermal and the catalytic-thermal set. In what follows we will refer to such a set of states $\mathsf{D}(\boldsymbol{p}) := \mathsf{T}(\boldsymbol{p}) \setminus \mathsf{S}(\boldsymbol{p})$ as the *catalytic activation set* (CAS). Naturally, the volume of this region in the space of distributions largely varies between different initial states $\boldsymbol{p}$.

The aim of the next numerical experiment is to extend the results from the previous section to arbitrary initial states. We again fix a small error on the catalyst $\epsilon_C = \mu\epsilon_C^{bnd}$ to assure that we do not work in the embezzlement regime. We then sample uniformly the initial state $\boldsymbol{p}$ and compute the associated CAS, denoted $\mathsf{D}(\boldsymbol{p})$. For each point $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{p})$ we estimate the probability that a random catalyst can be used to transform $\boldsymbol{p}$ into $\boldsymbol{q}$ using the methods described in Alg. 1. Finally, we calculate the number of states inside $\mathsf{D}(\boldsymbol{p})$ for which the probability of catalysing using random catalysts, $p_{\text{succ}}(\epsilon_C)$, is larger than a fixed threshold value γ_{thd} [2]. This allows us to estimate, for each $\boldsymbol{p}$, the fraction $f(\boldsymbol{p})$ of all possible transformations which can be catalyzed using random catalysts with probability at least γ_{thd}, i.e.:

$$f(\boldsymbol{p}) := \frac{|\widetilde{\mathsf{D}}(\boldsymbol{p})|}{|\mathsf{D}(\boldsymbol{p})|}, \tag{3.17}$$

where $\widetilde{\mathsf{D}}(\boldsymbol{p}) \subseteq \mathsf{D}(\boldsymbol{p})$ and $\boldsymbol{q} \in \widetilde{\mathsf{D}}(\boldsymbol{p})$ if and only if $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C) \geq \gamma_{thd}$. The quantity $f(\boldsymbol{p})$ can be estimated using the following simple algorithm:

[2] It's worth to remind that the transformations which we consider here are always deterministic and the probability $p_{\text{succ}}(\epsilon_C)$ refers to the sampled catalysts, i.e. how probable it is to *deterministically* catalyze a given transformation using a catalyst drawn at random.

Algorithm 2: Estimating $f(\boldsymbol{p})$ by sampling

Input: $\boldsymbol{p}, d_C, \epsilon_C, \gamma_{thd}, \mathcal{P}_{dist}$
Output: Estimate of $f(\boldsymbol{p})$
Parameters: N_S

$\mathsf{A}(\boldsymbol{p}) \leftarrow$ initialise uniformly N_S states of dimension d_S
$\mathsf{S}(\boldsymbol{p}) \leftarrow$ all states in $\mathsf{A}(\boldsymbol{p})$ satisfying (3.15)
$\mathsf{T}(\boldsymbol{p}) \leftarrow$ all states in $\mathsf{A}(\boldsymbol{p})$ satisfying (3.16)
$\mathsf{D}(\boldsymbol{p}) \leftarrow \mathsf{T}(\boldsymbol{p}) \setminus \mathsf{S}(\boldsymbol{p})$

$pos = 0$
foreach $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{p})$ **do**
  $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C) \leftarrow$ compute using Alg. 1
  if $p_{succ}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C) \geq \gamma_{thd}$ **then**
    $pos \leftarrow pos + 1$
$f(\boldsymbol{p}) \leftarrow pos/|D(\boldsymbol{p})|$

The results of this numerical experiment are summarized in Fig. 3.3. Interestingly, even for relatively small catalyst dimensions (e.g. $d_C = 16$) there is a modest fraction of possible transformations that with high probability can be catalysed with a random catalyst. Furthermore, by increasing dimension of the catalyst this fraction improves significantly, so that already for moderate sized catalysts ($d_C = 256$) most of the possible transformations can be catalysed using random catalysts with a large probability. As a consequence, this numerical investigation allows us to infer the following.

Random states are universal catalysts

Most of all possible transformations can be catalytically activated with high probability using random states as catalysts.

In Sec. 3.5 we give analogous plots for a different choice of the relative error μ, the threshold value γ_{thld} and system dimension d_S, to demonstrate that this behavior is generic, i.e. it does not depend on our particular choice of parameters.

Conjecture for multi-copy states

In the previous sections we numerically showed that it is plausible to expect that catalysts with a sufficiently large dimension can, with high probability, be used as universal catalysts. We tried to describe this effect analytically and realized that the analysis greatly simplifies under the assumption that the catalyst starts in a multi-copy form. In that case a protocol that uses a multi-copy catalyst to activate all possible catalytic transformations has been proposed (Lipka-Bartosik and Skrzypczyk, 2021a). Later on we realised that our proof technique has a serious flaw and the protocol relies on the fact that, for any thermal operation, one can always find a suitable unitary that does not correlate the system with its environment. Since we do not know if this is true, the main protocol from (ibid.) may not be valid in general. We leave solving this conundrum for the near future. Still, we believe that the insights gained from the presented numerics, as well as the protocol described in (ibid.), make it a fair point to conjecture that catalytic universality might hold for all multi-copy catalysts comprised of sufficiently many copies.

Comparison with multi-copy states

In the previous sections we studied how useful random states are in catalysing thermodynamic transformations. In the final numerical experiment we check how these insights compare with the conjecture that multi-copy states make universal catalysts. In particular, we compute the quantity $f(\boldsymbol{p})$ defined in Eq. (3.17) using as a catalyst multiple copies of a fixed state.

We again fix the error on the catalyst to be $\epsilon_C = \mu\epsilon_C^{bnd}$ and compute the quantity $f(\boldsymbol{p})$ using Alg. 2, with the only difference that now $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C)$ is computed using a fixed multi-copy catalyst. Hence, it can be either 0 (the multi-copy catalyst does not allow to transform $\boldsymbol{p}$ into $\boldsymbol{q}$ within the allowed error on the catalyst) or 1 (when $\boldsymbol{p}$ can be transformed into $\boldsymbol{q}$ within the allowed error). The single-copy catalyst is chosen to be a qubit in a state $\omega_C = \text{diag}(1-r, r)$ with $0 \leq r < 1/2$. The results of this numerical experiment are summarised in Fig. 3.4. Comparing these with Fig. 3.3 illustrates that multi-copy catalysts generally achieve a larger fraction $f(\boldsymbol{p})$ *(i.e. fraction of activated transformations / all activable transformations) for a fixed catalyst dimension*. However, in the multi-copy case the improvement in $f(\boldsymbol{p})$ obtained by increasing the catalyst's dimension is generally smaller than in the case of random catalysts. This indicates that: catalytic universality using only few copies might be generally possible when the catalyst is sufficiently mixed. Finally, Fig. 3.4 demonstrates that some marks of catalytic universality can be observed for a relatively modest number of particles comprising the catalyst.

Summary of numerical findings

Although still somewhat preliminary in nature, these numerical findings strongly suggest that high-dimensional states, with high probability, will act as catalysts.. It seems reasonable to expect that a potential route to further analytic results will be to look for statements which hold with high probability. We leave this tantalising extension of this results for future work.

Finally, in the Sec. 3.5 we provide further numerical evidence that the catalytic universality phenomenon is a generic feature of sufficiently high dimensional catalysts. In particular, we (i) perform numerical calculations for larger dimensional systems S, (ii) change the sampling of catalysts to other distributions and (iii) choose different thresholds for the catalyst error ϵ_C and δ.

3.4 Discussion and open problems

In this chapter we have presented a surprising property of large-dimensional states: that, given sufficiently large dimension, generic states can act as a universal catalyst for most of potential transformations. We believe that this realisation is a step forward in our understanding of catalysis and provides new insights both in the field of quantum thermodynamics and resource theories, as finding the right catalyst is one of the main difficulties when employing catalysis in any resource theory.

What is more, we believe this opens the door for new avenues of exploration which will be of independent interest. In the following subsections we briefly sketch the most promising, in our opinion, directions of extending the results presented in this chapter.

3.4.1 The mechanism of catalysis

Since the seminal paper of Jonathan and Plenio (Jonathan and Plenio, 1999) our understanding of catalysis has grown significantly. However, we still do not fully understand the real mechanism behind catalysis and how it allows for lifting some of the restrictions imposed by allowable operations.

Here we made a step forward in explaining this mechanism. However, before a satisfactory understanding can be reached, several important challenges still need to be tackled. In particular, a long-standing open problem is determining which physical properties of states are important for catalysis? Moreover, we do not know how is the set of states reachable via catalytic transformations modified when additional constraints on the catalyst are made - e.g. in terms of energy, entropy or the distribution of its eigenvalues. What is the main property or "resource" relevant for catalysis? Our analytic results and preliminary numerics strongly suggest that dimension of the catalyst and distribution of its eigenvalues are both important properties which make a good catalyst. This also indicates a trade-off relation between catalyst dimension and its ability to catalyze transformations. Quantifying and understanding this potential trade-off will significantly advance our understanding of catalysis.

3.4.2 Catalytic universality for generic states

In Sec. 3.3.1 we presented a simple numerical evidence which indicates that the catalytic universality might appear for arbitrary large-dimensional catalysts. We believe that solving this problem will shed more light on the fundamental problem of what does the catalyst really do to facilitate the transformation. In particular, should we expect to find only specific catalysts if we modify some of our initial assumptions?

Another interesting way to proceed would be to study how important for the catalytic universality are correlations between the subsystems which form the catalyst. Looking more closely at the proofs presented here we can see that both in the embezzlement and genuine catalysis regime the main catalytic transformation $\mathcal{E}_{SC}$ does not build such correlations. In this respect, the only time where correlations can increase is during the pre and post-processing steps. However, since this potential increase in correlations is only due to the transformation error, we conjecture that it is not a necessary requirement for our results to hold.

In this respect, it would be also interesting to revisit the results from (Müller, 2018) and check whether in the regime where only correlations are allowed to build up (that is when the reduced state of the catalyst subsystems remain undisturbed), multi-copy catalysts can be still viewed as universal catalysts.

3.4.3 Catalytic universality and second laws for coherence

In our work we have not explored catalysis in the regime where states ρ_S and σ_S contain coherences between energy levels. It is known that in this case the second laws (3.2) provide only necessary but not sufficient conditions for state transformations. When considering fully general states with coherences, one has to additionally satisfy a completely new set of conditions resulting from the time symmetry constraints (Lostaglio, Jennings, and Rudolph, 2015). Loosely speaking, these new laws tell us that coherences between energy levels must decrease during thermodynamic transformation. In that case it would be interesting to see if the catalytic universality phenomenon can appear also for fully general coherent states.

3.4.4 Other potential directions

Extending catalytic universality to arbitrary QRTs. A natural question which arises when studying catalysis in the context of majorization-based QRT's is whether the catalysis phenomenon can be properly defined and studied for general resource theories as well. Interestingly, there are examples of QRTs for which catalysis does not enlarge the set of states which can be reached using free operations (Schmid et al., 2020). This leads to an interesting question: what are the necessary properties of a general QRT which allow it to have a nontrivial catalysis? Consequently, one can further ask if the catalytic universality phenomenon can also emerge for such theories. If not, then it would mean that catalytic universality is a unique feature of majorization-based QRTs and it would be interesting to see which special aspects of such theories allow for the catalytic universality?

3.5 Additional numerics

3.5.1 Numerical analysis for higher dimensional systems

In this section we present the results of a supplementary numerical computation that provides further evidence that catalytic universality is a generic phenomenon. In particular, we compute the fraction $f(\boldsymbol{p})$ for different dimensions of the system d_S and different choices of the error parameter μ and threshold value γ_{thld}. The results are presented in Fig. 3.5.

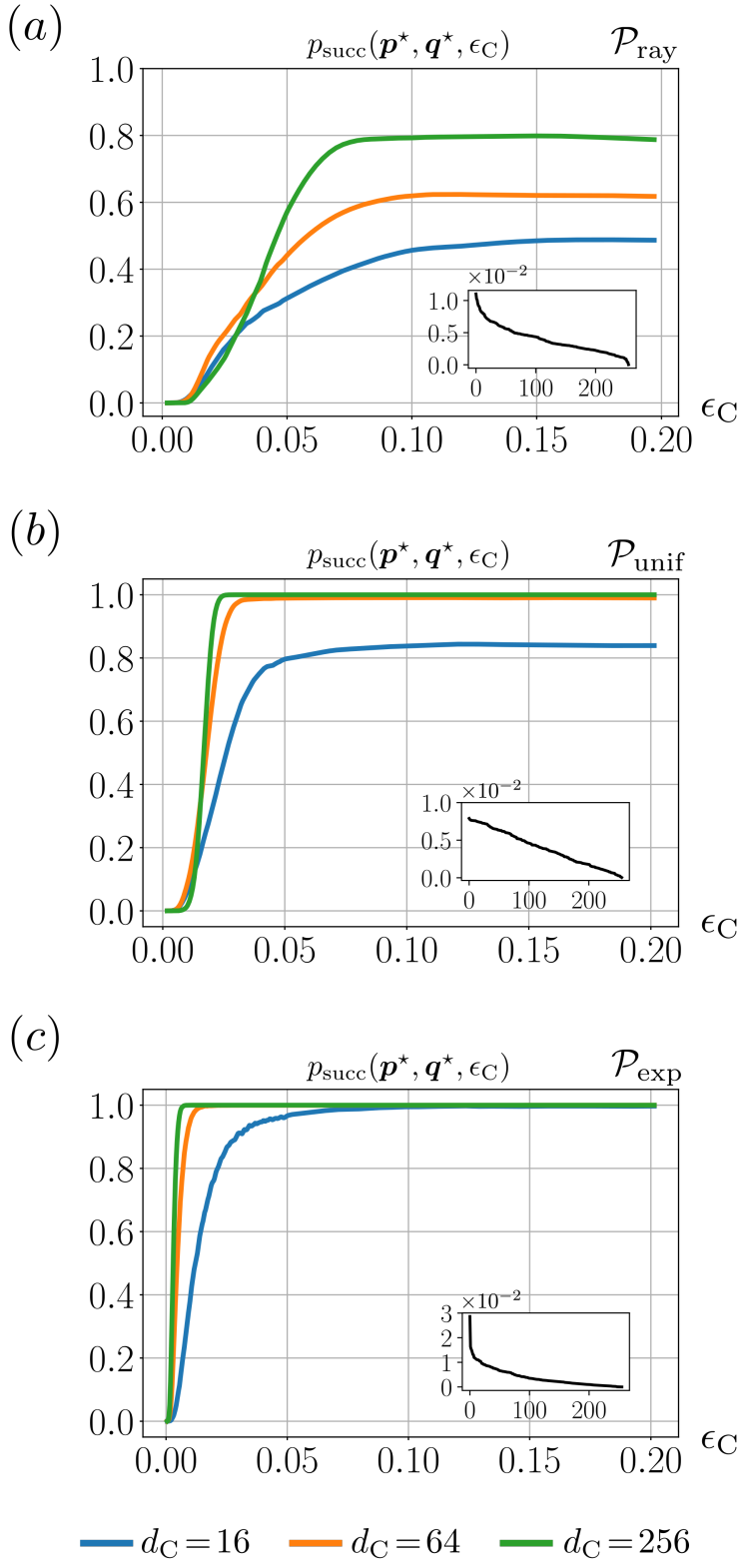

FIGURE 3.1: The probability $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$ that a random state of dimension d_C drawn from probability distribution $\mathcal{P}_{\text{dist}}$ can catalyze a fixed state transformation $\boldsymbol{p}^\star \to \boldsymbol{q}^\star$. Each panel corresponds to a different distribution: (*a*) Rayleigh $\mathcal{P}_{\text{ray}}$, (*b*) uniform $\mathcal{P}_{\text{unif}}$ and (*c*) exponential $\mathcal{P}_{\text{exp}}$. Inset plots illustrate an exemplary distribution of eigenvalues $c_k^{(i)}$ of a random catalyst $\boldsymbol{c}^{(i)}$ drawn according to a respective distribution and then normalized. We can see that in all cases $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}^\star, \epsilon_C)$ saturates at some point; This is means that for a fixed catalyst dimension, it may be impossible to activate all potential transformations.

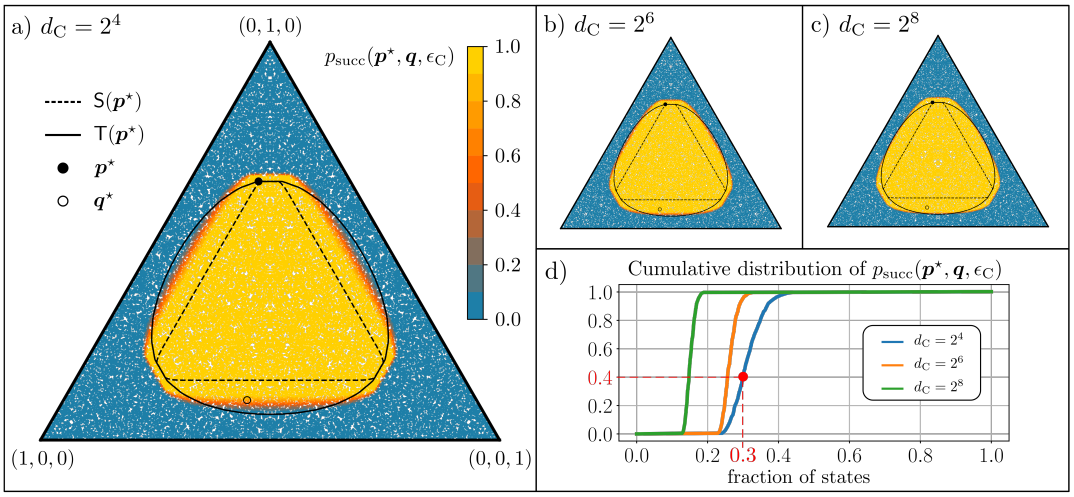

FIGURE 3.2: The probability $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C)$ that a random state can be used to enable the transformation from $\boldsymbol{p}^\star$ to $\boldsymbol{q}$ approximately catalytically, i.e. with an error $\epsilon_C = \mu\epsilon_C^{bnd}$ and $\mu = 0.1$. Plots $(a) - (c)$ correspond to different dimension of the random catalyst ($d_C = 2^4$, 2^6 and 2^8 respectively). States inside the region bounded by dashed lines define the thermal set $\mathsf{S}(\boldsymbol{p}^\star)$, consisting of all states that can be reached from $\boldsymbol{p}^\star$ using thermal operations. The solid line corresponds to the catalytic-thermal set $\mathsf{T}(\boldsymbol{p}^\star)$, consisting of all states which can be reached from $\boldsymbol{p}^\star$ with the help of some (potentially finely-tuned) catalyst. Note that the probability of success $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C)$ is generally very close to 1 for most final states $\boldsymbol{q}$ inside $\mathsf{D}(\boldsymbol{p}^\star) := \mathsf{T}(\boldsymbol{p}^\star) \setminus \mathsf{S}(\boldsymbol{p}^\star)$, even when the dimension of the catalyst is relatively small. Plot (d) illustrates the cumulative distribution of $p_{\text{succ}}(\boldsymbol{p}^\star, \boldsymbol{q}, \epsilon_C)$. To simplify interpretation, an exemplary point is drawn in red. It corresponds to the case when 30% of all states $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{p}^\star)$ can be reached using random catalysts of dimension $d_C = 2^4$, with probability less than 0.4. In other words, 70% of all possible catalytic transformations can be realised using random catalysts with probability 0.4 or higher.

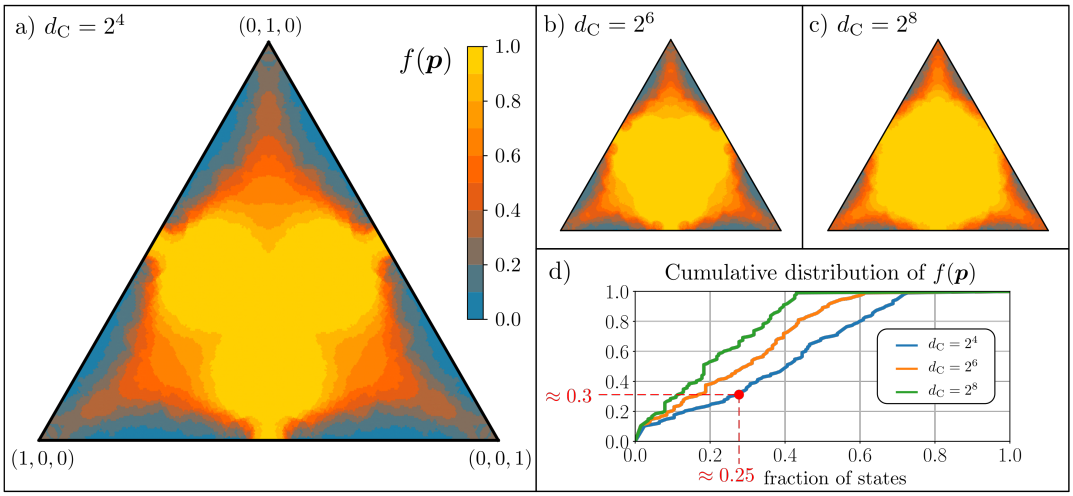

FIGURE 3.3: The fraction $f(\boldsymbol{p})$ of all states inside the catalytic activation set $\mathsf{D}(\boldsymbol{p})$ (CAS) for which the probability $p_{\text{succ}}(\boldsymbol{p}, \boldsymbol{q}, \epsilon_C)$ that a random state can be used as a catalyst, is larger than the threshold value $\gamma_{thd} = 0.9$. Plots $(a - c)$ correspond to random catalysts of dimensions 2^4, 2^6 and 2^8 respectively. Plot (d) illustrates the cumulative distribution of $f(\boldsymbol{p})$, therefore indicating what fraction of all possible transformations can ever be activated using a given catalyst. As an example, the red point corresponds to the situation where random states of dimension $d_C = 2^4$ are used to catalyze possible transformations. These random catalysts are not useful for roughly 25% of all possible initial states $\boldsymbol{p}$, i.e. for each such state they allow at most $\approx 30\%$ of all output states $\boldsymbol{q}$ in CAS to be reached with probability equal to or greater than γ_{thld}. In other words, such random catalysts can be used to reach more than 30% of all possible output states $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{p})$, with probability at least γ_{thld}, for at least 75% of all initial states $\boldsymbol{p}$.

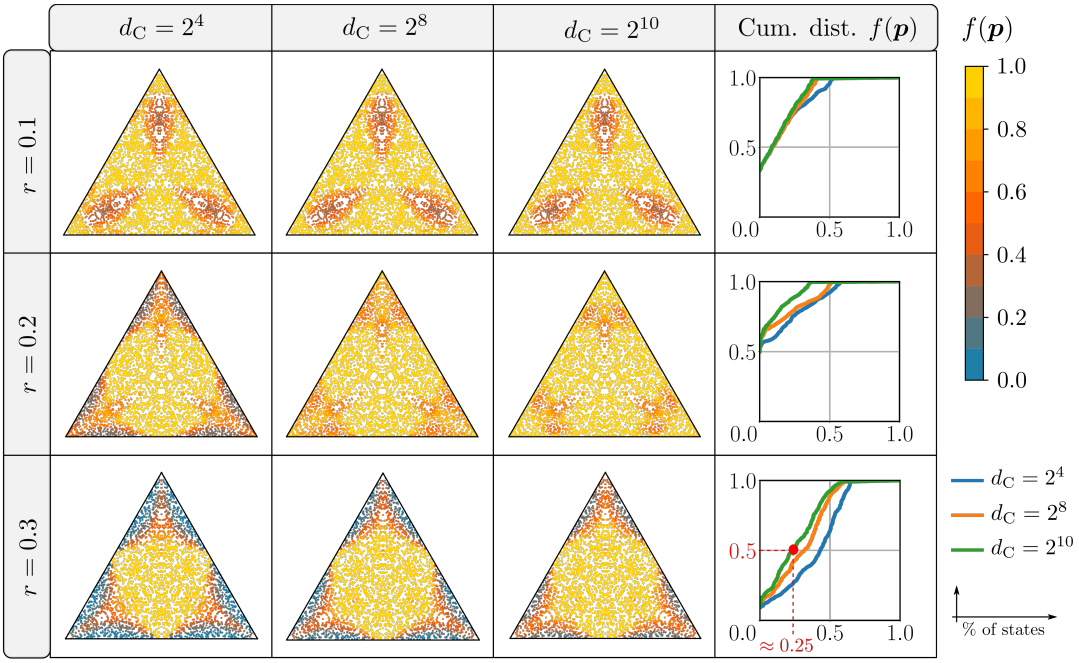

FIGURE 3.4: The fraction $f(\boldsymbol{p})$ of all states inside the catalytic activation set of $\boldsymbol{p}$ which can be catalysed using a multi-copy catalyst composed of $n \in \{4, 8, 10\}$ qubits $\omega_C = \mathrm{diag}(1-r, r)$, for different values of parameter r. Column 4 shows a cumulative distribution of $f(\boldsymbol{p})$, i.e. it illustrates the fraction of catalytic transformations that can be activated using a multi-copy catalyst. For example, the red exemplary point indicates that for around 0.25 of all possible input states $\boldsymbol{p}$, a quantum system composed of $n = 10$ qubits in a state with $r = 0.3$ acts as a catalyst for less than 0.5 of all possible output states $\boldsymbol{q}$ in CAS. Equivalently, for this choice of parameters, the multi-copy catalyst can activate more than 0.5 of all possible transformations for approximately 0.75 of all possible input states.

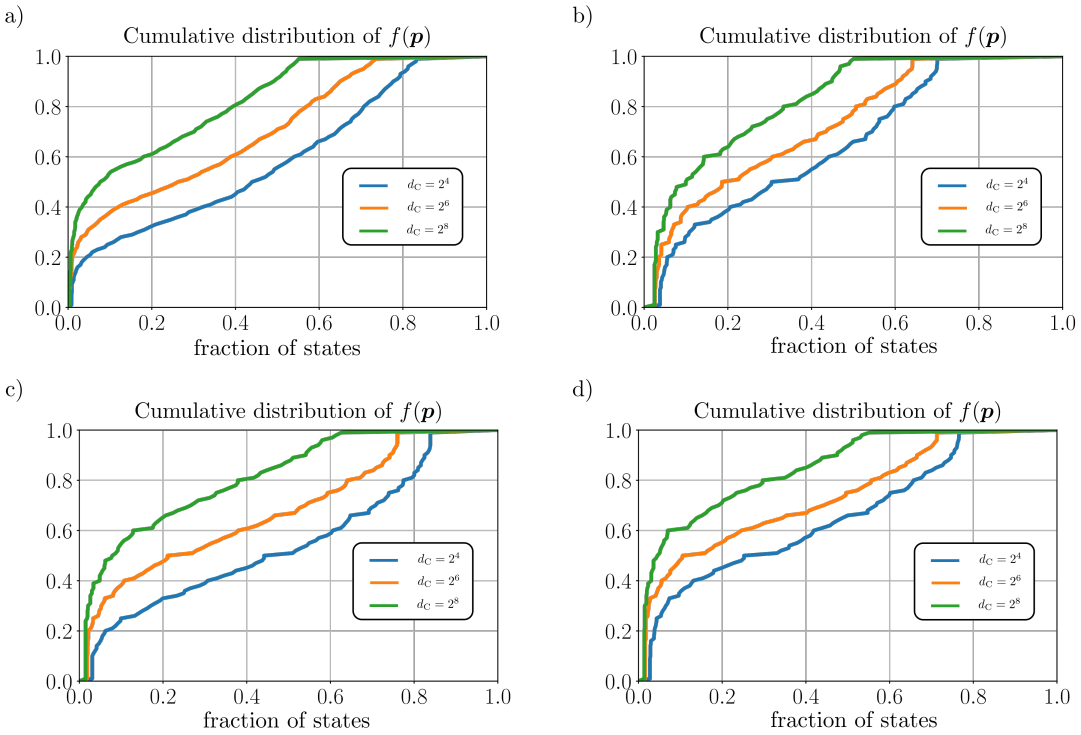

FIGURE 3.5: Cumulative distribution of $f(\boldsymbol{p})$ for different choices of parameters: (a) $[d_S = 4, \mu = 0.05, \gamma_{\text{thld}} = 0.9]$, (b) $[d_S = 5, \mu = 0.05, \gamma_{\text{thld}} = 0.9]$, (c) $[d_S = 4, \mu = 0.1, \gamma_{\text{thld}} = 0.8]$ (d) $[d_S = 5, \mu = 0.1, \gamma_{\text{thld}} = 0.8]$.

Chapter 4

Quantum batteries with a ground state

4.1 Introduction

4.1.1 Background

The second law of thermodynamics sets limits for all thermodynamic processes. It determines which state transformations are possible, regardless of the microscopic details of the process. From a practical point of view it imposes fundamental restrictions on the amount of average work $\langle w \rangle$ performed by the system which evolves from the state ρ towards ρ' and interacts with a thermal reservoir at a fixed temperature, i.e

$$\langle w \rangle \leq F(\rho) - F(\rho') = -\Delta F, \tag{4.1}$$

where $F(\rho)$ is the non-equilibrium free energy. The second law of thermodynamics is a statistical law and as such, it governs how thermodynamic systems behave when averaged over many realizations of the experiment. This information is relevant for a macroscopic observer, however, it does not provide much information about the microscopic details of the occuring process.

Recent developments in experimental techniques allow for manipulating and measuring systems at the nanoscale level, see for example (Koski et al., 2014; Koski et al., 2015; Chida et al., 2017; Camati et al., 2016; P. et al., 2016; Zanin et al., 2019). In order to take full advantage of these techniques it is crucial to understand how thermodynamic laws translate into the non-equilibrium domain, where fluctuations of thermodynamic quantities begin to play a significant role and averaged quantities are no longer enough to characterize their thermodynamic behaviour. This motivates extending thermodynamic framework to systems driven out of equilibrium, a setting which has been extensively studied in the recent literature (Jarzynski, 1997b; Jarzynski, 1997a; Alhambra et al., 2016; Richens and Masanes, 2016; Ito et al., 2019; Debarba et al., 2019).

Fortunately, there exist much stronger constraints on the possible distributions of thermodynamic work than the second law of thermodynamics (4.1). These constraints are often referred to as *fluctuation relations*. Arguably the best-known relation from this family is the Jarzynski equality which for a system composed of n particles implies that the probability of extracting more work per particle than the free energy difference vanishes exponentially the number of particles composing the system (Jarzynski, 1997b). The Jarzynski equality has been thoroughly studied both theoretically (Cohen and Mauzerall, 2004; Palmieri and Ronis, 2007) and experimentally

(Douarche et al., 2005; An et al., 2015; Hoang et al., 2018), and also formulated in the fully coherent case (Kurchan, 2000; Tasaki, 2000; Alhambra et al., 2016).

4.1.2 Motivation

A standard and often implicitly accepted assumption in thermodynamics is the existence of a perfect work receiver. This allows to treat the whole free energy difference as thermodynamic work. However, in some cases the existence of such a perfect receiver is difficult to motivate — especially in the quantum regime, where the device has to be modelled itself as a quantum system. The common approach is to consider the explicit work receiver (or the "battery") as an ancillary system which interacts with the working body and the heat bath during the transformation. In such situations the battery device is often modelled using either a classical Hamiltonian (Deffner and Jarzynski, 2013) or a Hamiltonian unbounded from below (Alhambra et al., 2016). Both of these approaches, although perfectly valid in the classical regime, cannot be justified when the battery's energy is close to its ground state. Therefore, in order to understand the effects arising in small-scale thermodynamics, it is important to understand how the presence of the battery ground state influences thermodynamic protocols.

This approach is fundamental to the modern (resource-theoretic) program of thermodynamics which attempts to systematically account for all possible resources involved in the process. One then studies their conversions under a restricted set of physically motivated interactions between the system, the heat bath and the battery device. In the most general case this interaction is modelled using an arbitrary energy-preserving unitary acting on these three subsystems. This is precisely the framework of thermal operations which we have already seen in Chapter 2. Thermodynamic work is then defined by specifying an "explicit battery model", which amounts to:

(i) specifying the Hamiltonian of the ancillary battery system,

(ii) the work quantifier (e.g. average energy),

(iii) the allowed interactions between the battery, the system and the bath (or restrictions on the allowable battery states).

The two most widely used models in the literature involve a qubit battery (wit) where the battery is in a pure state at all times and work is defined using average energy, and an ideal weight battery. In this model work is also defined using average energy of the battery, however the global unitary in addition has to commute with the translations on the battery, leading to the notion of translational invariance (TI). This is, however, only possible when the energy spectrum of the battery ranges over the whole real line, i.e. the battery is unbounded from both above and below. The ideal weight model is a common way of defining work in the quantum regime which was considered for the first time in (Skrzypczyk, Short, and Popescu, 2014) and then utilized to prove several fundamental results in the field of quantum thermodynamics (Alhambra et al., 2016; Åberg, 2014; Åberg, 2018; Masanes and Oppenheim, 2017; Richens and Masanes, 2016). Importantly, this additional assumption has powerful physical implications: it assures that work satisfies the second law of thermodynamics (Masanes and Oppenheim, 2017) and furthermore leads to the Jarzynski equality (Alhambra et al., 2016).

Naturally, one can argue that the ideal weight is not a realistic model of a battery as it does not have a ground state energy. On the other hand, one can also argue that

if the transformation is performed sufficiently far from the vacuum, the evolution assisted with a battery with a ground state (which we will refer to as the *physical battery*) should be equivalent to the evolution assisted with the ideal weight (Åberg, 2014; Faist, 2016) and should reproduce the same quantitative results (i.e. work distributions) as a physical battery. However, the nature we observe often does not follow this scheme; the existence of the ground state can be very often perceived, no matter how far we are from it. A basic question then appears:

Main question

Does the existence of the vacuum state of the battery have any implications for thermodynamic processes?

In this chapter we will carefully examine this question and will try to develop a satisfactory and comprehensive answer.

4.1.3 Contribution

In this Chapter we examine two different regimes of battery operation: the *high-energy* regime in which the average energy of the battery is far from the ground state energy (i.e. population of the ground state is small) and *low-energy* regime in which the occupation of the ground state cannot be ignored. While we corroborate the intuition that the battery with vacuum essentially behaves like the ideal weight in the high energy regime, we investigate quantitatively different predictions to which it leads in the low-energy regime, in particular with respect to the Jarzynski equality, second law of thermodynamics and fluctuations of work.

More precisely, we introduce a generalization of the translational invariance property which we term effective translational invariance (ETI). This allows to consider thermodynamic processes with broken translational symmetry, and hence provides a convenient tool for studying quantum thermodynamics using arbitrary physical batteries.

We further show that ETI in the framework of thermal operations implies a generalization of the standard Jarzynski equality, i.e., a family of inequalities which impose looser constraints on the allowable work distributions than the standard Jarzynski equality. This is our first main result.

Using these new inequalities we show that one can still recover the second law of thermodynamics in an approximate form, i.e., with correction terms decaying exponentially fast with the distance to the ground state. This allows to identify heat contributions to the average energy change when the battery is bounded from below and shows that global translational symmetry is not necessary to satisfy the second law of thermodynamics. Importantly, these deviations are related to the average work rather than just work and hence they are of a completely different character than the deviations reported by standard fluctuation theorems. This is the second main contribution of this chapter.

In the next part of the chapter we show that our model correctly reproduces the single-shot results on the work of formation originally derived using qubit as the battery system (Horodecki and Oppenheim, 2013). This answers an open problem from the field of quantum thermodynamics by showing that the notions of *single-shot* and *fluctuating* work can be both properly defined and studied for a battery with a ground state. This is the third main result of this chapter.

In the last part of the chapter we study a paradigmatic example of Landauer erasure and compare how the two battery models (the ideal weight battery and the harmonic oscillator with vacuum) behave when assisting this type of transformation. Notice that the second case contrasts with the classical treatment of work as the battery can no longer supply arbitrary values of energy when it operates close to its vacuum state. Because of this both batteries experience quantitative differences in their behavior in the low-energy regime. We show that for any battery model the very existence of the vacuum state implies a lower bound on the size of work fluctuations whenever energy is taken from the battery. In this way the occupation of the vacuum state is the fundamental factor which forbids performing Landauer erasure with deterministic (or arbitrarily concentrated) work distribution. We further show that the ideal weight violates this bound and hence allows for transformations which cannot be achieved in the low-energy regime.

We finish the chapter with a short summary and present several related open problems which we believe to be relevant for the field of quantum thermodynamics.

4.1.4 Structure

In Sec. 4.2 we describe an extension of the framework of thermal operations that accounts for using battery systems. Then in Sec. 4.3 we describe our main results. In particular, in Sec. 4.3.1 and Sec. 4.3.2 we describe the consequences of introducing the battery's ground state for thermodynamic protocols. In Sec. 4.3.3 we show that the model of a harmonic oscillator battery reproduces deterministic work and in Sec. 4.3.4 we analyse the problem of fluctuations during Landauer erasure using two battery models: a harmonic oscillator and an ideal weight. We summarise this Chapter in Sec. 4.4 and discuss several open problems.

4.2 Thermal operations with work

In Chapter 2 we described the resource theory of thermal operations, our main working horse throughout this chapter. Let us start by recalling that in this resource theory, the necessary and sufficient conditions for state transformations involving the system and external bath, are given by thermomajorisation criteria. However, in order to provide a complete description, we must extend this model to the case when thermodynamic work can be performed.

A natural question to ask is how thermomajorisation criteria change when thermodynamic work is allowed. Arguably the most widely-studied modification of the framework involves adding an ancillary battery system W and studying transformation of the form (2.77) where system S is now replaced by the joint system SW. One then identifies thermodynamic work as the change in the average energy of this ancillary system. More precisely, the battery is measured in the energy eigenbasis before and after the process, yielding outcomes ϵ and $\epsilon + w$ respectively. Work w is then a random variable distributed according to:

$$p(w) = \int_{\epsilon} \mathrm{d}\epsilon \ \mathrm{tr}\left[(\mathbb{1}_S \otimes \Pi_{\epsilon+w} \otimes \mathbb{1}_B)\, U(\rho_S \otimes \Pi_\epsilon\, \rho_W\, \Pi_\epsilon \otimes \tau_B)U^\dagger \right] \tag{4.2}$$

where $\Pi_x = |x\rangle\langle x|_W$ is a projector onto one of the energy eigenstates of W. This way of defining work, however, leads to two fundamental questions:

(i) When can we treat work stored in the battery as a legitimate thermodynamic work?

(ii) What are the necessary and sufficient conditions for the *work-assisted* state transformations $\rho_S \xrightarrow{TO} \sigma_S$?

It is perhaps surprising that answering those questions relies on which model of the battery we use. In the next sections we will describe and discuss properties of the main battery models found in the literature.

4.2.1 Thermal operations with a qubit battery

The simplest approach proposed in (Horodecki and Oppenheim, 2013) is to use as a battery a two-level qubit (wit) with a tunable energy gap δ and a Hamiltonian:

$$H_W = \delta \, |1\rangle\langle 1|_W \tag{4.3}$$

Furthermore, one then assumes that the battery always starts either in the ground or the excited state. This simple construction allows to introduce a notion of work valid in the single-shot regime and was described in Chapter 2. The arising work is referred to as *deterministic work* and is defined as the optimal δ for which there exists a TO realizing:

$$\rho_S \otimes |i\rangle\langle i|_W \xrightarrow{TO} \sigma_S \otimes |j\rangle\langle j|_W \, . \tag{4.4}$$

where $(i, j) = (0, 1)$ when the work is stored in the battery (distillation) and $(i, j) = (1, 0)$ when it is consumed (formation). Since the battery starts and ends up in a pure energy eigenstate, its energy difference can be fully associated with the work that needs to be performed (or can be extracted) during the transformation. To determine whether a given state ρ_S can be converted into σ_S using δ of work one has to check the thermomajorization criteria between the joint states of SW [see Chapter 2 for the details].

Naturally, transformations in which the wit ends up in a pure state are practically impossible to achieve. That is why for a theory to be applicable to realistic protocols one should consider transformations of the form (4.4), where now the battery is allowed to finish in a slightly mixed state. However, at this point it is not clear if we can interpret the average energy change of the wit as valid thermodynamic work. Since its entropy has changed, it is impossible to differentiate it between work and heat. To see this more explicitly, consider the following example.

Example 1: Thermalization of a wit

Consider a process which outputs a joint Gibbs state irrespective of the input, i.e.:

$$\forall \, \rho_S, \rho_W \qquad \rho_S \otimes \rho_W \xrightarrow{TO} \tau_S \otimes \tau_W . \tag{4.5}$$

Notice that this is a valid thermal operation as it always preserves the total Gibbs state. Consider applying this map to the state $\rho_S \otimes \rho_W = \tau_S \otimes |0\rangle\langle 0|_W$. The average work associated with this transformation is given by:

$$\begin{aligned} \langle w \rangle &= \mathrm{tr}\left[H_W \left(\tau_W - |0\rangle\langle 0|_W\right)\right] \\ &= \frac{\delta}{1 + e^{\beta\delta}} \geq 0. \end{aligned} \tag{4.6}$$

Hence, thermalization of a wit yields a positive amount of work on average, in contradiction to the second law of thermodynamics (4.1). Notice that the above example does not show that wit is a deficient battery model, but rather that it does not behave well in a non-ideal scenario. It may give us reasonable estimates for the work cost of a process only in the regime where the battery does not change its entropy. In this line, allowing for changing its entropy (however slightly) poses certain difficulties in interpreting its average energy change as thermodynamic work.

4.2.2 Thermal operations with an ideal weight battery

To resolve the problem we observed in the previous section we can use a different battery model, i.e. an ideal weight with Hamiltonian:

$$H_W = \int_{-\infty}^{\infty} x \, |x\rangle\langle x|_W \, \mathrm{d}x \tag{4.7}$$

where the basis $\{|x\rangle_W \mid x \in \mathbb{R}\}$ is formed from continuous orthonormal states representing the position of the weight. Notice that assuming this particular Hamiltonian is in some sense an arbitrary choice. We can also choose to store work in the kinetic energy of a moving particle or angular momentum of a rotating wheel. What is important is that the model is doubly-infinite, i.e. it has the capacity to store and provide arbitrary amounts of work for any possible transformation. This is of course not a realistic assumption as every physical implementation of the ideal weight must have an energetic minimum, as well as moving particles and rotating wheels eventually halt.

Furthermore, the global unitary U is assumed to commute with the shift operator on the weight. The shift operator is defined as $\Delta_y := \int_{-\infty}^{\infty} |x+y\rangle\langle x|_W \, \mathrm{d}x$. The commutation assumption can be then written as

$$\forall\, y \qquad [U_{SBW},\ \mathrm{id}_{SB} \otimes \Delta_y] = 0. \tag{4.8}$$

This additional assumption is often referred to as *translational invariance* (TI) and its importance was highlighted in (Masanes and Oppenheim, 2017) where it was shown that this combined with the energy conservation assures that dumping entropy into the joint state of the system and the battery (as we saw in the previous example) is impossible. TI is also enough to recover not only the second law of thermodynamics, but also quantum versions of the Jarzynski and Crooks fluctuation theorems (Alhambra et al., 2016). These results provide solid grounds to interpret shifts on the ideal weight as legitimate thermodynamic work. Moreover, the authors of (ibid.) also proved that the necessary and sufficient conditions for a work-assisted transformation between incoherent states in this case can be expressed by a generalization of thermomajorization which they termed Gibbs-stochasticity, i.e.:

$$\forall\, s' \qquad \sum_{s,w} p(s',w|s) e^{\beta(\widetilde{E}_{s'} - E_s + w)} = 1, \tag{4.9}$$

where $p(s',w|s)$ is the conditional probability of the final state of the system having energy levels $\widetilde{E}_{s'}$ (energy of the final Hamiltonian $\widetilde{H}_S$) and work w being done by the system, given that the initial state of the system had energy level E_s (energies of the

initial Hamiltonian H_S). Interestingly, by setting $w = 0$ one can recover from (4.9) the thermomajorization criteria, as described in (ibid.).

However, even though this approach is convenient mathematically and recovers the standard thermodynamic results, this proposal leads to new problems. First, Nature does not allow for Hamiltonians which are unbounded from below and thus it is not clear how well the ideal weight model describes any physical battery. Secondly, the conditions (4.9) lead to certain limitations when one tries to study deterministic work, as described by the following example:

Example 2: No deterministic work for ideal weight battery

Consider the following process realized using the ideal weight and a qubit system S with a fully-degenerate Hamiltonian $H_S = 0$, i.e.:

$$\forall \rho_S \qquad \rho_S \xrightarrow{TO} a\,|0\rangle\langle 0|_S + b\,|1\rangle\langle 1|_S\,, \tag{4.10}$$

where a, b are positive such that $a + b = 1$. When explicitly including the ideal weight we have that for all ρ_S the input state $\rho_S \otimes |0\rangle\langle 0|_W$ is transformed into:

$$\sum_{s,s',w} p_S(s)\, p(s', w|s)\, |s'\rangle\langle s'|_S \otimes |w\rangle\langle w|_W\,, \tag{4.11}$$

where $\rho_S = \sum_s p_S(s)\,|s\rangle\langle s|_S$ and we used the fact that due to TI we can start in an arbitrary energy eigenstate of the battery. Now the probability distribution $p(s', w|s)$ reads:

$$p(s', w|s) = p(s'|s)\, p(w|s, s'). \tag{4.12}$$

In our particular example we have $p(s' = 0|s) = a$ and $p(s' = 1|s) = b$ for all $s \in \{0, 1\}$. However, notice that if we now demand a deterministic work cost, i.e. $p(w) = \delta_{w,w^*}$ for some real number w^*, then conditions (4.9) imply:

$$s' = 0: \qquad w_* = -kT \log 2 - kT \log a, \tag{4.13}$$
$$s' = 1: \qquad w_* = -kT \log 2 - kT \log b. \tag{4.14}$$

This can only be satisfied if $a = b = 1/2$ and hence shows that it is impossible to obtain a fluctuation-free work when performing (4.10) using an ideal weight battery.

It is tempting to think that this behavior is a consequence of the third law of thermodynamics. However, note that in this case we have access to two infinitely big systems: an infinite heat bath and an (infinite) ideal weight and we put no constraints on how many degrees of freedom we may access. Notice also that if we lift the TI constraint and choose two distinct energy levels of the battery separated by an energy difference δ then we are able to exactly recover any deterministic transformation of the form (4.4) [see e.g. (Faist, 2016) for a proof of this statement]. Hence the problem must be somehow related to the TI property. Surprisingly, we will see that by assuming a less demanding notion of translational invariance we can recover both a proper behavior in the macroscopic limit (second law) and the desired behavior in the microscopic limit (deterministic work).

4.2.3 Thermal operations with a harmonic oscillator battery

Motivated by these realizations we now consider a battery model which has an energy spectrum bounded from below. Arguably the simplest model which satisfies this property is a harmonic oscillator:

$$H_W = \sum_{k=0}^{N} \epsilon_k \, |\epsilon_k\rangle\langle\epsilon_k|_W \,, \tag{4.15}$$

with $\epsilon_k := k\,\delta$ and N is the number of energy levels of the oscillator. For diagonal states any thermal operation $\mathcal{T}_{SW}$ acting on a harmonic oscillator battery can be fully characterized by a set of transition probabilities $\{r(s'k'|sk)\}$ which describe the probability that the state $|s\rangle_S \otimes |\epsilon_k\rangle_W$ gets mapped to another state $|s'\rangle \otimes |\epsilon_{k'}\rangle_W$. They can be extracted from $\mathcal{T}_{SW}$ via:

$$r(s'k'|sk) := \mathrm{tr}\Big(|s'\rangle\langle s'| \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}| \, \mathcal{T}_{SW}\big[\, |s\rangle\langle s| \otimes |\epsilon_k\rangle\langle\epsilon_k| \,\big]\Big). \tag{4.16}$$

For $\mathcal{T}_{SW}$ to be a valid thermal operation the associated transition probabilities must necessarily satisfy:

$$\forall s,k \qquad \sum_{s',k'} r(s'k'|sk) = 1, \tag{4.17}$$

$$\forall s',k' \qquad \sum_{s,k} r(s'k'|sk) e^{\beta(\widetilde{E}_{s'} - E_s + \epsilon_{k'} - \epsilon_k)} = 1. \tag{4.18}$$

The first line of these conditions means that $\mathcal{T}_{SW}$ is a trace-preserving map whereas the second line assures that the channel preserves the Gibbs state. Note that for a battery with spectrum bounded from below the commutation relation (4.8) cannot be satisfied for all y (unless the map is the identity map), and therefore a more general notion of translational invariance is needed. In many physically relevant situations it is impossible to have precise control over the global unitary acting on SBW and hence it may be very difficult to impose the TI condition in practice. Here, instead of constraining the global unitary, we will put constraints on the effective map on the system and the battery arising from this unitary. This is a much looser constraint than (4.8) and is potentially easier to implement in practice.

Effective translational invariance

Let us consider a discrete energy translation operator acting on the harmonic oscillator battery:

$$\Delta_n := \sum_{k \geq n} |k-n\rangle\langle k|_W \tag{4.19}$$

We are going to assume that the thermal operation $\mathcal{T}_{SW}$ is invariant with respect to translations of the battery only above a certain threshold energy $\epsilon_{\min} := \delta\, k_{\min}$ for some $0 \leq k_{\min} \leq N$. We will refer to this notion as effective translational invariance (ETI). Formally this means that the channel commutation relation:

$$\Big[\mathcal{T}_{SW}, \mathrm{id}_S \otimes \Delta_n[\cdot]\Delta_n^\dagger\Big] \, \big[\rho_{SW}\big] \, = 0, \tag{4.20}$$

holds for all states of the battery whose minimum occupied energy is above the threshold energy $\epsilon_{\min}$, i.e. $\forall\, \rho_{SW}$ s.t. $\mathrm{tr}\left[(\mathbb{1}_S \otimes \Pi_\epsilon)\rho_{SW}\right] = 0$ for $\epsilon < \epsilon_{\min}$. Notice that this condition is less stringent than the commutation relation (4.8), i.e. TI implies ETI but the converse is not true in general. In terms of transition probabilities $\{r(s'k'|sk)\}$ this condition can be restated as (see Sec. 4.5 for details):

$$r(s'k'|sk) = r(s', k' + n|s, k + n), \tag{4.21}$$

for all integer n such that $0 \leq k' + n \leq N$ and $k_{\min} \leq k + n \leq N$. In what follows we will refer to the set of energy levels below and above $\epsilon_{\min}$ as the *vacuum* and *invariant regime* respectively. Intuitively, the ETI assumption means that whenever a given process has a non-zero probability of taking the battery from $|\epsilon_k\rangle_W$ to $|\epsilon_{k'}\rangle_W$, then all transitions related to the associated work cost $w = \epsilon_{k'} - \epsilon_k$ are equally probable. For a graphical explanation see Fig. 4.1.

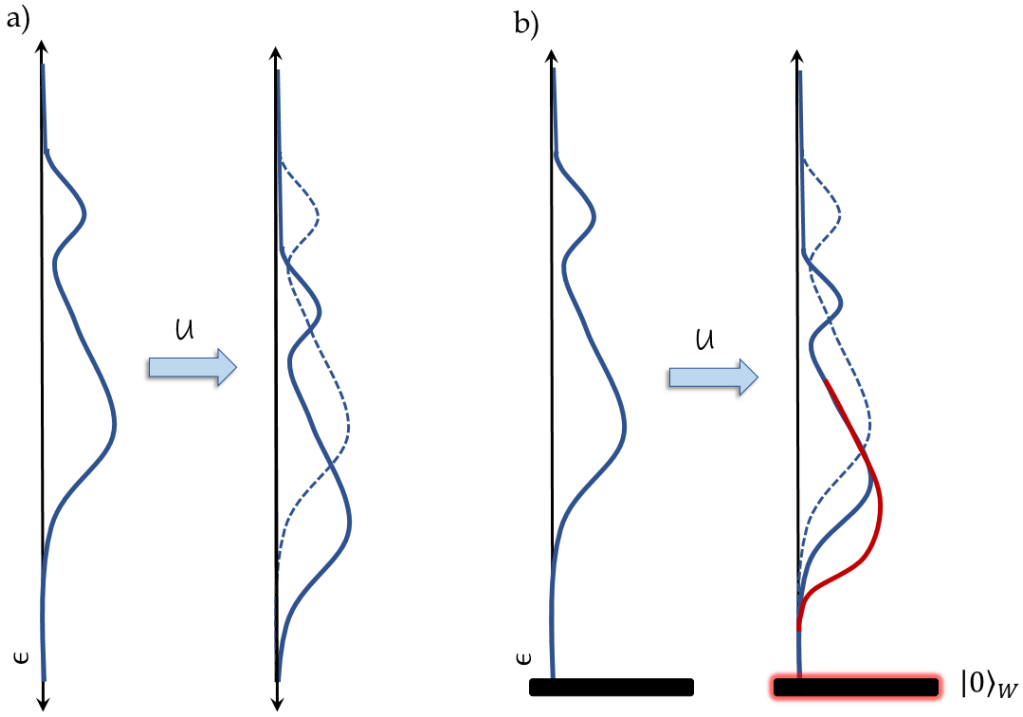

FIGURE 4.1: A battery with unbounded spectrum (a), attached to the system and environment, works as a tool to define work by changes of its average energy $\langle w \rangle$ during operations described by unitaries applied to the system, battery and environment. A common assumption that the unitaries commute with the shift operator on the unbounded battery (a) leads to the conclusion that probabilities of transitions on the battery are the same for the same energy gain on the battery, disregarding what is its initial state. For a physical model of a battery bounded from below, unitaries cannot commute with the shift operator, as the presence of the vacuum affects the final distribution of battery populations (b). We model this by allowing transitions emerging from levels below $\epsilon_{\min}$ to break translational invariance. This introduces corrections to the standard second law inequality for $\langle w \rangle$, which shows that the average change of energy of the battery ceases to serve as a good measure of work.

4.3 Results

4.3.1 Jarzynski equality for physical batteries

Let us now move to the more general case of thermodynamic protocols that allow for the Hamiltonian of the system to be changed. In this case the standard Jarzynski equality can be written as (Jarzynski, 1997b; Jarzynski, 1997a):

$$\langle e^{\beta w} \rangle = \frac{\widetilde{Z}_S}{Z_S}, \tag{4.22}$$

where $\widetilde{Z}_S$ and Z_S are partition function associated with the final and initial Hamiltonian on the system S. Recently, Alhambra *et. al* in (Alhambra et al., 2016) using the framework of thermal operations with the ideal weight derived a quantum analog of (4.22) relating fluctuations of work in an arbitrary thermodynamic process. In particular they showed that for states block-diagonal in the energy eigenbasis the following identity holds:

$$\langle e^{\beta (w - f_s)} \rangle = \widetilde{Z}_S, \tag{4.23}$$

where the fine-grained free energy $f_i := E_i + \frac{1}{\beta} \log p_S(i)$ and w are random variables, E_i is the energy of i-th level of the system S and averaging is over $p(s, w)$, that is, the probability that system S starts in level $|s\rangle_S$ and performs work w. The above equality can be thought of as an extension of (4.22) to the case when the initial state of the system is out of equilibrium. Moreover, it can be easily verified that when the initial state is thermal (i.e., $\rho_S \propto e^{-\beta H_S}$) then (4.23) reduces to the standard Jarzynski equality (4.22).

The proof of (4.23) presented in (ibid.) relies crucially on the TI assumption imposed on thermal operations. In the following theorem we show that when translational symmetry is broken the equality (4.23) turns into a family of Jarzynski-like inequalities. The crucial difference here is that in the regime of broken translational symmetry (vacuum regime) there are no constraints of this form. At first glance this may lead to the conclusion that Jarzynski equality in the form (4.23) can be violated arbitrarily well. However, in the following theorem we show that by looking at an analog of (4.23) conditioned on the battery level k, we can deduce nontrivial bounds which hold in the translationally invariant regime of our model. Moreover, in the next section we will show that these relations imply that the work fluctuations are constrained in such a way that the second law of thermodynamics still holds.

Before we present the main result of this section let us first rewrite (4.23) in a form which explicitly demonstrates its dependence on the initial battery state:

$$\rho_W = \sum_{k=0}^{N} p_W(k) \, |\epsilon_k\rangle\langle\epsilon_k|_W \, . \tag{4.24}$$

Rewriting (4.23) we get:

$$\langle e^{\beta(w - f_s)} \rangle = \sum_{k,k'} p_W(k) \langle e^{\beta(w_{kk'} - f_s)} \rangle_k, \tag{4.25}$$

where we denoted $w_{kk'} = \epsilon_{k'} - \epsilon_k$ and labelled:

$$\langle e^{\beta(w_{kk'}-f_s)}\rangle_k := \sum_{s,s',k'} p_S(s)\, r(s'k'|sk) e^{\beta(w_{kk'}-f_s)}.$$

With this in mind we can present our first main result of this section.

Jarzynski-like inequalities

Result 4.1. Let $\mathcal{T}_{SW}$ be a thermal operation acting on a harmonic oscillator battery and satisfying (4.20). Then for all $k \geq k_{min}$,

$$\langle e^{\beta(w_{kk'}-f_s)}\rangle_k \leq \widetilde{Z}_S \left(1 + e^{-\beta\delta_k}\right), \tag{4.26}$$

where $\delta_k := \epsilon_k - \epsilon_{\min} + \delta$ is the energy difference between the state $|\epsilon_k\rangle_W$ and the top of the vacuum regime $|\epsilon_{\min} - \delta\rangle_W$.

Importantly, this result does not constrain fluctuations of work when the battery starts in the energy subspace without translational symmetry ($k < k_{\min}$). In other words, if the symmetry (4.20) is violated for some energy subspace and the battery starts in that subspace, the r.h.s. of (4.23) can be made arbitrarily large (and in fact depends on the dimension of the battery as can be seen from the proof of the theorem). This phenomenon did not occur for the ideal weight because in that case (4.20) was satisfied for all energy levels (meaning that the battery had effectively no vacuum regime). When the initial state of the system is the equilibrium state then (4.26) leads to the following family of inequalities:

$$\forall\, k \geq k_{\min} \quad \langle e^{\beta w_{kk'}}\rangle_k \leq \frac{\widetilde{Z}_S}{Z_S}\left(1 + e^{-\beta\delta_k}\right), \tag{4.27}$$

Finally, the following example explicitly demonstrates that work can fluctuate arbitrarily for a valid thermal operation when the battery starts in the vacuum regime.

Deviation from the Jarzynski expression

Consider a qubit system S and a thermal operation $\mathcal{T}_{SW}$ with a constant trivial Hamiltonian $H_S = \widetilde{H}_S = 0$ and described by its action on the basis states:

$$|s\rangle\langle s|_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W \to |0\rangle\langle 0|_S \otimes |\epsilon_{k-1}\rangle\langle\epsilon_{k-1}|_W\,, \tag{4.28}$$
$$|s\rangle\langle s|_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W \to |1\rangle\langle 1|_S \otimes \gamma_W, \tag{4.29}$$

where $\gamma_W := \sum_{k'=0}^{N} 2^{-k'-1} |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W$ and $k > 0$. Below, we will take the limit $N \to \infty$, which assures proper normalisation, and with the energy gap $\delta = \beta^{-1}\log 2$ implies that the map is a valid thermal operation. The process can be fully characterized using transition probabilities:

$$k > 0: \qquad r(0,\, k-1|s,\, k) = 1, \tag{4.30}$$
$$k = 0: \qquad r(1,\, k'|s,\, 0) = 2^{-k'-1}, \tag{4.31}$$

where all other transition probabilities are equal to zero. Notice that the process satisfies (4.21) for all $k > 0$, i.e. the vacuum regime is spanned just by

one energy level $|\epsilon_0\rangle\langle\epsilon_0|_W$. Let us now compute the averaged term from the Jarzynski equality (4.25). We have:

$$\langle e^{\beta(w_{kk'}-f_s)}\rangle_k = \begin{cases} \widetilde{Z}_S & \text{if } k>0, \\ \widetilde{Z}_S(N+1) & \text{if } k=0. \end{cases} \tag{4.32}$$

This means that the average (4.25) goes to infinity (in the limit $N \to \infty$) whenever the initial occupation of the battery ground state is non-zero.

We described a thermal operation for which Jarzynski equality does not provide a meaningful bound on work fluctuations when the battery initially occupies an energy eigenstate for which ETI does not hold. According to Result 4.1, when the battery is initiated in such an eigenstate, these fluctuations may be higher than in the case of the ideal weight (for which the standard Jarzynski equality holds "everywhere"). Surprisingly, in such cases it is still possible to derive a form of the second law of thermodynamics with corrections depending on these occupations and the distance of the battery state to the vacuum regime. These correction terms quantify which part of the average energy change on the battery must be associated with heat rather than work. In the next section we present a general theorem which allows to quantitatively determine which part of the average energy change on the battery can be considered as heat and which as useful work.

4.3.2 Second law of thermodynamics for physical batteries

Thermodynamic work can be largely influenced by energy fluctuations in the system. Although both of them contribute to the change in the system's average energy, the work should be stored in an ordered form so that it can be later used for another transformation. At the same time heat is irreversibly dissipated and lost. In this section we describe the second main contribution of this chapter which is a modified version of the second law of thermodynamics valid for batteries bounded from below. This allows to estimate how much of the average energy change on the battery can be associated with thermodynamic work when its spectrum is bounded from below.

Second law for batteries with vacuum

Result 4.2. Let $\mathcal{T}_{SW}$ be a thermal operation satisfying (4.20). Then for all battery states of the form:

$$\rho_W = \sum_{k=0}^{N} p_W(k)\,|\epsilon_k\rangle\langle\epsilon_k|_W \tag{4.33}$$

the average work $\langle w\rangle$ satisfies

$$\langle w\rangle \leq -\Delta F_S + A_\beta(\rho_W,\rho_S) + B_\beta(\rho_W), \tag{4.34}$$

where the correction terms are defined as

$$A_\beta(\rho_W, \rho_S) := \sum_{k<k_{\min}} p_W(k)\left[E^{\max}_{S'} - F(\rho_S) - \eta_S \frac{\partial \eta_k}{\partial \beta}\right], \tag{4.35}$$

$$B_\beta(\rho_W) := \frac{1}{\beta}\log\left[1 + \sum_{k\geq k_{\min}}^{N} p_W(k)\, e^{-\beta\delta_k}\right], \tag{4.36}$$

where $\Delta F_S = F(\rho'_{S'}) - F(\rho_S)$ is the change in the non–equilibrium free energy of the system and the following notation is used:

$$\eta_k := Z_W\, e^{\beta\epsilon_k},\ \ \eta_S := Z_S\, e^{\beta E^{\max}_S},\ \ E^{\max}_{S'} := \max_s \widetilde{E}_s,\ \ \delta_k := \epsilon_k - \epsilon_{\min} + \delta. \tag{4.37}$$

Let us analyze the terms appearing in (4.34). The upper bound for average work has two correction terms which both depend on the initial state of the battery. First of them is proportional to the occupation below the threshold energy $\epsilon_{\min}$ and describes the contribution to the battery's average energy related to changing its entropy (using the battery in its vacuum regime spoils the battery). The second term decreases exponentially fast with the distance to the threshold $\epsilon_{\min}$ and effectively vanishes when battery starts far away from it. In particular, if the battery operates sufficiently far from the threshold energy, that is if $\sum_{k<k^*} p_W(k) \approx 0$ for some $k^* \gg k_{\min}$, then both correction terms vanish and (4.34) reduces to the ordinary form of the second law.

We stress here that the possible violation of the second law inequality, which can occur when the battery is initialized in the proximity of the vacuum regime, is an indication that in such case the average change of the energy of the battery can no longer be considered as a valid description of thermodynamic work. Essentially, this can be interpreted as showing that in principle, there are no ideal work reservoirs. In this way we can interpret Result 4.2 as a quantitative tool for determining the part of the average energy change on the battery that cannot be associated with thermodynamic work for arbitrary thermodynamic processes. In other words, Result 4.2 allows to determine how far the battery must be initialized in order to interpret the energy change on the battery as a genuine thermodynamic work.

Proof of Result 4.2 (sketch). The standard approach in deriving the second law of thermodynamics (4.1) is to start with the fluctuation theorem (4.23) and upper bound the average work using convexity of the exponential function. However, as we saw in a previous example, if the battery has an energy subspace for which the map is not translationally invariant, the r.h.s of (4.23) can be made arbitrarily large for some battery states. Hence we need to modify the method to obtain informative bounds. We start by decomposing the average work $\langle w\rangle$ into two terms, each related to a different regime of the initial state of the battery:

$$\langle w\rangle = \sum_w p(w)\, w = \langle w\rangle_{\text{vac}} + \langle w\rangle_{\text{inv}}, \tag{4.38}$$

where we labelled:

$$\langle w\rangle_{\text{vac}} := \sum_{w}\sum_{k<k_{\min}} p_W(k)\, p(w|k)\, w \tag{4.39}$$

$$\langle w\rangle_{\text{inv}} := \sum_{w}\sum_{k\geq k_{\min}} p_W(k)\, p(w|k)\, w. \tag{4.40}$$

Our strategy is to independently bound both terms appearing in (4.38). Regarding the first term note that all we know is that the transition probabilities $\{r(s'k'|sk)\}$ come from a stochastic map (4.17) which preserves the associated Gibbs state (4.18). In particular, this means that they are all upper bounded by respective Gibbs factors. This means that for all input and output pairs (s,k) and (s',k') we can write $r(s'k'|sk) \leq e^{-\beta(\widetilde{E}_{s'}-E_s+w_{kk'})}$, which after some manipulation leads to:

$$\langle w\rangle_{\text{vac}} \leq -\eta_S \sum_{k<k_{\min}} p_W(k)\frac{\partial \eta_k}{\partial \beta} \tag{4.41}$$

Consider now the second term from (4.38). Notice that now the sum runs over $k \geq k_{\min}$ and so the assumptions of Result 4.1 are satisfied. Hence, using Result 4.1 and the convexity of the exponential function it can be shown that:

$$\langle w\rangle_{\text{inv}} \leq -\Delta F_S + B_\beta(\rho_W) + \left(\sum_{k=0}^{k_{\min}-1} p_W(k)\right)\cdot (E_S^{\max} - F(\rho_S)). \tag{4.42}$$

The result follows by combining bounds (4.41) and (4.42). □

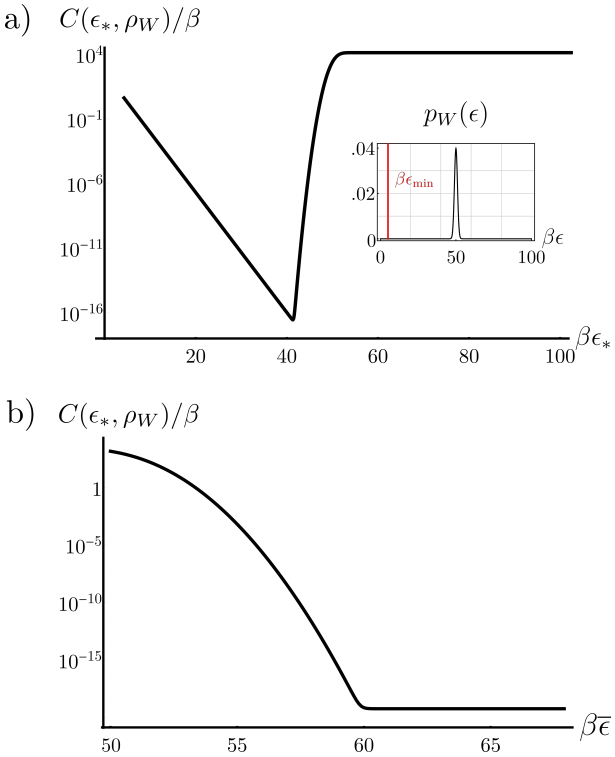

FIGURE 4.2: (a) Term $C(\epsilon_*,\rho_W)$ computed from the bound on average work from Corollary 6.1. The initial state of the battery with average energy $\beta\bar{\epsilon} = 50$ has a Gaussian profile $p_W(\epsilon) \propto e^{\beta^2(\epsilon-\bar{\epsilon})^2/2}$. The energy spacing is set to $\beta\delta = 0.1$ and vacuum regime to $\beta\epsilon_{\min} = 5$. System S is a qubit with $H_S = \widetilde{H}_S = 0$. Notice that in this case the initial state of the battery is fixed. Therefore, by increasing ϵ_*, we reach a point where $p(\epsilon < \epsilon_*)$ is no longer small ($\beta\epsilon_* \approx 40$) and the correction term blows-up. (b) Term $C(\epsilon_*,\rho_W)$ for the same setting, but with a fixed $\epsilon_* = 10\times\epsilon_{\min}$, plotted as a function of the average energy $\beta\bar{\epsilon}$ of the Gaussian profile. The kink around $\beta\bar{\epsilon} = 60$ appears as a consequence of fixing the threshold value ϵ_* in Eq. (4.45). In the same time, the first term from Eq. (4.45) vanishes exponentially fast with $\bar{\epsilon}$.

In order to illustrate Result 4.2 more clearly consider an infinite-dimensional battery ($N \to \infty$) with initial energy population $p(\epsilon < \epsilon_*)$ below some energetic

cut-off $\epsilon_* > \epsilon_{\min}$, that is:

$$p(\epsilon < \epsilon_*) = \sum_{k:\, \epsilon_k \leq \epsilon_*} \langle \epsilon_k | \, \rho_W \, | \epsilon_k \rangle_W \, . \tag{4.43}$$

The parameter ϵ_* will serve us to describe the initial state of the battery. In the following corollary we present a simplified (though slightly looser) bound derived from Result 4.2. As we shall see, in this bound the dependence of the correction term on the initial state of the battery is much simpler and in fact can be described by using just the function $p(\epsilon < \epsilon^*)$.

Corollary 4.1. For any thermal operation $\mathcal{T}_{SW}$ acting on a harmonic oscillator battery W with threshold energy ϵ_{min} and initial state ρ_W, satisfying (4.43), we have:

$$\langle w \rangle \leq -\Delta F_S + \frac{1}{\beta} C(\epsilon_*, \rho_W), \tag{4.44}$$

with

$$C(\epsilon_*, \rho_S) = p(\epsilon < \epsilon_*) \Big[c_S \, h(\beta, \delta, \epsilon_{min}) + \log c_S \Big] + c_S e^{-\beta(\epsilon_* - \epsilon_{\min})},$$

where we defined:

$$h(\beta, \delta, \epsilon_{\min}) := e^{-\beta\delta}[1 + \beta\delta e^{\beta\epsilon_{min}}(1 - e^{-\beta\delta})^{-2}], \tag{4.45}$$

$$c_S := d_S e^{\beta E_S^{max}} \, . \tag{4.46}$$

Notice that for battery states concentrated far from the vacuum regime the term $C(\epsilon_*, \rho_W)$ vanishes exponentially fast in the low energy regime (see Fig. 4.2). Proof of the corollary can be found in Sec. 4.5.

4.3.3 Recovering deterministic work

All operations allowed in the framework involving the ideal weight can by carried out using the harmonic oscillator battery with a bounded spectrum from below, as long as we use the battery sufficiently high above the ground state [see e.g. (Åberg, 2014; Åberg, 2018)]. However, the converse statement is not true. In this section we show that there are thermal operations which can only be accomplished using batteries bounded from below. That means that the sets of operations generated by these two models are not equivalent.

In this section we describe a method of extending arbitrary thermal operations defined on a wit to thermal operations acting on a harmonic oscillator battery. First of all, this construction shows that it is possible to recover the notion of deterministic (i.e. fluctuation-free) work for batteries with a ground state whenever they operate above the regime of broken translational symmetry. More intuitively this means that by properly breaking translational symmetry one can minimize to zero the fluctuations of work, while still satisfying the second law of thermodynamics.

Secondly, *any* thermal operation acting on a wit can be extended using the construction provided below. In Sec. 4.5 we show that any thermal operation arising from this construction satisfies the second law in the sense of Result 4.1. In this way, the maps which leave the wit in a mixed state can still be "salvaged" and lead to the average work which approximately obeys the second law of thermodynamics and for which the size of violations can be easily controlled.

Constructing thermal operations using harmonic oscillator batteries

Let Γ_{wit} be an arbitrary thermal operation acting on S and a two-level battery (wit) performing the transformation:

$$\rho_S \otimes |0\rangle\langle 0|_W \to \mathcal{R}_{00}(\rho_S) \otimes |0\rangle\langle 0|_W + \mathcal{R}_{01}(\rho_S) \otimes |1\rangle\langle 1|_W \,, \tag{4.47}$$

$$\rho_S \otimes |1\rangle\langle 1|_W \to \mathcal{R}_{10}(\rho_S) \otimes |0\rangle\langle 0|_W + \mathcal{R}_{11}(\rho_S) \otimes |1\rangle\langle 1|_W \tag{4.48}$$

Note that full information about Γ_{wit} is contained in the set of subchannels $\{\mathcal{R}_{kk'}\}$ which we will refer to as *battery subchannels*. Let $\{\mathcal{R}_{kk'}\} = \{\mathcal{R}_{00}, \mathcal{R}_{01}, \mathcal{R}_{10}, \mathcal{R}_{11}\}$ be (arbitrary) battery subchannels associated with the transformation on the wit. The transition probabilities $\{r(s'k'|sk)\}$ can be extracted from $\{\mathcal{R}_{kk'}\}$ by taking

$$r(s'k'|sk) = \mathrm{tr}\left[|s'\rangle\langle s'|\, \mathcal{R}_{kk'}[|s\rangle\langle s|]\right] \tag{4.49}$$

We will use the map Γ_{wit} as a primitive in constructing a family of thermal operations acting on a harmonic oscillator battery. We define a transformation Γ_{osc} acting on S and a harmonic oscillator battery W in the following way:

Construction 4.1. The action of Γ_{osc} is given by:

$$\rho_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W \to \sum_{i=0}^{\infty} \mathcal{R}_{00}\mathcal{R}_{01}^i(\rho_S) \otimes |\epsilon_i\rangle\langle\epsilon_i|_W \,, \tag{4.50}$$

$$\rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W \to \mathcal{R}_{10}(\rho_S) \otimes |\epsilon_{k-1}\rangle\langle\epsilon_{k-1}|_W + \sum_{i=0}^{\infty} \mathcal{R}_{00}\mathcal{R}_{01}^i\mathcal{R}_{11}(\rho_S) \otimes |\epsilon_{k+i}\rangle\langle\epsilon_{k+i}|_W \,, \tag{4.51}$$

for all $k > 0$.

In Sec. 4.5 we show that any map arising from this construction is a valid thermal operation. This becomes clear once we realise that the building block of the map is the wit map Γ_1 that is applied in a consequtive way (see Fig. 4.3). Moreover, the map additionally satisfies the ETI property (4.20) with $\epsilon_{\min} = \epsilon_1$. This implies that the assumptions of Result 4.2 are met and Γ_{osc} satisfies the second law in the form (4.34). In fact, with a small modification we can easily define the maps Γ_{osc} for a harmonic oscillator battery with finite number of energy levels. To simplify presentation we postpone the details of this construction to Sec. 4.5.

Finally, we apply the above construction to a primitive map which uses work stored in the battery to form an arbitrary (energy-incoherent) quantum state out of a thermal state. As a result we obtain a method of recovering deterministic work for transformations acting on a harmonic oscillator battery with a spectrum bounded from below. In this sense we can think about the harmonic oscillator battery as owning both benefits of the previous two battery models: it allows for studying deterministic work which is an important concept in the resource-theoretic approach (property of the wit) while satisfying the second law of thermodynamics (property of the ideal weight).

Deterministic work for harmonic oscillator batteries

Result 4.3. Let Γ_{wit} be a thermal operation acting on system S and a two-level battery with Hamiltonian $H_W = \delta\, |1\rangle\langle 1|_W$ with $\delta \geq 0$ and performing the

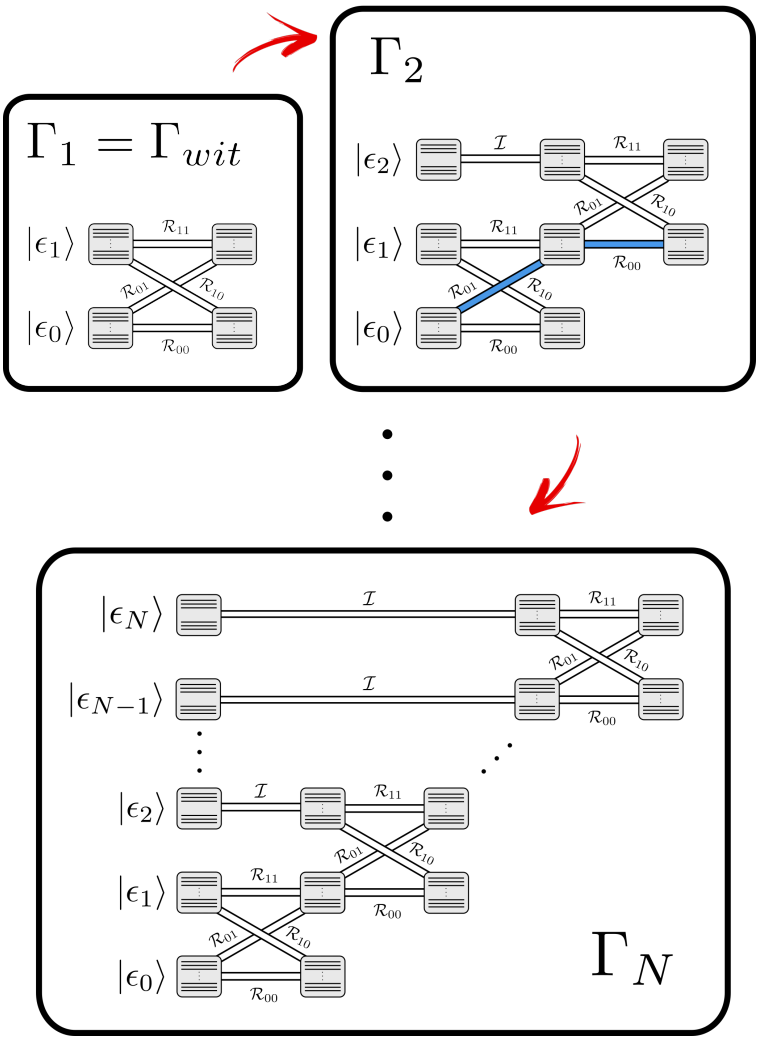

FIGURE 4.3: Graphical explanation of the construction extending thermal operation Γ_{wit} defined on a qubit battery W_{wit} into a map Γ_N acting on an $(N+1)$-level harmonic oscillator. When the battery has an infinite spectrum the construction yields a map $\Gamma_{osc}[\cdot] := \lim_{N\to\infty} \Gamma_N[\cdot]$. In Sec. 4.5 we show that for every $N \in \mathbb{N}$ channel Γ_N is a thermal operation and in the limit $N \to \infty$ becomes effectively translationally invariant with threshold energy $\epsilon_{\min} = \delta$. Blue color corresponds to an exemplary battery subchannel for a 3–level battery, that is $\mathcal{R}^{(2)}_{01}[\cdot] = \mathcal{R}_{00}\mathcal{R}_{01}[\cdot]$. Other battery sub-channels can be determined in an analogous way.

transformation:

$$\Gamma_{wit}\left[\rho_S \otimes |1\rangle\langle 1|_W\right] = \sigma_S \otimes |0\rangle\langle 0|_W\,, \tag{4.52}$$

for some energy-incoherent states ρ_S and σ_S. Then there is a thermal operation Γ_{osc} acting on S and an infinite harmonic oscillator battery such that

$$\Gamma_{osc}\left[\rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W\right] = \sigma_S \otimes |\epsilon_k - \delta\rangle\langle\epsilon_k - \delta|_W\,, \tag{4.53}$$

for all $k > 0$. The average work associated with this transformation is given by:

$$\langle w \rangle = w_F(\rho), \tag{4.54}$$

where $w_F(\rho)$ is the work of formation defined in Eq. (2.97).

Proof. Notice that (4.52) corresponds to a set of battery subchannels with $\mathcal{R}_{11}[\rho_S] = 0$ and $\mathcal{R}_{10}[\rho_S] = \sigma_S$. Hence using Construction 4.1 leads to a map Γ_{osc} which performs (4.53) for all $k > 0$. □

If the map (4.52) is not exact but holds up to some ϵ error on the wit then deterministic work will have an additional contribution surpassing the free energy change, as specified by the correction terms in Result 4.2. However, this contribution can be suppressed by increasing the distance of the initial state of the battery to the vacuum regime. Thus we can start in an arbitrary (e.g. mixed) state of the battery and due to Result 4.2 the approximate version of the second law of thermodynamics will still be satisfied. In this way the notion of deterministic work can be recovered when the energy spectrum of the battery is bounded from below.

Before we finish this section let us briefly note that the construction we present here allows one to recover proper thermodynamic work (i.e. satisfying the second law of thermodynamics) for all primitive maps on the wit. In particular, this means that also the thermalizing map, when extended to a harmonic oscillator battery, no longer leads to violations of the second law, as long as the battery is initialized sufficiently high above the vacuum state.

4.3.4 Landauer erasure and measures of work fluctuations

In this section we examine the main differences between the ideal weight and the harmonic oscillator battery in the low-energy regime, that is when population of the vacuum state cannot be ignored. Naturally we expect that this would lead to certain limitations which do not occur for ideal weight battery. A question arises: what are the consequences of these limitations, and most importantly, what are their implications for general thermodynamic protocols? Here we provide a partial answer to this problem. Our main goal is to explore how the presence of the ground state is reflected in the fluctuations of thermodynamic work.

In what follows we focus our attention on the paradigmatic process of Landauer erasure and implement it using two different battery models: the ideal weight and the harmonic oscillator. We will compare the minimal size of fluctuations that are necessary for these two battery models to perform Landauer erasure with certain fidelity.

Before going into the details, a remark has to be made about possible measures of fluctuations. The very notion of a process being deterministic is always related to the fact that, for this process, a selected measure of fluctuations of work distribution takes zero value. For the sake of comparison with earlier results for translationally invariant batteries (Richens and Masanes, 2016), we start with a measure of fluctuations which is based on the distance to the average work:

$$Z_1[w] = \max_{w:p(w)\neq 0} |w - \langle w \rangle| \tag{4.55}$$

Thanks to Example 2 we see that, when using an ideal weight battery, it is not possible to obtain a non-trivial transformation via a deterministic process with respect to this measure. However, in Section 4.3.4 we show that this can be achieved when the same transformation is performed using a harmonic oscillator battery with unoccupied vacuum state. Later in Section 4.3.4 we compare the differences between the two battery models when fluctuations are quantified by the variance of the work distribution:

$$Z_2[p(w)] = \int \mathrm{d}w\, p(w)(w - \langle w \rangle)^2. \tag{4.56}$$

Finally in Section 4.3.4 we investigate a general quantifier of fluctuations:

$$Z[p(w)] = \int \mathrm{d}w\, p(w) f(w - \langle w \rangle), \tag{4.57}$$

where $f(x)$ is an arbitrary real function satisfying $f(0) = 0$.

Landauer erasure: fluctuations measured via $Z_1[p(w)]$.

Consider a qubit S with a constant Hamiltonian $H_S = \widetilde{H}_S = 0$ prepared in the Gibbs state $\tau_S = \frac{1}{2}\mathbb{1}_S$. Suppose that we also have an imperfect machine which is able to erase qubits with some small failure probability ε and which can tolerate fluctuations of work up to a certain value c[1]. We therefore demand that a bound on the distance to the average energy is satisfied for all registered values of w:

$$|w - \langle w \rangle| \leq c. \tag{4.58}$$

Our goal is to check if putting constraints on the allowed fluctuations of work of the form (4.58) can limit the ultimate precision of erasure that can be achieved in the process. Let us assume that our machine maps the qubit with probability $1 - \varepsilon$ to the state $|0\rangle_S$ and with probability ε fails and outputs the orthogonal state $|1\rangle_S$. The action of the machine on S can be described by the effective transformation:

$$\Gamma_S[\tau_S] = \rho_S(\varepsilon), \tag{4.59}$$

where $\rho_S(\varepsilon) = (1 - \varepsilon)\,|0\rangle\langle 0|_S + \varepsilon\,|1\rangle\langle 1|_S$ and $\Gamma_S = \mathrm{tr}_W\, \mathcal{T}_{SW}$ is a thermal operation reduced to system S. We will implement this effective transformation using two different battery models: the ideal weight and the harmonic oscillator. For the ideal weight we will look for an optimal transformation whereas for the harmonic oscillator we will apply Construction 4.1 to extend the Landauer erasure map on the wit to the harmonic oscillator battery. Our goal here is to find the minimal ε which can be achieved when work fluctuations are bounded by c according to (4.58).

Ideal weight. Let $\mathcal{T}_{SW} = \Gamma_{weight}$ be the thermal operation that performs (4.59) using the ideal weight battery. Since any such operation commutes with the shift operator on the weight (4.8) the transformation will be independent of the initial state of the battery. Hence, without loss of generality, we can choose the initial battery state to be $\rho_W = |0\rangle\langle 0|_W$. The necessary and sufficient conditions for the existence of a thermal operation realizing (4.59) in this case are given by (4.9). In our particular example

[1]Note that we use ε to denote the transformation error, whereas the energies of the battery system are labelled by ϵ.

(fully-degenerate Hamiltonian) they reduce to:

$$\forall\, s' \qquad \sum_{s,w} p(s', w|s) e^{\beta w} = 1, \tag{4.60}$$

Given a process Γ_{weight} described by $p(s', w|s)$ and satisfying (4.60), the work generated in a particular state transition $|s\rangle_S \to |s'\rangle_S$ has probability distribution given by:

$$p(w|s, s') = \frac{p(s', w|s)}{p(s'|s)}, \tag{4.61}$$

where $p(s'|s) = \sum_w p(s', w|s)$ is the probability of a given transition on the system. In general work as described by (4.61) fluctuates around the mean value $\langle w \rangle$. The authors of (Richens and Masanes, 2016) (Result 2, Supplementary notes) showed that the process which minimizes work fluctuations (4.58) without increasing its average value $\langle w \rangle$ satisfies:

$$p(w|s, s') = \delta(w - w_{ss'}). \tag{4.62}$$

In other words, the optimal process which minimizes work fluctuations has a fixed value of work $w_{ss'}$ for any given state transition $|s\rangle_S \to |s'\rangle_S$.

Sometimes in the literature such processes are called "deterministic" to reflect the fact that for each transition there is a well-defined work cost. In this work we will use the term "deterministic" only with respect to the distribution $p(w)$ and only when a specific measure of fluctuations becomes zero. For the measures of fluctuation we will consider here this means that $p(w)$ becomes a point distribution. For simplicity of presentation we will assume that $w_{00} = w_{10} = w_0$ and $w_{01} = w_{11} = w_1$, i.e. the random variable w takes the value w_0 when we erase the state and w_1 when we fail and output an orthogonal state. Applying transformation Γ_{weight} to the initial state of the qubit and the ideal weight battery leads to the joint state:

$$\Gamma_{weight}\Big[\tau_S \otimes |0\rangle\langle 0|_W\Big] = \sum_{s,s',w} p(s', w|s)\, |s'\rangle\langle s'|_S \otimes |w\rangle\langle w|_W \tag{4.63}$$

$$= \sum_{s,s'} p(s'|s)\, |s'\rangle\langle s'|_S \otimes |w_{s'}\rangle\langle w_{s'}|_W\,, \tag{4.64}$$

where the probabilities $p(s'|s)$ are chosen such that the transformation reduced to S correctly reproduces (4.59). The action of Γ_{weight} is summarized in Tab. 4.1.

$\|\mathbf{s}\rangle \to \|\mathbf{s'}\rangle$	$\mathbf{p(s'\|s)}$	**work**
$\|0\rangle \to \|0\rangle$	$1-\varepsilon$	w_0
$\|1\rangle \to \|0\rangle$	$1-\varepsilon$	w_0
$\|0\rangle \to \|1\rangle$	ε	w_1
$\|1\rangle \to \|1\rangle$	ε	w_1

Table 4.1: The action of map Γ_{weight} on system S with the associated work costs.

Notice that the conditions (4.60) imply that the shifts $\{w_{s'}\}$ must necessarily satisfy:

$$(s'=0) \qquad e^{\beta w_0} = \frac{1}{2(1-\varepsilon)}, \tag{4.65}$$

$$(s'=1) \qquad e^{\beta w_1} = \frac{1}{2\varepsilon}. \tag{4.66}$$

Using these we find that the random variable w can take on one of the following values:

$$w_{s'} = \begin{cases} -kT\log 2 - kT\log(1-\varepsilon) & s'=0, \\ \\ -kT\log 2 - kT\log\varepsilon & s'=1. \end{cases} \tag{4.67}$$

The average work during this process can be calculated as:

$$\begin{aligned} \langle w\rangle_{\text{weight}} &= \sum_w p(w)\, w \\ &= \sum_{s,s'} p(s)\, p(s'|s) w_{ss'} \\ &= (1-\varepsilon)w_0 + \varepsilon\, w_1 \\ &= -kT\log 2 + kTh(\varepsilon), \end{aligned} \tag{4.68}$$

where $h(x) := -(1-x)\log(1-x) - x\log x$ is the binary entropy function. In the case when fluctuations of work are bounded by a constant c as in (4.58) we have:

$$\begin{aligned} c &\geq \max_{w:\, p(w)>0} |w - \langle w\rangle_{\text{weight}}| \\ &= \max_{s'} |w_{s'} - \langle w\rangle_{\text{weight}}| \\ &= kT\,|h(\varepsilon) + \log\varepsilon| \\ &\geq -kT(\log\varepsilon + \log 2). \end{aligned} \tag{4.69}$$

This limits the range of ε which can be achieved when the fluctuations are constrained by c, i.e.:

$$\varepsilon \geq \frac{1}{2} e^{-c/kT}. \tag{4.70}$$

If we now take the error $\varepsilon \to 0$ the average work $\langle w\rangle$ will approach a finite value of $\langle w\rangle = -kT\log 2$. However, this can be accomplished only when access to high energies is provided, so that fluctuations of work must be unconstrained, and $c \to \infty$.

Harmonic oscillator battery. Let us consider again the process from (4.59), but now implemented using a harmonic oscillator battery. In order to construct the desired thermal operation we start with a primary process acting on the wit battery with energy separation δ, i.e.:

$$\Gamma_{wit}\left[\tau_S \otimes |1\rangle\langle 1|_W\right] = \rho_S(\varepsilon) \otimes |0\rangle\langle 0|_W\,. \tag{4.71}$$

Using standard methods (e.g. thermomajorization curves) we can determine the minimal value of δ for which (4.71) is a valid thermal operation, i.e.

$$\delta = -k_B T D_{\max}(\rho(\epsilon)||\tau) = kT\log 2 + kT\log(1-\varepsilon). \tag{4.72}$$

Recall that the action of any thermal operation Γ_{wit} can be written as:

$$\begin{aligned}\Gamma_{wit}\Big[\rho_S \otimes \rho_W\Big] =& \Big(\mathcal{R}_{00}\left[\rho_S\right] + \mathcal{R}_{10}\left[\rho_S\right]\Big) \otimes |0\rangle\langle 0|_W + \\ & \Big(\mathcal{R}_{01}\left[\rho_S\right] + \mathcal{R}_{11}\left[\rho_S\right]\Big) \otimes |1\rangle\langle 1|_W .\end{aligned} \tag{4.73}$$

For diagonal input states we can encode the action of subchannels $\{\mathcal{R}_{kk'}\}$ using a set of substochastic matrices $\{R_{kk'}\}$ acting on the diagonals of respective states. Let us denote the vector of initial probabilities of system S with $\mathbf{x} = \mathrm{diag}(\rho_S)$. A simple analysis shows that the action of Γ_{wit} on diagonal states can be expressed using the set of matrices:

$$R_{00} = \begin{bmatrix} 0 & 0 \\ \frac{1-2\varepsilon}{2(1-\varepsilon)} & \frac{1-2\varepsilon}{2(1-\varepsilon)} \end{bmatrix}, \quad R_{10} = \begin{bmatrix} 1-\varepsilon & 1-\varepsilon \\ \varepsilon & \varepsilon \end{bmatrix}, \tag{4.74}$$

$$R_{01} = \begin{bmatrix} \frac{1}{2(1-\varepsilon)} & 0 \\ 0 & \frac{1}{2(1-\varepsilon)} \end{bmatrix}, \quad R_{11} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.75}$$

In the above we chose R_{00} and R_{01} such that Γ_{wit} preserves the Gibbs-state and hence is a valid thermal operation. Using Construction 4.1 we can now extend Γ_{wit} to a thermal operation Γ_{osc} acting on a harmonic oscillator battery. The matrices $\{\widetilde{R}_{kk'}\}$ describing this new process are given by:

For $k = 0$:

$$\begin{aligned}\widetilde{R}_{kk'} &= R_{00}\, R_{01}^{k'} \\ &= \begin{bmatrix} 0 & 0 \\ \frac{1-2\varepsilon}{[2(1-\varepsilon)]^{k'+1}} & \frac{1-2\varepsilon}{[2(1-\varepsilon)]^{k'+1}} \end{bmatrix}\end{aligned} \tag{4.76}$$

For $k > 0$:

$$\widetilde{R}_{kk'} = \begin{cases} R_{10}, & \text{if } k' = k-1, \\ 0, & \text{otherwise.} \end{cases} \tag{4.77}$$

We will now consider two different cases: an ideal one in which the vacuum state of the battery is not occupied and the more physical one in which we assume a small (but nonzero) population in the vacuum state.

Harmonic oscillator battery: ground state not occupied. The transformation constructed using matrices $\{\widetilde{R}_{kk'}\}$ in Construction 4.1 has the same action for all battery states above the ground state, that is:

$$\Gamma_{osc}\left[\tau_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W\right] = \rho_S(\varepsilon) \otimes |\epsilon_k - \delta\rangle\langle\epsilon_k - \delta|_W , \tag{4.78}$$

This holds for all initial states of the battery above the ground state, i.e. for all

$\epsilon_k > \epsilon_{\min} = 0$. This means that any initial state ρ_W of the battery (with $\langle\epsilon_0|\rho_W|\epsilon_0\rangle = 0$) will lead to the same transformation. Furthermore, it can be shown (see Sec. 4.5 for the details) that for any primitive map Γ_{wit} and any initial state of the battery above the ground state the average work associated with the extended map Γ_{osc} can always be expressed as:

$$\langle w\rangle_{osc} = \delta \cdot \left[\mathbf{1}^T(\mathbb{1} - R_{01})^{-1}R_{11}\,\mathbf{x} - 1\right] = -\delta,$$

where $\mathbf{1}^T = (1, 1, \ldots, 1)$ is the (horizontal) identity vector and $\mathbb{1}$ is the identity matrix. From (4.78) it is clear that whenever the battery starts above its ground state the shifts $\{w_{s'}\}$ are the same whenever erasure succeeds or fails, i.e. for all s':

$$w_{s'} = -\delta = -kT\log 2 - kT\log(1-\epsilon). \tag{4.79}$$

This also implies that the random variable w does not fluctuate. Indeed we have:

$$\begin{aligned} c &\geq \max_{w:\,p(w)>0} |w - \langle w\rangle_{\text{osc}}| \\ &= |\delta - \langle w\rangle_{\text{osc}}| \\ &= 0. \end{aligned} \tag{4.80}$$

In particular this means that ε does not depend on c and hence there is no fundamental limit on ϵ allowed by the transformation, i.e.:

$$\varepsilon \geq 0. \tag{4.81}$$

In this way a harmonic oscillator battery faithfully reproduces the amount of deterministic work needed to perform erasure for an arbitrary low error rate ε, a feat which the ideal weight could not perform as described in the example from previous section.

Harmonic oscillator battery: ground state occupied. While the above is a valid mathematical construction, in a physical situation the vacuum state will be inevitably occupied, and therefore one should expect the harmonic oscillator battery to show fluctuations of work. To quantify them, we assume that the vaccum state is occupied with probability γ. In this case the situation is different to the ideal ($\gamma = 0$) case. This is because applying the thermal map Γ_{osc} when the battery is in its ground state thermalizes the battery, i.e. returns a state with all levels of the harmonic oscillator battery occupied. However, although the map brings the battery to a full-rank state, the probability of occupying higher energy levels decays exponentially with the energy. Hence the entropy change in the battery is always finite and, in particular, does not diverge in the limit of infinite dimension of the battery (though with probability decaying exponentially fast with the energy). Furthermore, the occupation of the vacuum induces a proportional error on the system. In order to find the action on the system and the total distribution of work $p(w)$ let us note that our construction assures translational invariance for all energy levels above energy $\epsilon_{min} = \epsilon_0$. This means that without loss of generality we can prepare the battery in the initial state:

$$\rho_W = (1-\gamma)\,|\epsilon_1\rangle\langle\epsilon_1|_W + \gamma\,|\epsilon_0\rangle\langle\epsilon_0|_W\,, \tag{4.82}$$

where $\gamma \in [0,1]$. Applying our Landauer erasure map leads to the following state on the system and the battery:

$$\Gamma_{osc}[\tau_S \otimes \rho_W] = (1-\gamma)\Gamma_{osc}[\tau_S \otimes |\epsilon_1\rangle\langle\epsilon_1|_W] + \gamma\Gamma_{osc}[\tau_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W] \tag{4.83}$$

$$= (1-\gamma)\rho_S(\epsilon) \otimes |\epsilon_0\rangle\langle\epsilon_0|_W + \sum_{i=0}^{\infty} \mathcal{R}_{01}^i \mathcal{R}_{00}[\tau_S] \otimes |\epsilon_i\rangle\langle\epsilon_i|_W \,. \tag{4.84}$$

The map Γ_{osc} implements Landauer erasure on the system with the total error

$$\varepsilon_{tot} = \varepsilon(1-\gamma) + \gamma, \tag{4.85}$$

where ε is the parameter inherited from Γ_{wit}, based on which Γ_{osc} is constructed. Therefore, even for $\varepsilon = 0$ perfect erasure is not possible and ultimately depends also on the occupation of the vacuum state.

In this case the protocol can only be implemented if we allow $c \to \infty$, i.e. work w can take all possible values between 0 and ∞ (although this happens with exponentially small probabilities). In this way, even though in the ideal case we could in principle always choose $\varepsilon = 0$ for any value of $c \geq 0$, when the vacuum state is populated we must allow for unbounded fluctuations of work in order to carry out the erasure perfectly.

To summarize, we see that using a battery with broken translational symmetry can lead to an arbitrarily good precision of Landauer erasure just as in the case of the ideal weight battery, but with work fluctuations reduced to zero. However, once we consider a more physical situation in which the battery's ground state is occupied, both models behave essentially in the same way. It is an interesting open question whether the different behavior certified for $\gamma = 0$ can lead to some physical advantages over the ideal weight (e.g. resulting from lifting the translational invariance constraint), or it is just a mathematical idealization, similar as in the case of wit. This reasoning comes from the fact that the set of operations generated when using a bounded battery is strictly larger that that of the ideal weight; Whether these additional manipulation abilities can ever lead to any useful advantages, we as an open problem for future.

Landauer erasure: fluctuations measured via $Z_2[p(w)]$.

We saw above that for a non-zero occupation of the vacuum state, Γ_{osc} returns a state with all levels of the harmonic oscillator battery occupied (though with probability decaying exponentially with the energy), while the average energy remains finite. Therefore, the previously used measure of fluctuations, $F_1[w]$, takes infinite values. In order to give a quantitative description of the functioning of physical batteries, we therefore switch to considerations of statistical moments of energy changes, as they take into account not only values of registered work, but also probabilities of obtaining it.

Immediately we can draw some conclusions about the functioning of physical batteries (i.e. with ground states), even without invoking their translational invariance properties in some regime above the cut-off energy. Whenever the vacuum state (or the lowest energy eigenstate on which the process acts non-trivially) is occupied, stochasticity of the map implies that the work distribution has positive contributions resulting from the vacuum being populated (as population on this level cannot be mapped to lower levels). On the other hand, Landauer erasure is a type of transformation which requires work (hence in our notation it is associated with a negative

average work). Therefore higher energy levels of the battery have to be initially occupied, in order to assure the dominant negative contribution to the average. As a consequence, the variance can only be reduced to zero when the vacuum is not populated. Moreover, as we show below in Result 4.4, for a fixed occupation of the vacuum, the variance of Landauer erasure cannot even be brought down arbitrarily close to zero. This can be treated as another argument suggesting that the ideal weight may not be able to encapsulate important physical effects in the regimes where the occupation of the vacuum cannot be ignored. In other words, it allows for processes which cannot be realised in practise, i.e. when one only has access to physical batteries.

Before delving into quantitative analysis let us start with a simple fact. The below theorem provides a bound on the fluctuations of work which is valid for any battery with the ground state and any stochastic transformation on the system and the battery (i.e. it is not necessarily translationally-invariant or Gibbs-preserving).

No-go for work fluctuations close to the ground state

Result 4.4. Let ρ_W be the state of the battery with the ground state occupation γ, i.e.

$$\gamma = \langle \epsilon_0 | \rho_W | \epsilon_0 \rangle. \tag{4.86}$$

The variance of the work distribution $p(w)$ arising from any thermodynamic protocol with $\langle w \rangle \leq 0$ is bounded by:

$$\mathrm{Var}[w] \geq \gamma \langle w \rangle^2. \tag{4.87}$$

Proof. The variance $\mathrm{Var}[w]$ of a random variable w distributed according to $p(w)$ is given by:

$$\mathrm{Var}[w] = \sum_w p(w)(w - \langle w \rangle), \tag{4.88}$$

Let s be the probability of performing work $w = 0$ when the battery is in the ground state. Using the assumption $\langle w \rangle \leq 0$ we can write:

$$\begin{aligned}
\mathrm{Var}[w] &\geq p(w=0)(0 - \langle w \rangle)^2 + \sum_{w>0} p(w)(w - \langle w \rangle)^2 \\
&\geq p(w=0)\langle w \rangle^2 + \langle w \rangle^2 \sum_{w>0} p(w) \\
&\geq \gamma s \langle w \rangle^2 + \gamma(1-s)\langle w \rangle^2 \qquad (4.89) \\
&= \gamma \langle w \rangle^2. \qquad (4.90)
\end{aligned}$$

This proves the claim. □

Harmonic oscillator battery. In this section the average energetic cost of performing the transformation is measured in terms of the average work $\langle w \rangle_{osc}(\epsilon, \gamma)$, while the

variance of the work distribution:

$$\begin{aligned} Z_2[p(w)] &= \mathrm{Var}_{osc}(\epsilon,\gamma) \\ &= \int \mathrm{d}w\, p(w)(w-\langle w\rangle)^2 \end{aligned} \tag{4.91}$$

will be used as a measure of fluctuations. For the map from Construction 4.1 based on (4.74) and (4.75), we can calculate the corresponding measures:

$$\langle w\rangle_{osc}(\varepsilon,\gamma)/k_BT = -\delta\left(1-\frac{2\gamma(1-\varepsilon)}{1-2\varepsilon}\right) \tag{4.92}$$

and

$$\frac{\mathrm{Var}_{osc}(\varepsilon,\gamma)}{(k_BT)^2} = \gamma\delta^2\,\frac{2(1-\varepsilon)(-2\gamma(1-\varepsilon)-2\varepsilon+3)}{(1-2\varepsilon)^2}. \tag{4.93}$$

where $\delta := \log 2(1-\varepsilon)$ is defined as before. It can be easily shown that, for a given value of $\varepsilon_{\mathrm{tot}} = \varepsilon(1-\gamma)+\gamma$ both $\langle w\rangle_{osc}(\varepsilon,\gamma)$ and $\mathrm{Var}_{osc}(\varepsilon,\gamma)$ are maximized for $\varepsilon = 0$, $\gamma = \epsilon_{tot}$ (see also Fig. 4.4). As $\langle w\rangle_{osc}(\epsilon,\gamma)$ is always non-positive, its maximization corresponds to minimization of work spent on the erasure. However, the less work we spend on the erasure, the higher are its fluctuations, as measured by $F_2[p(w)]$.

Ideal weight battery. These fluctuations on the harmonic oscillator battery should be compared with fluctuations of work registered on the ideal weight battery for an analogous process, i.e. a process implementing Landauer erasure with the same failure probability $\varepsilon_{\mathrm{tot}}$. As a promising candidate for minimizing fluctuations in this regime, we take the process (4.63) with work values specified in Table 4.1 (which minimizes the measure $F_1[w]$). Therefore, we denote the corresponding values of the average work and variance with subscript "1", i.e.:

$$\langle w\rangle_1(\varepsilon_{tot})/k_BT = -\log 2 + h(\varepsilon_{tot}), \tag{4.94}$$

and

$$\frac{\mathrm{Var}_1(\varepsilon_{tot})}{(k_BT)^2} = \varepsilon_{tot}\big(-\log(\varepsilon_{tot})-h(\varepsilon_{tot})\big)^2+(1-\varepsilon_{tot})\big(-\log(1-\varepsilon_{tot})-h(\varepsilon_{tot})\big)^2, \tag{4.95}$$

where $h(x) = -(1-x)\log(1-x) - x\log x$. In the regime of small errors ε_{tot}, fluctuations on the harmonic oscillator battery are smaller compared to the ones on the weight (see Fig. 4.4). As indicated above, this range grows with decreasing occupation of the vacuum state, in agreement with the description of the no-fluctuation case of the idealised harmonic oscillator battery presented in the previous section. We also observe that working with an oscillator battery leads to higher expenditures of work, as $\langle w\rangle_{osc}(0,\varepsilon_{\mathrm{tot}}) \leq \langle w\rangle_1(\epsilon_{\mathrm{tot}})$, with equality for $\epsilon_{\mathrm{tot}} = 0$ and $\epsilon_{\mathrm{tot}} = 1$.

Note however that it is not clear that the process (4.63) minimizes the measure $F_2[w]$ for the ideal weight. In fact, the following probabilities define a process on the ideal weight which in the limit $\lambda \to 0$ is deterministic with respect to $Z_2[w]$:

$$p(w|s'=0) = \delta(w-w_0), \tag{4.96}$$

$$p(w|s'=1) = (1-\lambda)\delta(w-w_0)+\lambda\delta(w-w_1), \tag{4.97}$$

where $\lambda \in [0,1]$. Since in our case we have $p(s'=0) = 1-\varepsilon$ and $p(s'=1) = \varepsilon$ we can write the total distribution of work $p(w)$ as:

$$p(w) = (1-\lambda\varepsilon)\delta(w-w_0) + \lambda\varepsilon\delta(w-w_1). \tag{4.98}$$

Such a thermal operation exists if the Gibbs-stochasticity conditions are satisfied. This means that we must have:

$$e^{\beta w_0} = \frac{1}{2(1-\varepsilon)}, \tag{4.99}$$

$$(1-\lambda)e^{\beta w_0} + \lambda e^{\beta w_1} = \frac{1}{2\varepsilon}. \tag{4.100}$$

Notice now that for an arbitrarily small ε and an arbitrarily small λ the work value w_1 can be always chosen large enough so that these two constraints are satisfied. In particular for any $\lambda, \varepsilon > 0$ the choice

$$w_1 = \frac{1}{\beta}\log\left[\frac{1}{2\lambda\epsilon} - \frac{1-\lambda}{2(1-\varepsilon)}\right] \tag{4.101}$$

leads to a legitimate thermal operation. Direct calculations in the limit $\lambda \to 0$ show:

$$\langle w\rangle_2(\varepsilon) = w_0 = -k_B T[\log 2 + \log(1-\varepsilon)], \tag{4.102}$$

$$\mathrm{Var}_2(\varepsilon) = 0. \tag{4.103}$$

In this way by allowing for larger amounts of work which occur with respectively smaller probabilities, we can recover deterministic work (with respect to $Z_2[w]$). In fact, in the limit $\lambda \to 0$, process (4.96) leads to the same average and variance of work as the process from Construction 4.1, with no vacuum state occupied. Note that, due to the convexity of the exponential function in (4.99) we have $\langle w\rangle_2 \leq \langle w\rangle_1$, so here also a reduction of fluctuations is obtained at the cost of performing additional work.

Landauer erasure: generalized model of fluctuations $Z[w]$.

Let us now consider a simple generalization of the quantifiers we studied in the previous sections, i.e. we will look at the following quantity:

$$Z[w] := \int \mathrm{d}w\, p(w) f(w - \langle w\rangle). \tag{4.104}$$

In the above $f(x)$ can be any function which satisfies $f(0) = 0$. Intuitively, the above describes the (weighted) average work fluctuations that one is willing to tolerate in their protocol. In this way bounding (4.104) with some constant value c provides a more general description of our willingness to tolerate fluctuations of work depending on their size. To gain some more intuition consider the following choice of the cost function $f(x)$:

$$f(x) = \begin{cases} 0 & \text{if} \quad |x| \leq c, \\ \infty & \text{else.} \end{cases} \tag{4.105}$$

This corresponds to the situation of *c-bounded work* which we considered in Section 4.3.4, when setting $c = 0$ leads to a deterministic work extraction, with respect to the

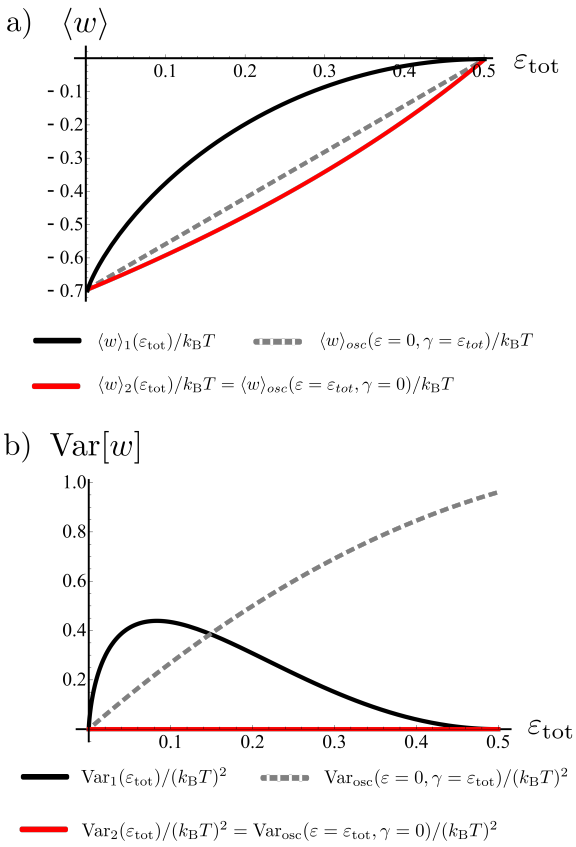

FIGURE 4.4: Comparison between different processes performing Landauer erasure on ideal weight and physical battery, with error ϵ_{tot} on a system, stated in terms of (a) average work $\langle w \rangle$ and (b) variance of work distribution $\mathrm{Var}[w]$. On a physical battery, a process based on Construction 4.1 with (4.74) and (4.75), whose error ε_{tot} stems entirely from vacuum occupation γ, is characterized by minimal work cost and maximal fluctuations (for this particular battery model) indicated by the grey dashed lines. For the nonphysical case of the vacuum being not occupied, fluctuations can be reduced to zero, at the expense of increasing work cost (red lines). On the other hand, the same statistics can be achieved on the ideal weight for the process (4.96) in the limit $\lambda \to 0$. For $\lambda = 1$, solid black lines depict minimal possible work expense and corresponding variance of work distribution on the ideal weight.

measure $F_1[p(w)]$. On the other hand, choosing $f(x) = x^2$ leads to the variance of the work distribution (see Section 4.3.4).

We will now again consider the example of approximate Landauer erasure, but now expressing constraints in terms of the function F defined in (4.104). Our goal here is similar as before: we would like to check if putting constraints on the allowed fluctuations of work can lead to any bounds on the ultimate precision of the Landauer erasure process. We will again look at the problem first from the perspective of using the ideal weight as the battery and second using the harmonic oscillator with the ground state.

The ideal weight battery. Let us consider the approximate Landauer erasure process performed on the weight described by the process (4.96). As mentioned before, by allowing larger amounts of work W_1 occurring with respectively smaller probabilities $\varepsilon\lambda$, we can recover deterministic work (with respect to $Z_2[p(w)]$). This analysis, however, does not take into account the size of fluctuations. We should expect that small fluctuations around the average $\langle w \rangle$ should not be as adversarial as the large ones, and similarly large fluctuations should be more costly to tolerate. Hence we instead put a constraint on the average cost of fluctuations:

$$Z[w] \leq c. \tag{4.106}$$

To focus our attention let us choose an exponential cost function $f(x) = e^{|x|} - 1$, i.e. we are going to treat small fluctuations as almost free and larger ones as significantly more costly. Any function with exponential or better decay in x would lead to the

same qualitative conclusions. In this case our constraint (4.106) becomes:

$$(1-\lambda\varepsilon)e^{\beta(w_0-\langle w\rangle)}+\lambda\varepsilon e^{\beta(w_1-\langle w\rangle)}\leq c. \tag{4.107}$$

Since both $\lambda, \varepsilon > 0$ this also implies:

$$\lambda\varepsilon e^{\beta w_1}\leq c\, e^{\beta\langle w\rangle}\leq c, \tag{4.108}$$

since $\langle w\rangle \leq 0$ which follows by applying Jensen's inequality to (4.99). Let us now look at the Gibbs-stochasticity conditions (4.99). By solving for w_0 we can express the second condition as:

$$\lambda\varepsilon e^{\beta w_1}=\frac{1-2\varepsilon+\varepsilon\lambda}{2(1-\varepsilon)}\geq\frac{1-2\varepsilon}{2(1-\varepsilon)}. \tag{4.109}$$

Combining (4.108) with (4.109) leads to the following bound on ε:

$$\varepsilon\geq\frac{1}{2}-\frac{c}{2(1-c)}. \tag{4.110}$$

Hence we have a non-trivial bound for all $c \in [0, \frac{1}{2})$. Importantly, this bound holds for all initial states of the weight battery and states that its impossible to obtain lower ε for any possible thermal operation.

Harmonic oscillator battery. Let us now consider the same transformation but performed using the harmonic oscillator battery. In the ideal case when the vacuum state is not occupied the whole analysis can be performed just as in the previous section. However, the situation changes if we start with a non-zero population of the vacuum state. In that case Γ_{osc} returns a state with all levels of the harmonic oscillator battery occupied (though with probability decaying exponentially fast with the energy). Furthermore, occupation of the vacuum induces a proportional error on the system. In order to find the action on the system and the total distribution of work $p(w)$ let us note that our construction assures translational invariance for all energy levels above $k = 0$. This means that without loss of generality we can prepare the battery in the following initial state:

$$\rho_W=(1-\gamma)\,|\epsilon_1\rangle\langle\epsilon_1|_W+\gamma\,|\epsilon_0\rangle\langle\epsilon_0|_W\,, \tag{4.111}$$

where $\gamma \in [0, 1/2]$. Applying our Landauer erasure map leads to the following final state on the system and the battery:

$$\begin{aligned}\Gamma_{osc}[\tau_S\otimes\rho_W]&=(1-\gamma)\Gamma_{osc}[\tau_S\otimes|\epsilon_1\rangle\langle\epsilon_1|_W]+\gamma\,\Gamma_{osc}[\tau_S\otimes|\epsilon_0\rangle\langle\epsilon_0|_W]\\&=(1-\gamma)\rho_S(\varepsilon)_S\otimes|\epsilon_0\rangle\langle\epsilon_0|_W+\gamma\sum_{i=0}^{\infty}\mathcal{R}_{01}^i\mathcal{R}_{00}[\tau_S]\otimes|\epsilon_i\rangle\langle\epsilon_i|_W\,.\end{aligned} \tag{4.112}$$

Notice that we also have:

$$\mathcal{R}_{01}^i\mathcal{R}_{00}[\tau_S]=\frac{1-2\varepsilon}{2^{i+1}(1-\varepsilon)^{i+1}}\,|1\rangle\langle1|_S\,. \tag{4.113}$$

To simplify notation let us denote $\alpha_i := \mathrm{tr}\,\mathcal{R}_{01}^i\mathcal{R}_{00}[\tau_S]$. The final state on the system becomes:

$$\rho_S'=\mathrm{tr}_W\,\Gamma_{osc}[\tau_S\otimes\rho_W]=(1-\varepsilon_{tot})\,|0\rangle\langle0|_S+\varepsilon_{tot}\,|1\rangle\langle1|_S\,,$$

where $\varepsilon_{tot} = \varepsilon + \gamma(1-\varepsilon)$ is the effective total error on the system. Hence it is now not only the channel imperfection but also occupation of the vacuum which bounds the precision of the Landauer erasure. The work distribution can be now computed as:

$$p(w) = \big(1-\gamma\big)\ \delta(w+\delta) + \gamma \sum_{i=0}^{\infty} \alpha_i \delta(w - i\delta). \tag{4.114}$$

The average work $\langle w \rangle_{osc}$ then becomes:

$$\langle w \rangle_{osc} = -\delta \left(1 - \frac{2\gamma(1-\varepsilon)}{1-2\varepsilon}\right) \tag{4.115}$$

In this case our average cost function (4.106) with exponential cost $f(x) = e^{|x|}$ becomes:

$$Z[p(w)] = (e^{2\beta\delta\gamma} - 1) - \frac{\gamma}{2} e^{-2\beta\delta\gamma} \left(1 - \frac{e^{\beta\delta}}{2 - e^{\beta\delta}}\right) \tag{4.116}$$

Notice that if we now choose the smallest possible $\delta = k_B T \log[2(1-\varepsilon)]$ we obtain an expression diverging to infinity with $\varepsilon \to 0$. This also holds if we choose a bigger work value, i.e. $\delta = k_B T \log 2$.

We therefore see that, for a broad class of fluctuation measures $Z[p(w)]$ (4.104) with $f(x)$ increasing at least exponentially, physical batteries with occupied vacuum lead to infinite deviations from deterministic work. This is in line with our observation that occupation of the ground state leads to fluctuations of work which can never be brought down arbitrarily close to zero. In this respect the two battery models (the ideal weight and the harmonic oscillator battery) lead to different thermodynamic predictions in the low energy regime, i.e. when the occupation of the ground state cannot be ignored.

4.4 Discussion and open problems

In classical thermodynamics the existence of a perfect work receiver is a standard and very often implicitly accepted assumption. This is also the case for many contemporary results in quantum and stochastic thermodynamics (e.g. fluctuation theorems) where work is described by energy changes on the system. In this work we argued towards breaking this paradigmatic assumption and studying a truly physical (i.e. *explicit* and *bounded*) work storage device, a harmonic oscillator battery.

We have observed that the effects related to the ground state of the work-storage device manifest in the Jarzynski fluctuation theorem. We determined the estimates on fluctuations of energy changes of the battery which replace the standard fluctuation theorem for the case when the battery is bounded from below.

Furthermore, using these estimates we derived corrections to the second law of thermodynamics and showed that they vanish exponentially fast with the distance of the initial state of the battery to the bottom of its spectrum. These corrections allow to determine when contributions to the average energy of the battery can be treated as average work, hence providing a quantitative tool to differentiate when the battery operates accurately and when it acts as an entropy sink. In other words, the bounds we find lead to a new type of dissipated work that is not associated with the particular process, but to the non-ideal character of the work reservoir.

Notably, mathematical forms of our bounds on work fluctuactions and average work remain valid even when coherence is present in the system, though terms which in the incoherent case refer to free energy loose this association. As a consequence, the bounds can no longer be used as tools for the assesment of physicality of work. It would be interesting to see if there is a positive interplay between quantum effects resulting from the existance of a ground state of the battery and coherence between its energy states, when it comes to amount of extractable average work.

The bounds presented here are a step forward towards understanding to what extent the energy changes on physical batteries can be treated as a genuine thermodynamic work. We emphasize that the deviations from the second law which we described concern the average work rather than the work itself, hence they are of completely different character than deviations reported by standard fluctuation theorems. In the regime of translational symmetry these corrections are negligible, hence the shifts on the battery can be treated as thermodynamic work. This naturally leads to an important open question: what is work if we are not using the battery in the translationally invariant regime? Another interesting investigation in this direction would be to asses the corrections to the second law in terms of the entropy decrease of the system and the bath, rather than in terms of the average work, as we have done in this chapter. Furthermore, one can look for saturable upper bounds on the l.h.s of (4.26), so that in the limit $k \to \infty$ the standard Jarzynski identity can be recovered.

In this work we also laid the first step in answering the fundamental question: how thermodynamic laws and fluctuation relations modify if we take into account the limitations imposed by the presence of the lowest energy state of the battery? This question is relevant both in classical and quantum thermodynamics and is not exclusively related to the resource-theoretic framework which we use here. In this sense we believe that the results we present here open the door for deriving more universal corrections that would apply to arbitrary models of work reservoirs, as well as other frameworks for thermodynamics.

We also showed that the harmonic oscillator battery model correctly reproduces the amount of single-shot work of formation, originally introduced using qubit as the battery system in (Horodecki and Oppenheim, 2013). Not only this allows to study the two seemingly contradictory notions of *single-shot* and *fluctuating* work using a single battery model, but also, due to Result 2, allows to determine when we can treat deterministic work as a genuine thermodynamic work satisfying the second law of thermodynamics. A natural open question is whether this result can be extended to the case of work of distillation, i.e. when work is distilled from a quantum state and stored in the battery.

Furthermore, we demonstrated that in the regime where the population of the battery ground state cannot be neglected, the ideal weight battery does not provide a proper description of physical processes. In the presented example of Landauer erasure we showed that there is a fundamental lower bound on the minimal variance of work distribution that any physical model of a battery has to satisfy. This bound is violated for the ideal weight and shows that fluctuations of thermodynamic work in the low-energy regime behave significantly different than in the regime of high energies.

Finally, we showed that physical batteries will suffer from the existence of the ground state. This influences our understanding of the notion of work as well as asks for reformulation of the whole domain of fluctuation relations whenever the battery is used in the "close to the vacuum" regime. However, one may also ask an orthogonal question of whether physical batteries can offer any kind of advantage

over the ideal weight. This must necessary happen in a way that this advantage be "larger" than the violation of the second law as quantified by our correction terms. The arising advantage could not be then attributed only to the second law violation but rather would describe a genuine improvement. For example, we could fix the allowable amount of work dissipated into the battery due to the presence of the ground state (i.e. $w_{diss} = \Delta F - \langle w \rangle$) and run a thermodynamic protocol in the intermediate regime (i.e. many copies but not infinitely large). The work dissipated in the process then would not scale with the size of the system. This can lead to improved manipulation, in a similar way that is demonstrated in the case of large catalysts. Such a behavior is indeed not excluded since in our model, by accepting a weaker notion of translational invariance, we effectively allow for a larger class of operations than in the case of the ideal weight model.

4.5 Proofs

In this section we present all the proofs we have postponed throughout this chapter. In Sec. 4.5.1 we prove Result 4.1, in Sec. 4.5.2 we prove Result 4.2 and in Sec. 4.5.3 we prove Corollary 6.1. Finally, in Sec. 4.5.4 we formally derive the main properties of the map Γ_{osc}.

4.5.1 Proof of Result 4.1

Let us start by defining a random variable:

$$f_s := E_s + kT \log p_S(s) \tag{4.117}$$

occuring with probability $p_S(s)$ and whose average is the free energy of a system in state $\rho_S = \sum_s p_S(s) |s\rangle\langle s|_S$, i.e.

$$\langle f_s \rangle = \sum_s p_S(s) \left[E_s + kT \log p_S(s) \right] = F(\rho_S) \tag{4.118}$$

Now consider the following average:

$$\langle e^{\beta(f_{s'} - f_s + w)} \rangle = \sum_{s,s',w} p(s,s',w)\, e^{\beta(f_{s'} - f_s + w)}. \tag{4.119}$$

We can compute the distribution $p(s,s',w)$ explicitly from the map $\mathcal{T}_{SW}$ and write it using transition probabilities $\{r(s'k'|sk)\}$ in the following way:

$$p(s,s',w) = p_S(s)\, p(s'w|s) = \sum_{k,k'} p_W(k)\, p_S(s)\, r(s'k'|sk)\, \delta_{w,w_{kk'}}, \tag{4.120}$$

where we defined $w_{kk'} := \epsilon_{k'} - \epsilon_k$. We also define an analogous average quantity conditioned on the battery eigenstate $|\epsilon_k\rangle_W$, i.e.:

$$\langle e^{\beta(f_{s'} - f_s + w_{kk'})} \rangle_k := \sum_{s,s',k'} p_S(s)\, r(s'k'|sk) e^{\beta(f_{s'} - f_s + w_{kk'})}. \tag{4.121}$$

This allows us to write (4.119) as:

$$\langle e^{\beta(f_{s'} - f_s + w)} \rangle = \sum_k p_W(k) \langle e^{\beta(f_{s'} - f_s + w_{kk'})} \rangle_k\, \delta_{w,w_{kk'}}$$

We are now ready to prove Result 4.1.

Proof. We start by rewriting:

$$\langle e^{\beta(f_{s'}-f_s+w_{kk'})}\rangle_k = \sum_{s'} p_S(s')\, h(s',k), \tag{4.122}$$

where we labeled $h(s',k) = \sum_{s,k'} r(s'k'|sk) e^{\beta(\widetilde{E}_{s'}-E_s+w_{kk'})}$. It will be also convenient to introduce:

$$g_{s'}(k'|k) := \sum_s e^{\beta(\widetilde{E}_{s'}-E_s)}\, r(s'k'|sk). \tag{4.123}$$

Notice that for each $k \geq k_{\min}$ we can rewrite $h(s',k)$ as:

$$\begin{aligned} h(s',k) &= \sum_{k'=0}^{N-k_{\min}} g_{s'}(k'|k) e^{\beta(\epsilon_{k'}-\epsilon_k)} + \sum_{k'=N-k_{\min}+1}^{N} g_{s'}(k'|k) e^{\beta(\epsilon_{k'}-\epsilon_k)} \\ &= \sum_{k'=0}^{N-k_{\min}} g_{s'}(k'|k) e^{\beta(\epsilon_{k'}-\epsilon_k)} + \sum_{i=1}^{k_{\min}} g_{s'}(N-k+i|k_{\min}) e^{\beta(\epsilon_{N-k+i}-\epsilon_{\min})}, \end{aligned}$$

where in the last equation we changed the summation index to $i = k' - N + k_{\min}$ and used the ETI property (4.21) with $n = k_{\min} - k$. Using again the stochasticity and the Gibbs-preserving conditions we can rewrite the first sum in the above equality as::

$$\begin{aligned} \sum_{k'=0}^{N-k_{\min}} g_{s'}(k'|k) e^{\beta(\epsilon_{k'}-\epsilon_k)} &= \sum_{k'=0}^{N-k_{\min}} g_{s'}(N-k|N-k') e^{\beta(\epsilon_{k'}-\epsilon_k)} \\ &= \sum_{l'=k_{\min}}^{N} g_{s'}(N-k|l') e^{\beta(\epsilon_{N-l'}-\epsilon_k)} \\ &= 1 - \sum_{l=0}^{k_{\min}-1} g_{s'}(N-k|l) e^{\beta(\epsilon_{N-k}-\epsilon_l)} \\ &= 1 - A_{s'}(N-k), \end{aligned} \tag{4.124}$$

where we labelled $A_{s'}(x) := \sum_{l=0}^{k_{\min}-1} g_{s'}(x|l)\, e^{\beta(\epsilon_x-\epsilon_l)}$. Using again (2.86) with $k' = N-k$ we get:

$$A_{s'}(N-k) + \sum_{l=k_{\min}}^{N-1} g_{s'}(N-k|l) e^{\beta(\epsilon_{N-k}-\epsilon_l)} + g_{s'}(N-k|N) e^{\beta(\epsilon_{N-k}-\epsilon_N)} = 1. \tag{4.125}$$

On the other hand for $k' = N-k+1$ we get:

$$A_{s'}(N-k+1) + g_{s'}(N-k+1|k_{\min})\, e^{\beta(\epsilon_{N-k+1}-\epsilon_{\min})} + \tag{4.126}$$

$$\sum_{l=k_{\min}+1}^{N} g_{s'}(N-k+1|l) e^{\beta(\epsilon_{N-k+1}-\epsilon_l)} = 1. \tag{4.127}$$

Notice that due to the ETI property we also have:

$$\sum_{l=k_{\min}+1}^{N} g_{s'}(N-k+1|l)e^{\beta(\epsilon_{N-k+1}-\epsilon_l)} = \sum_{l=k_{\min}}^{N-1} g_{s'}(N-k|l)e^{\beta(\epsilon_{N-k}-\epsilon_l)}, \tag{4.128}$$

Combining (4.125) with (4.126) and the last equation yields:

$$A_{s'}(N-k) = A_{s'}(N-k+1) + g_{s'}(N-k+1|k_{\min})e^{\beta(\epsilon_{N-k+1}-\epsilon_{\min})} \tag{4.129}$$

$$- g_{s'}(N-k|N)e^{\beta(\epsilon_{N-k}-\epsilon_N)}. \tag{4.130}$$

We now repeat this step $k_{\min}$ times to express $A_{s'}(N-k)$ using $A_{s'}(N-k+k_{\min})$, i.e.

$$A_{s'}(N-k) = A_{s'}(N-k+k_{\min}) + \sum_{i=1}^{k_{\min}} g_{s'}(N-k+i|k_{\min})e^{\beta(\epsilon_{N-k+i}-\epsilon_{\min})} \tag{4.131}$$

$$- \sum_{i=0}^{k_{\min}-1} g_{s'}(N-k+i|N)e^{\beta(\epsilon_{N-k+i}-\epsilon_N)}. \tag{4.132}$$

Notice now that the first sum on the RHS of (4.131) appears exactly in (4.124). This allows to write for all $k \geq k_{\min}$

$$h(s',k) = 1 + \sum_{i=0}^{k_{\min}-1} g_{s'}(N-k+i|N)e^{\beta(\epsilon_{N-k+i}-\epsilon_N)} - A_{s'}(N-k+k_{\min}) \tag{4.133}$$

$$= 1 + \sum_{i=0}^{k_{\min}-1} g_{s'}(N-k+i|N)e^{\beta(\epsilon_{N-k+i}-\epsilon_N)} - \sum_{i=0}^{k_{\min}-1} g_{s'}(N-k+k_{\min}|i)e^{\beta(\epsilon_{N-k+k_{\min}}-\epsilon_i)} \tag{4.134}$$

$$= 1 + \sum_{i=0}^{k_{\min}-1} \left[g_{s'}(N-k+i|N)e^{\beta(\epsilon_i-\epsilon_k)} - g_{s'}(N-k+k_{\min}|i)e^{\beta(\epsilon_{N-k+k_{\min}}-\epsilon_i)} \right] \tag{4.135}$$

$$\leq 1 + e^{-\beta\epsilon_k} \cdot \sum_{i=0}^{k_{\min}-1} \left[g_{s'}(N-k+i|N)e^{\beta\epsilon_i} \right] \tag{4.136}$$

$$\leq 1 + e^{-\beta\epsilon_k} e^{\beta(\epsilon_{k_{\min}}-\delta)} g_{s'}(N-(k-k_{\min})|N) \tag{4.137}$$

Where the first inequality follows since $g_{s'}(k'|k)$ are positive. The last inequality follows since we are looking for a bound which holds for all Gibbs-preserving channels and hence we have to choose the worst-case set of transition probabilities $\{r(s'k'|sk)\}$. Due to the stochasticity of the channel which implies that for all s,k we have $\sum_{s',k'} r(s'k'|sk) = 1$ w.l.o.g we can choose a probability distribution which maximizes the sum in (4.136). The maximal value of the sum is achieved when the total probability mass is placed in the transition corresponding to output level i with the largest value of $e^{\beta\epsilon_i}$ (i.e. $i = k_{\min}-1$). This also means that we have:

$$\forall\, s,k \quad \sum_{s',k'} r(s'k'|sk) = 1 \quad \Longrightarrow \quad \forall\, s \quad \sum_{s'} r(s', N-(k-k_{\min})|s,N) = 1. \tag{4.138}$$

Let us denote $q(s'|s) := r(s', N-(k-k_{\min})|s,N)$ which due to (4.138) satisfies $\sum_{s'} q(s'|s) = 1$ for each s. Using now (4.123) and defining $\delta_k := \epsilon_k - \epsilon_{\min} + \delta = \delta(k-k_{\min}+1)$ we

can further rewrite the bound for $h(s', k)$ as:

$$\forall_{k \geq k_{\min}} \qquad h(s', k) \leq 1 + e^{-\beta\delta_k} \sum_s e^{\beta(\widetilde{E}_{s'} - E_s)} r(s', N - (k - k_{\min})|s, N) \tag{4.139}$$

$$= 1 + e^{-\beta\delta_k} e^{\beta\widetilde{E}_{s'}} \sum_s e^{-\beta E_s} q(s'|s). \tag{4.140}$$

Writing above expression explicitly using the definition of $h(s', k)$ from below (4.122) we get:

$$\forall\, s',\ \forall\, k \geq k_{\min} \sum_{s,k'} r(s'k'|sk) e^{\beta(\widetilde{E}_{s'} - E_s + w_{kk'})} \leq 1 + e^{-\beta\delta_k} e^{\beta\widetilde{E}_{s'}} \sum_s e^{-\beta E_s} q(s'|s) \tag{4.141}$$

Let us now multiply both sides of (4.141) by a factor $e^{-\beta\widetilde{E}_{s'}}$ and sum over s'. This leads to:

$$\forall\, k \geq k_{\min} \sum_{s,s',k'} p_S(s)\, r(s'k'|sk) e^{\beta(w_{kk'} - f_s)} \leq \widetilde{Z}_S + e^{-\beta\delta_k} \sum_s e^{-\beta E_s} \left(\sum_{s'} q(s'|s) \right) \tag{4.142}$$

$$= \widetilde{Z}_S \left(1 + e^{-\beta\delta_k} \right), \tag{4.143}$$

where $\widetilde{Z}_S$ is the partition function of the final Hamiltonian on S. Let us now denote the conditional probability distribution $p(s, k'|k) := \sum_{s'} p_S(s)\, r(s'k'|sk)$. This allows us to write:

$$\forall\, k \geq k_{\min} \qquad \sum_{s,k'} p(s, k'|k) e^{\beta(w_{kk'} - f_s)} = \langle e^{\beta(w_{kk'} - f_s)} \rangle_k \leq \widetilde{Z}_S \left(1 + e^{-\beta\delta_k} \right), \tag{4.144}$$

where averaging $\langle \cdot \rangle_k$ is over random variables f_s and $w_{kk'}$ for a given k and can be obtained by setting $f_{s'} = 0$ in (4.121). This proves the Result. Notice further that we prove the inequality by choosing a particular subset of feasible transition probabilities $\{r(s'k'|sk)\}$ and hence the above bound can be saturated. □

It turns out that a more refined version of Result 4.1 will be useful when proving Result 4.2. Let us return to the expression (4.122). For all $k \geq k_{\min}$ we have:

$$\langle e^{\beta(f_{s'} - f_s + w_{kk'})} \rangle_k = \sum_{s'} p_S(s')\, h(s', k) \leq 1 + e^{-\beta\delta_k} \sum_{s,s'} p_S(s') q(s'|s) e^{\beta(\widetilde{E}_{s'} - E_s)} \tag{4.145}$$

$$\leq 1 + e^{-\beta\delta_k} e^{-\beta E^{\max}_{S'}} Z_S \tag{4.146}$$

$$= 1 + \eta_S e^{-\beta\delta_k}, \tag{4.147}$$

where $\eta_S := Z_S e^{\beta E^{\max}_{S'}}$ and $E^{\max}_{S'} := \max_{s'} \widetilde{E}_{s'}$. We will use this expression when proving Result 4.2.

4.5.2 Proof of Result 4.2

In this section we present the proof of Result 4.2.

Proof. The average work $\langle w \rangle$ can be written as:

$$\langle w \rangle = \sum_w p(w) \cdot w = \sum_w \sum_{k=0}^{N} p_W(k)\, p(w|k) = \langle w \rangle_{\text{vac}} + \langle w \rangle_{\text{inv}}, \tag{4.148}$$

where we labeled:

$$\langle w \rangle_{\text{vac}} := \sum_w \sum_{k<k_{\min}} p_W(k)\, p(w|k) \cdot w, \tag{4.149}$$

$$\langle w \rangle_{\text{inv}} := \sum_w \sum_{k\geq k_{\min}} p_W(k)\, p(w|k) \cdot w \tag{4.150}$$

We can further write this using transition probabilities $\{r(s'k'|sk)\}$ as:

$$\langle w \rangle_{\text{vac}} = \sum_{k=0}^{k_{\min}-1} \sum_{s,s',k'} p_W(k)\, p_S(s)\, r(s'k'|sk)\, w_{kk'}, \tag{4.151}$$

and similarily for $\langle w \rangle_{\text{inv}}$. The transition probabilities $\{r(s'k'|sk)\}$ satisfy Gibbs-preserving conditions, i.e. $r(s'k'|sk) \leq e^{-\beta(\epsilon_{k'}-\epsilon_k)}\, e^{-\beta(\widetilde{E}_{s'}-E_s)}$. This implies:

$$\sum_{s,s'} p_S(s)\, r(s'k'|sk) \leq \sum_{s,s'} p_S(s)\, e^{-\beta w_{kk'}}\, e^{-\beta(\widetilde{E}_{s'}-E_s)} \tag{4.152}$$

$$= e^{\beta E^{\max}_{S'}}\, e^{-\beta w_{kk'}}\, Z_S \tag{4.153}$$

$$= \eta_S e^{-\beta w_{kk'}}, \tag{4.154}$$

Plugging this into our expression for $\langle w \rangle_{\text{vac}}$ gives:

$$\begin{aligned}
\langle w \rangle_{\text{vac}} &\leq \eta_S \cdot \sum_{k=0}^{k_{\min}-1} \sum_{k'=0}^{N} p_W(k)\, e^{-\beta(\epsilon_{k'}-\epsilon_k)} \cdot (\epsilon_{k'} - \epsilon_k) \\
&= -\eta_S \cdot \sum_{k=0}^{k_{\min}-1} p_W(k) \cdot \frac{\partial}{\partial \beta} \left(\sum_{k'=0}^{N} e^{-\beta(\epsilon_{k'}-\epsilon_k)} \right) \\
&= -\eta_S \frac{\partial}{\partial \beta} \left[\sum_{k=0}^{k_{\min}-1} \frac{p_W(k)}{g_W(k)} \right].
\end{aligned} \tag{4.155}$$

Consider now the second term from (4.148):

$$\langle w \rangle_{\text{inv}} = \sum_{k=k_{\min}}^{N} \sum_{s,s',k'} p_W(k)\, p_S(s)\, r(s'k'|sk)\, w_{kk'}.$$

Notice that for $k \geq k_{\min}$ we satisfy the assumptions of Result 4.1. Thus, multiplying (4.145) by $p_W(k)$ and summing over $k \geq k_{\min}$ we obtain the following bound:

$$\sum_{k=k_{\min}}^{N} \sum_{s,s',k'} p_W(k) p_S(s) r(s'k'|sk) e^{\beta(f_{s'}-f_s+w_{kk'})} \leq 1 + \eta_S \left(\sum_{k=k_{\min}}^{N} p_W(k) e^{-\beta\delta_k} \right). \tag{4.156}$$

Using the convexity of the exponential function, taking the logarithm of both sides of (4.156) and using the definition of $B_\beta(\rho)$ from the theorem we find:

$$\begin{aligned} B_\beta(\rho_W) &\geq \sum_{k=k_{\min}}^{N} \sum_{s,s',k'} p_S(s) p_W(k) r(s'k'|sk)(f_{s'} - f_s + w_{kk'}) \\ &= \sum_{k=k_{\min}}^{N} \sum_{s,s',k'} p_S(s) p_W(k) r(s'k'|sk)(f_{s'} - f_s) + \langle w \rangle_{\text{inv}} \\ &= \Delta F_S - \sum_{k=0}^{k_{\min}-1} \sum_{s,s',k'} p_S(s) p_W(k) r(s'k'|sk)(f_{s'} - f_s) + \langle w \rangle_{\text{inv}}. \end{aligned}$$

where we used:

$$\Delta F_S = \sum_{s,k,s',k'} p_S(s) p_W(k) r(s'k'|sk)(f_{s'} - f_s). \tag{4.157}$$

Notice now that we have the following bound:

$$\sum_{k=0}^{k_{\min}-1} \sum_{s,s',k'} p_S(s) p_W(k) r(s'k'|sk) f_{s'} \leq \sum_{k=0}^{k_{\min}-1} \sum_{s'} p_W(k) r(s'|k) f_{s'} \tag{4.158}$$

$$\leq \left(\sum_{k=0}^{k_{\min}-1} p_W(k) \right) \max_{s'} f_{s'} \tag{4.159}$$

$$= \left(\sum_{k=0}^{k_{\min}-1} p_W(k) \right) E_{S'}^{\max}. \tag{4.160}$$

On the other hand we also have:

$$\sum_{k=0}^{k_{\min}-1} \sum_{s,s',k'} p_S(s) p_W(k) r(s'k'|sk) f_s = F(\rho_S) \cdot \left(\sum_{k=0}^{k_{\min}-1} p_W(k) \right), \tag{4.161}$$

where we used the fact that $\sum_{s',k'} r(s'k'|sk) = 1$ for all s, k and $F(\rho_S) = \sum_s p_S(s) f_s$. Combining (4.160) and (4.161) gives:

$$\langle w \rangle_{\text{inv}} \leq -\Delta F_S + B_\beta(\rho_W) + \left(\sum_{k=0}^{k_{\min}-1} p_W(k) \right) \cdot (E_{S'}^{\max} - F(\rho_S)). \tag{4.162}$$

The theorem is proven by combining bounds (4.155) and (4.162). □

4.5.3 Proof of Corrolary 6.1

Let us denote the energy populations of the harmonic oscillator battery by:

$$x_< = \sum_{k<k_{\min}} p_W(k), \qquad x_> = \sum_{k_{\min}\leq k<k^*} p_W(k), \qquad 1-x = \sum_{k\geq k^*} p_W(k), \tag{4.163}$$

where $x_> + x_< = x$ is the occupation of the battery below the energy cut-off $\epsilon_* = k_*\delta$. By rewriting (4.155), we obtain

$$\langle w\rangle_{\text{vac}} \leq \eta_S \sum_{k=0}^{k_{\min}-1} \sum_{k'=0}^{N} p_W(k)\, e^{-\beta(\epsilon_{k'}-\epsilon_k)}(\epsilon_{k'} - \epsilon_k) \tag{4.164}$$

$$= \eta_S Z_W \left(\sum_{k=0}^{k_{\min}-1} p_W(k) e^{\beta\epsilon_k} (\langle E\rangle_\beta - \epsilon_k) \right), \tag{4.165}$$

with $Z_W = \sum_{k'=0}^{N} e^{-\beta\epsilon_{k'}}$ and $\langle E\rangle_\beta = Z_W^{-1} \sum_{k'=0}^{N} e^{-\beta\epsilon_{k'}}\epsilon_{k'}$. We see that (4.165) is upper bounded by

$$\langle w\rangle_{\text{vac}} \leq x_< \cdot \eta_S\, Z_W\, e^{\beta(k_{\min}-1)x} \langle E\rangle_\beta \leq \eta_S\, Z_W\, x\, e^{\beta\epsilon_{\min}} \langle E\rangle_\beta. \tag{4.166}$$

On the other hand notice that the free energy difference $\langle f_{s'} - f_s\rangle_{\text{vac}}$ can be upper bounded in the following way:

$$\langle f_{s'} - f_s\rangle_{\text{vac}} \leq \left(\sum_{k=0}^{k_{\min}} p_W(k) \right) \cdot \left(E_{S'}^{\max} - F_S \right) \leq x \left(E_{S'}^{\max} + \frac{1}{\beta} \log d_S \right). \tag{4.167}$$

Combining the last two expressions yields an upper-bound for $A_\beta(\rho_W, \rho_S)$:

$$A_\beta(\rho_W, \rho_S) \leq x \left(\eta_S\, Z_W e^{\beta\epsilon_{\min}} \langle E\rangle_\beta + E_{S'}^{\max} + \frac{1}{\beta} \log d_S \right) \tag{4.168}$$

Furthermore, the correction stemming from the distance to the vacuum regime can be rewritten as

$$\beta B_\beta(\rho_W) \leq \log\left[1 + \eta_S \sum_{k=k_{\min}}^{k^*-1} p_W(k) e^{-\beta\delta(k-k_{\min}+1)} + \eta_S \sum_{k=k_*}^{N} p_W(k) e^{-\beta\delta(k-k_{\min}+1)} \right)\Bigg] \tag{4.169}$$

$$\leq \eta_S \left(x_>\, e^{\beta\Delta} + (1-x) e^{-\beta\delta(k^*-k_{\min}+1)} \right) \tag{4.170}$$

$$\leq \eta_S \left(x e^{-\beta\delta} + e^{-\beta D} \right), \tag{4.171}$$

with $D := \delta(k^* - k_{\min})$. In the limit $N \to \infty$, we have $Z_W = (1 - e^{-\beta\delta})^{-1}$ and $\langle E\rangle_\beta = \delta \cdot e^{-\beta\delta}(1-e^{-\beta\delta})^{-1}$. Therefore, combining bounds (4.168) with (4.171), together with $\eta_S \leq d_S e^{\beta E_{S'}^{\max}}$, allows us to find the following bound:

$$\langle w\rangle \leq -\Delta F_S + C(\epsilon_*), \tag{4.172}$$

with:

$$C(\epsilon_*) = x(\epsilon_*) \left(c_S e^{-\beta\delta} \left(\frac{1}{\beta} + \frac{\delta e^{\beta\epsilon_{\min}}}{(1-e^{-\beta\delta})^2} \right) + \frac{1}{\beta} \log c_S \right) + \frac{c_S}{\beta} e^{-\beta D(\epsilon_*)}, \tag{4.173}$$

where $c_S := d_S e^{\beta E_{S'}^{\max}}$. We also explicitly marked the dependence of $D(\epsilon_*)$ and $\delta(\epsilon_*)$ on a selection of threshold energy ϵ_* for a given initial state of the battery.

4.5.4 Properties of the map Γ_{osc}

The map is a thermal operation

In this section we show that the map defined via Construction 4.1 is a valid thermal operation. This is equivalent to showing that Γ_{osc} preserves the Gibbs state for block-diagonal states, i.e. $\Gamma_{osc}[\tau_S \otimes \tau_W] = \tau_S \otimes \tau_W$.

Let us begin by taking an arbitrary thermal operation Γ_{wit} defined using a qubit battery. Its action in terms of subchannels $\{\mathcal{R}_{kk'}\}$ can be written as:

$$\Gamma_{wit}[\rho_S \otimes \rho_W] = \sum_{k,k'}^{N=1} p_W(k)\, \mathcal{R}_{kk'}(\rho_S) \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W \,. \tag{4.174}$$

Since by assumption Γ_{wit} preserves the Gibbs state, we can write:

$$\mathcal{R}_{00}(\tau_S) + e^{-\beta\delta}\mathcal{R}_{10}(\tau_S) = \tau_S, \tag{4.175}$$

$$\mathcal{R}_{01}(\tau_S) + e^{-\beta\delta}\mathcal{R}_{11}(\tau_S) = e^{-\beta\delta}\tau_S. \tag{4.176}$$

To show that the map Γ_{osc} is a valid thermal operation we will first consider a finite-dimensional harmonic oscillator battery (i.e. we additionally specify the action of the map on the highest energy level N) and show that this completed map, denoted $\widetilde{\Gamma}_{osc}$, is a thermal operation for all N. We will then take the limit $N \to \infty$ and show that in this limit the completed map is indistinguishable from the map described via Construction 6.1 in the main text.

Let us consider a finite dimensional harmonic oscillator battery with Hamiltonian:

$$H_W = \sum_{k=0}^{N} \epsilon_k \, |\epsilon_k\rangle\langle\epsilon_k|_W \,, \tag{4.177}$$

where as before $\epsilon_k = k\delta$.

Gibbs-preserving property. We define the completed map $\widetilde{\Gamma}_{osc}$ as:

$$\rho_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W \to \sum_{k'=0}^{N-1} \mathcal{R}_{00}\mathcal{R}_{01}^{k'}(\rho_S) \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W + \mathcal{R}_{01}^{N}(\rho_S) \otimes |\epsilon_N\rangle\langle\epsilon_N|_W \,, \tag{4.178}$$

$$\rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W \to \mathcal{R}_{10}(\rho_S) \otimes |\epsilon_{k-1}\rangle\langle\epsilon_{k-1}|_W + \sum_{k'=k}^{N-1} \mathcal{R}_{00}\mathcal{R}_{01}^{k'-k}\mathcal{R}_{11}(\rho_S) \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W + \mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\rho_S) \otimes |\epsilon_N\rangle\langle\epsilon_N|_W \,, \tag{4.179}$$

$$\rho_S \otimes |\epsilon_N\rangle\langle\epsilon_N|_W \to \mathcal{R}_{10}(\rho_S) \otimes |\epsilon_{N-1}\rangle\langle\epsilon_{N-1}|_W + \mathcal{R}_{11}(\rho_S) \otimes |\epsilon_N\rangle\langle\epsilon_N|_W \tag{4.180}$$

where $0 < k < N$. Let us now apply the extended map $\widetilde{\Gamma}_{osc}$ to the Gibbs state $\tau_S \otimes \tau_W$ and look at a single energy level $|\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W$ of the battery, i.e:

For $k' = 0$ we have:

$$\mathrm{tr}_W \left[(\mathbb{1}_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W)\widetilde{\Gamma}_{osc}(\tau_S \otimes \tau_W)\,(\mathbb{1}_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W) \right] = \tag{4.181}$$

$$g_W(0)\,\mathcal{R}_{00} + g_W(1)\mathcal{R}_{10}(\tau_S) = g_W(0)\tau_S. \tag{4.182}$$

For $0 < k' < N$ we have:

$$\mathrm{tr}_W \left[(\mathbb{1}_S \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W)\, \widetilde{\Gamma}_{osc}(\tau_S \otimes \tau_W)\, (\mathbb{1}_S \otimes |\epsilon_{k'}\rangle\langle\epsilon_{k'}|_W) \right] = \tag{4.183}$$

$$= g_W(0)\, \mathcal{R}_{00}\mathcal{R}_{01}^{k'}(\tau_S) + \sum_{k=1}^{k'} g_W(k)\, \mathcal{R}_{00}\mathcal{R}_{01}^{k'-k}\mathcal{R}_{11}(\tau_S) + g_W(k'+1)\mathcal{R}_{10}(\tau_S) \tag{4.184}$$

$$= g_W(0)\mathcal{R}_{00}\mathcal{R}_{01}^{k'-1} \left[\mathcal{R}_{01}(\tau_S) + e^{-\beta\delta}\mathcal{R}_{11}(\tau_S)\right] + \tag{4.185}$$

$$\sum_{k=2}^{k'} g_W(k)\, \mathcal{R}_{00}\mathcal{R}_{01}^{k'-k}\mathcal{R}_{11}(\tau_S) + g_W(k'+1)\mathcal{R}_{10}(\tau_S). \tag{4.186}$$

Let us now repeat the last step $k' - 1$ times. This would give us:

$$= g_W(k'-1)\mathcal{R}_{00} \left[\mathcal{R}_{01}(\tau_S) + e^{-\beta\delta}\mathcal{R}_{11}(\tau_S)\right] \tag{4.187}$$

$$+ \sum_{k=k'}^{k'} g_W(k)\, \mathcal{R}_{00}\mathcal{R}_{01}^{k'-k}\mathcal{R}_{11}(\tau_S) + g_W(k'+1)\mathcal{R}_{10}(\tau_S) \tag{4.188}$$

$$= g_W(k')\mathcal{R}_{00}(\tau_S) + g_W(k'+1)\mathcal{R}_{10}(\tau_S) \tag{4.189}$$

$$= g_W(k') \left[\mathcal{R}_{00}(\tau_S) + e^{-\beta\Delta}\mathcal{R}_{10}(\tau_S)\right] \tag{4.190}$$

$$= g_W(k')\tau_S. \tag{4.191}$$

For $k' = N$ we have:

$$\mathrm{tr}_W \Big[(\mathbb{1}_S \otimes |\epsilon_N\rangle\langle\epsilon_N|_W)\, \widetilde{\Gamma}_{osc}(\tau_S \otimes \tau_W)\, (\mathbb{1}_S \otimes |\epsilon_N\rangle\langle\epsilon_N|_W) \Big] \tag{4.192}$$

$$= g_W(0)\mathcal{R}_{01}^{N}(\tau_S) + \sum_{k=1}^{N} g_W(k)\mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\tau_S) \tag{4.193}$$

$$= g_W(0)\mathcal{R}_{01}^{N-1} \left[\mathcal{R}_{01}(\tau_S) + e^{-\beta\delta}\mathcal{R}_{11}(\tau_S)\right] + \sum_{k=2}^{N} g_W(k)\mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\tau_S) \tag{4.194}$$

$$= g_W(1)\mathcal{R}_{01}^{N-1}(\tau_S) + \sum_{k=2}^{N} g_W(k)\mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\tau_S) \tag{4.195}$$

$$= \quad \ldots \text{repeat } (N-1) \text{ times} \ldots$$

$$= g_W(N-1)\mathcal{R}_{01}(\tau_S) + g_W(N)\mathcal{R}_{11}(\tau_S) \tag{4.196}$$

$$= g_W(N)\tau_S.$$

The total state of the system and the battery is then given by:

$$\widetilde{\Gamma}_{osc}[\tau_S \otimes \tau_W] = \sum_{k'=0}^{N} g_W(k')\tau_S \otimes \tau_W = \tau_S \otimes \tau_W. \tag{4.197}$$

Trace-preserving property property. It can be also checked that the map Γ_{osc} is trace preserving since Γ_{wit} is trace preserving, i.e. for any ρ_S we have:

$$\mathrm{tr}\, \mathcal{R}_{00}[\rho_S] + \mathrm{tr}\, \mathcal{R}_{01}[\rho_S] = \mathrm{tr}\, \rho_S, \tag{4.198}$$

$$\mathrm{tr}\, \mathcal{R}_{10}[\rho_S] + \mathrm{tr}\, \mathcal{R}_{11}[\rho_S] = \mathrm{tr}\, \rho_S. \tag{4.199}$$

Hence for $k = 0$ we have:

$$\begin{aligned}
\operatorname{tr}\widetilde{\Gamma}_{osc}[\rho_S \otimes |\epsilon_0\rangle\langle\epsilon_0|] &= \sum_{k'=0}^{N-1} \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{01}^{k'}(\rho_S) + \operatorname{tr}\mathcal{R}_{01}^{N}(\rho_S) && (4.200)\\
&= \operatorname{tr}\mathcal{R}_{00}(\rho_S) + \ldots + \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{01}^{N-1}(\rho_S) + \operatorname{tr}\mathcal{R}_{01}^{N}(\rho_S) && (4.201)\\
&= \operatorname{tr}\mathcal{R}_{00}(\rho_S) + \ldots + \operatorname{tr}\mathcal{R}_{01}^{N-2}(\rho_S) && (4.202)\\
&= \quad \ldots \text{repeat } (N-1) \text{ times} \ldots && (4.203)\\
&= \operatorname{tr}\mathcal{R}_{00}(\rho_S) + \operatorname{tr}\mathcal{R}_{01}(\rho_S) && (4.204)\\
&= \operatorname{tr}\rho_S. && (4.205)
\end{aligned}$$

For $0 < k < N$ we have:

$$\begin{aligned}
&\operatorname{tr}\widetilde{\Gamma}_{osc}[\rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|] && (4.206)\\
&= \operatorname{tr}\mathcal{R}_{10}(\rho_S) + \sum_{k'=k}^{N-1} \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{01}^{k'-k}\mathcal{R}_{11}(\rho_S) + \operatorname{tr}\mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\rho_S) && (4.207)\\
&= \operatorname{tr}\mathcal{R}_{10}(\rho_S) + \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{11}(\rho_S) + \ldots + \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{01}^{N-k-1}\mathcal{R}_{11}(\rho_S) + \operatorname{tr}\mathcal{R}_{01}^{N-k}\mathcal{R}_{11}(\rho_S) && (4.208)\\
&= \operatorname{tr}\mathcal{R}_{10}(\rho_S) + \operatorname{tr}\mathcal{R}_{00}\mathcal{R}_{11}(\rho_S) + \ldots + \operatorname{tr}\mathcal{R}_{01}^{N-k-1}\mathcal{R}_{11}(\rho_S) && (4.209)\\
&= \quad \ldots \text{repeat } (N-k-1) \text{ times} \ldots && (4.210)\\
&= \operatorname{tr}\mathcal{R}_{10}(\rho_S) + \operatorname{tr}\mathcal{R}_{11}(\rho_S) && (4.211)\\
&= \operatorname{tr}\rho_S. && (4.212)
\end{aligned}$$

For $k = N$ we have:

$$\operatorname{tr}\widetilde{\Gamma}_{osc}[\rho_S \otimes |\epsilon_N\rangle\langle\epsilon_N|] = \operatorname{tr}\mathcal{R}_{10}(\rho_S) + \mathcal{R}_{11}(\rho_S) = \operatorname{tr}\rho_S. \tag{4.213}$$

Hence $\widetilde{\Gamma}_{osc}$ is trace-preserving which completes the proof that it is a valid thermal operation.

Let us now take the limit $N \to \infty$ and study the action of the map $\widetilde{\Gamma}_{osc}$. Notice that due to (4.176) for all positive δ we have that $\mathcal{R}_{01}^{N}(\rho_S) \to 0$ for $N \to \infty$ for all ρ_S. Hence in the limit the map $\widetilde{\Gamma}_{osc}$ becomes:

$$\begin{aligned}
\rho_S \otimes |\epsilon_0\rangle\langle\epsilon_0|_W &\to \sum_{i=0}^{\infty} \mathcal{R}_{00}\mathcal{R}_{01}^{i}(\rho_S) \otimes |\epsilon_i\rangle\langle\epsilon_i|_W\,, && (4.214)\\
\rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|_W &\to \mathcal{R}_{10}(\rho_S) \otimes |\epsilon_{k-1}\rangle\langle\epsilon_{k-1}|_W + && (4.215)\\
&\qquad \sum_{i=0}^{\infty} \mathcal{R}_{00}\mathcal{R}_{01}^{i}\mathcal{R}_{11}(\rho_S) \otimes |\epsilon_{i+k}\rangle\langle\epsilon_{i+k}|_W\,, && (4.216)
\end{aligned}$$

where $0 < k < \infty$, and hence it is indistinguishable from Γ_{osc}.

The average work under Γ_{osc}

It is easy to see that the action of Γ_{osc} does not depend on the initial state of the battery ρ_W, provided that the state does not have support on the ground state $|\epsilon_0\rangle\langle\epsilon_0|_W$. This allows to easily compute the average work $\langle w\rangle_N$ associated with Γ_{osc}, as well as its effective action on the system S.

Let us begin by first studying the corresponding map $\widetilde{\Gamma}_{osc}$ acting on a finite $N+1$ dimensional harmonic oscillator battery. The energy change of the battery after applying $\widetilde{\Gamma}_{osc}$ to an input state $\rho_{SW} = \rho_S \otimes |\epsilon_k\rangle\langle\epsilon_k|$ will in general be different for different k. The energy change $\Delta E_W^{(k)}$ associated with battery starting in eigenstate $|\epsilon_k\rangle\langle\epsilon_k|_W$ is then given by:

$$\Delta E_W^{(k)} = \tilde{E}_W^{(k)} - E_W^{(k)}, \tag{4.217}$$

where $E_W^{(k)} = \mathrm{tr}[H_W |\epsilon_k\rangle\langle\epsilon_k|] = \epsilon_k$, $\tilde{E}_W^{(k)} = \mathrm{tr}\left[H_W \rho'_W\right]$, H_W is the finite harmonic oscillator Hamiltonian as in (4.177) and $\rho'_W = \mathrm{tr}_S\, \widetilde{\Gamma}_{osc}\left[\rho_{SW}\right]$. Denoting the number of energy levels above ϵ_k with n, i.e $n = N - k$, we can compute the final energy $\tilde{E}_W^{(k)}$ as:

$$\tilde{E}_W^{(k)} = (\epsilon_k - \delta)\,\mathrm{tr}\,\mathcal{R}_{10}\left[\rho\right] + \sum_{j=0}^{n-1}(\epsilon_k + j\Delta)\ \mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{j}\mathcal{R}_{11}\left[\rho_S\right] + (\epsilon_k + n\delta)\ \mathrm{tr}\,\mathcal{R}_{01}^{n}\mathcal{R}_{11}\left[\rho_S\right] \tag{4.218}$$

$$= (\epsilon_k - \delta)\Big(\mathrm{tr}\,\mathcal{R}_{10}\left[\rho_S\right] + \sum_{j=0}^{n-1}\mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{j}\mathcal{R}_{11}\left[\rho_S\right] + \mathrm{tr}\,\mathcal{R}_{10}^{n}\mathcal{R}_{11}\left[\rho_S\right]\Big) + \tag{4.219}$$

$$\delta\sum_{j=0}^{n-1}(j+1)\ \mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{j}\mathcal{R}_{11}\left[\rho_S\right] + \delta(n+1)\ \mathrm{tr}\,\mathcal{R}_{01}^{n}\mathcal{R}_{11}\left[\rho_S\right] \tag{4.220}$$

$$= (\epsilon_k - \delta)\left(\mathrm{tr}\,\mathcal{R}_{10}\left[\rho_S\right] + \mathrm{tr}\,\mathcal{R}_{11}\left[\rho_S\right]\right) + \delta\sum_{j=0}^{n-2}(j+1)\,\mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{j}\mathcal{R}_{11}[\rho_S] + \tag{4.221}$$

$$n\,\mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{n-1}\mathcal{R}_{11}[\rho_S] + n\,\mathrm{tr}\,\mathcal{R}_{01}\mathcal{R}_{01}^{n-1}\mathcal{R}_{11}[\rho_S] + \mathrm{tr}\,\mathcal{R}_{01}^{n}\mathcal{R}_{11}[\rho_S] \tag{4.222}$$

$$\overset{(2)}{=} (\epsilon_k - \delta) + \delta\sum_{j=0}^{n-2}(j+1)\,\mathrm{tr}\,\mathcal{R}_{00}\mathcal{R}_{01}^{j}\mathcal{R}_{11}[\rho_S] + \delta n\,\mathrm{tr}\,\mathcal{R}_{01}^{n-1}\mathcal{R}_{11}[\rho_S] + \mathrm{tr}\,\mathcal{R}_{01}^{n}\mathcal{R}_{11}[\rho_S] \tag{4.223}$$

$$= (\epsilon_k - \delta) + \mathrm{tr}\,\mathcal{R}_{11}[\rho_S] + \mathrm{tr}\,\mathcal{R}_{01}\mathcal{R}_{11}[\rho_S] + \mathrm{tr}\,\mathcal{R}_{01}^{2}\mathcal{R}_{11}[\rho_S] + \ldots + \mathrm{tr}\,\mathcal{R}_{01}^{n}[\rho_S]\mathcal{R}_{11}[\rho_S] \tag{4.224}$$

$$= (\epsilon_k - \delta) + \sum_{j=0}^{n}\mathrm{tr}\,\mathcal{R}_{01}^{j}\mathcal{R}_{11}[\rho_S] \tag{4.225}$$

$$= \epsilon_k + \delta\left(\sum_{j=0}^{n}\mathrm{tr}\,\mathcal{R}_{01}^{j}\mathcal{R}_{11}[\rho_S] - 1\right). \tag{4.226}$$

In line (2) we repeatedly applied the fact that Γ_{wit} is trace-preserving, i.e

$$\sum_{k'}\mathrm{tr}\,\mathcal{R}_{kk'}[\rho_S] = \mathrm{tr}\,\rho_S \qquad \text{for all} \quad k. \tag{4.227}$$

Taking the limit of infinite battery $N \to \infty$ implies $n \to \infty$, hence the energy change associated with map Γ_{osc} can be expressed as:

$$\Delta E_W = \langle w\rangle_{osc} = \delta\cdot\left(\sum_{i=0}^{\infty}\mathrm{tr}\,\mathcal{R}_{01}^{i}\circ\mathcal{R}_{11}[\rho_S] - 1\right), \tag{4.228}$$

Notice that this energy change (and hence work) does not depend on the choice of the initial state $|\epsilon_k\rangle\langle\epsilon_k|$. Due to the linearity of Γ_{osc} this means that the map would

yield the same amount of work if we instead chose any combination of energy levels, i.e. any initial state ρ_W.

The expression for average work (4.228) can be further simplified if ρ_S is a diagonal state. In that case the action of subchannels $\{R_{kk'}\}$ can be fully described by a set of substochastic matrices $\{R_{kk'}\}$ acting on vectors composed of the diagonal parts of the states. In that case the infinite matrix geometric series in (4.228) can be explicitly computed, i.e. $\sum_{i=0}^{\infty} R_{01}^i = (\mathbb{1} - R_{01})^{-1}$. Denoting the diagonal of ρ_S with a vector $\mathbf{x}$ the average work can be expressed as:

$$\langle w \rangle_{osc} = \delta \cdot \left(\mathbf{1}^T (\mathbb{1} - R_{01})^{-1} R_{11} \mathbf{x} - 1 \right), \tag{4.229}$$

where $\mathbf{1}^T = (1, 1, \ldots, 1)$ is the (horizontal) identity vector and $\mathbb{1}$ is the identity matrix.

The map Γ_{osc} satisfies the ETI property

Here we are going to show that for diagonal states the map Γ_{osc} satisfies the ETI property. Hence we are going to show that transition probabilities $\{r(s'k'|sk)\}$ corresponding to Γ_{osc} satisfy:

$$r(s'k'|sk) = r(s', k' + n|s, k + n) \tag{4.230}$$

for all s, s' and integers n such that $k' + n \geq 0$ and $k + n \geq k_{\min}$ with $k_{\min} = 1$.

Notice first that due to (4.49) the condition (4.230) can be equivalently written using battery subchannels $\{\mathcal{R}_{kk'}\}$ as:

$$\mathcal{R}_{kk'} = \mathcal{R}_{k+n,k'+n}, \tag{4.231}$$

for all integers n such that $k' + n \geq 0$ and $k + n \geq 1$. Consider the battery subchannels associated with the map Γ_{osc} for $k \geq k_{\min} = 1$, i.e.:

$$\mathcal{R}_{kk'} = \sum_{i=0}^{\infty} \mathcal{R}_{00} \mathcal{R}_{01}^i \mathcal{R}_{11} \delta_{i+k,k'} + \mathcal{R}_{10} \delta_{k-1,k'} = \mathcal{R}_{00} \mathcal{R}_{01}^{k'-k} \mathcal{R}_{11} + \mathcal{R}_{10} \delta_{k-1,k'} \tag{4.232}$$

The shifted battery subchannels are given by:

$$\begin{aligned} \mathcal{R}_{k+n,k'+n} &= \sum_{i=0}^{\infty} \mathcal{R}_{00} \mathcal{R}_{01}^i \mathcal{R}_{11} \delta_{i+k+n,k'+n} + \mathcal{R}_{10} \delta_{k+n-1,k'+n} && (4.233) \\ &= \mathcal{R}_{00} \mathcal{R}_{01}^{k'-k} \mathcal{R}_{11} + \mathcal{R}_{10} \delta_{k-1,k'} && (4.234) \\ &= \mathcal{R}_{kk'}. && (4.235) \end{aligned}$$

Hence we can conclude that the map Γ_{osc} satisfies the ETI property.

Chapter 5

Operational significance of nonclassical teleportation

5.1 Introduction

5.1.1 Background

Quantum teleportation (Bennett et al., 1993) has fundamentally changed our way of understanding information. Although the famous protocol proposed by Bennett *et. al.* (ibid.) will never allow us to "beam up", it is by no means less spectacular and has become one of the most important protocols in quantum information theory. In its standard form it involves transferring an unknown quantum state to a remote recipient using classical communication and pre-shared entanglement. Although nothing actually moves during the process (despite the classical information carriers), the situation can't be meaningfully distinguished from one in which the original state has been transported to another location. To date it has been demonstrated in a wide range of experiments (Bouwmeester et al., 1997; Boschi et al., 1998; Furusawa et al., 1998; Bao et al., 2012; Leuenberger, Flatté, and Awschalom, 2005; Pirandola et al., 2015; Vaidman, 1994; Sherson et al., 2006) and is currently one of the building blocks in many quantum information contexts, ranging from distributed quantum networks (Briegel et al., 1998), quantum repeaters (Hasegawa et al., 2019), quantum computers (Gottesman and Chuang, 1999) and even the future quantum internet (Kimble, 2008).

In the ideal version of teleportation Alice and Bob share a maximally entangled state and Alice is given a system in some unknown state. She performs a Bell-state measurement on the system and her share of the entangled state and communicates the result to Bob who applies an appropriate unitary correction to his share and transforms it into the state given to Alice.

5.1.2 Motivation

However, in realistic teleportation protocols the states and measurements used are never perfect. This motivates studying a more general teleportation scheme involving arbitrary states and measurements. We will adapt this approach here and assume that Alice and Bob share an arbitrary quantum state and introduce a third party, called the Verifier, who gives Alice states to be teleported. She then applies an arbitrary measurement on her share of the entangled state and the system given to her and communicates the measurement result to Bob, who performs a local correction on his state.

The standard figure of merit used to quantify how well a given teleportation protocol performs is the average fidelity of teleportation, denoted here by $\langle F \rangle$ and defined as the fidelity between the state to be teleported and the final state of Bob's

after the protocol is finished, averaged uniformly over all measurement results and input states. This quantity was first introduced in (Popescu, 1994) and since then has been used widely to quantify the usefulness of states for teleportation (Horodecki, Horodecki, and Horodecki, 1999; Linden and Popescu, 1999; Olmschenk et al., 2009). The average fidelity of teleportation is maximal when teleportation is perfect, i.e. as in the ideal version. If Alice and Bob do not share an entangled state, or are unable to perform an entangled measurement, then the corresponding teleportation scheme is said to be "classical". For all such schemes the average fidelity can never exceed the threshold value (Horodecki, Horodecki, and Horodecki, 1999)

$$\langle F_c \rangle = \frac{2}{d+1}, \tag{5.1}$$

where d is the dimension of the Hilbert space from which the states to be teleported are drawn. Importantly, it was shown that there exist entangled states in nature (e.g. bound-entangled states (Linden and Popescu, 1999; Bennett et al., 1999; Horodecki et al., 2003; Horodecki, Horodecki, and Horodecki, 1998)) which cannot surpass this classical threshold. This led to a common belief that not all entangled states are useful for quantum teleportation.

However, it was recently shown that the average fidelity is not sufficiently sensitive to probe all aspects of teleportation experiments (Cavalcanti, Skrzypczyk, and Supic, 2017; Supic, Skrzypczyk, and Cavalcanti, 2019). In particular, every entangled state can lead to non-classical teleportation if the full data from the experiment is taken into account (Cavalcanti, Skrzypczyk, and Supic, 2017). To show this, a geometric method of quantifying the non-classicality of teleportation data using a measure called the robustness of teleportation (RoT) was introduced. By showing that the RoT is non-zero whenever Alice and Bob share entanglement and Alice performs a Bell state measurement, it was demonstrated that every entangled state leads to experimental data which could not be produced without entanglement. However, the question of in what sense this non-classical data showed that the entanglement could be considered as being "useful" for teleportation in some operational sense has remained unanswered. We can therefore formulate the following main question.

Main question

What is the operational significance of nonclassical teleportation?

In what follows we will try to answer this question using the theory of duality of convex optimisation.

5.1.3 Contribution

In this work we construct a resource theory of quantum teleportation. Unlike many other resource-theoretic studies in the literature which address a single type of resource, quantum teleportation combines two distinct resources - shared entanglement and entangled measurement. A similar type of "interaction" between resources happens also for the resource theory of boxes (Bell and Buscemi nonlocality). Using this framework we show that RoT admits two natural operational interpretations.

Firstly, it quantifies the advantage enabled by an entangled state and entangled measurement in the task of teleporting unknown quantum correlations – rather than

unknown states – over all classical instruments. This task can be thought of as a natural generalization of entanglement swapping (Zukowski et al., 1993; Pan et al., 1998) where the goal is not only to "swap" entanglement but to achieve pre-defined quantum correlations between parties. We show that the average score in this task when teleporting classical correlations reduces to the average fidelity of teleportation. This also shows a surprising property of bound-entangled states (Horodecki, Horodecki, and Horodecki, 1998) (i.e. states from which no entanglement can be distilled) – they provide advantage over separable states in teleporting genuine quantum correlations. This also answers an open problem from (Cavalcanti, Skrzypczyk, and Supic, 2017) by specifying in what sense all entangled states are useful for teleportation.

Secondly, we show that RoT also quantifies the maximal achievable advantage in the task of subchannel discrimination with quantum side information. This reveals that RoT is another robustness-based quantifier which fits into the program of discrimination tasks, a class of problems with fundamental importance to the field of quantum information (Kitaev, 1997; Acín, 2001; Childs, Preskill, and Renes, 2000). Analogous results have been shown also for entanglement (Vidal and Tarrach, 1999; Bae, Chruscinski, and Piani, 2019; Takagi et al., 2019), coherence (Napoli et al., 2016), EPR-steering (Piani and Watrous, 2015), quantum measurement (Skrzypczyk and Linden, 2019; Ducuara and Skrzypczyk, 2019; Oszmaniec and Biswas, 2019), measurement incompatibility (Designolle et al., 2019; Designolle, Farkas, and Kaniewski, 2019b) and fault-tolerant quantum computation (Howard and Campbell, 2017). This surprising connection allows us to infer that every entangled state can act as a useful quantum memory for local subchannel discrimination.

Finally, by formulating teleportation in the language of resource theories, we show that both tasks provide complete sets of monotones for two natural notions of simulation (free operations), one classical and the other quantum. As we shall see, both of these have their drawbacks and benefits, depending on the specific task exploiting a resourceful teleportation instrument.

5.1.4 Structure

This chapter is structured as follows. In Sec. 5.2 we introduce the relevant framework for describing nonclassical teleportation. In Sec. 5.3 we describe its operational significance, which is also the main result of this Chapter. In Sec. 5.4 we give a summary of this Chapter and highlight a few open problems. Finally, in Sec. 5.5 we give proofs of the main statements discussed in this Chapter.

5.2 Nonclassical teleportation

Let us now describe the most general teleportation protocol that can be performed between two parties. We start by recalling the standard teleportation protocol, defining the relevant notation and summarising our assumptions.

5.2.1 Teleportation instruments

In our study we will assume that Alice and Bob share an arbitrary quantum state ρ^{AB} of dimension $d_A \times d_B$ and the third party, called the Verifier, provides quantum states from a collection of states $\{\omega_x^V\}$, $x = 0, 1, \ldots, n$ of dimension d_V, all of them known to Alice (but not the concrete realization in the experimental run). She then applies a measurement $M_a^{VA} \in$ POVM on her share of the entangled state and input

system, as a result projecting Bob's state into:

$$\rho^B_{a|\omega_x} = \frac{1}{p(a|x)} \mathrm{tr}_{VA}\left[(M_a^{VA} \otimes \mathbb{1}^B)\left(\omega_x^V \otimes \rho^{AB}\right)\right], \tag{5.2}$$

where $p(a|x) = \mathrm{tr}\left[(M_a^{VA} \otimes \mathbb{1}^B)\left(\omega_x^V \otimes \rho^{AB}\right)\right]$ is the probability of a particular outcome a given that state ω_x was provided by the Verifier. For our purposes it will be more convenient to work with unnormalized states and thus we define:

$$\sigma^B_{a|\omega_x} = p(a|x)\rho^B_{a|\omega_x} = \Lambda_a(\omega_x),$$

where $\Lambda_a(\cdot) = \Lambda_a^{V\to B}(\cdot)$ is a subchannel from V to B, labelled by a, which transforms the input states ω_x into (unnormalised) output states $\sigma_{a|\omega_x}$. We will refer to such a collection as a teleportation instrument and denote it with $\mathbb{\Lambda} = \{\Lambda_a\}$. Notice that since the operators M_a^{VA} form a POVM, the corresponding subchannels Λ_a satisfy:

$$\sum_a \Lambda_a(\omega) = \rho^B, \tag{5.3}$$

irrespective of the input state ω. This is reminiscent of the no-signaling condition, meaning that Bob's state cannot change if he does not know Alice's measurement result. In fact, any no-signalling instrument (any set of subchannels satisfying condition (5.3), for all input states) can be written as some teleportation instrument, and vice versa. We prove this important statement in Sec. 5.5.

When the set of states $\{\omega_x\}$ forms a tomographically-complete set, the experiment becomes effectively independent of the input (we show this in more detail in Sec. 5.5). This means that the full information about teleportation instrument can be obtained by probing it with $\{\omega_x\}$. Furthermore, whenever the set of input states is not tomographically-complete one can always reduce the Hilbert space to a smaller subspace for which the set forms a subspace of an operator space. This motivates introducing a notion of a *complete teleportation experiment*, i.e. an experiment in which the set of input states is tomographically-complete. In this chapter, we will focus exclusively on complete teleportation experiments.

Consider now the case when ρ^{AB} is a separable state, i.e. it can be written as

$$\rho^{AB} = \sum_\lambda p(\lambda)\, \rho^A_\lambda \otimes \rho^B_\lambda, \tag{5.4}$$

which we denote by $\rho^{AB} \in \mathsf{SEP}(A:B)$. The associated teleportation instrument takes the form:

$$\begin{aligned}\Lambda^c_a(\omega_x) &= \sum_\lambda p(\lambda)\, \mathrm{tr}_{VA}\left[\left(M_a^{VA} \otimes \mathbb{1}^B\right)\left(\omega_x \otimes \rho^A_\lambda \otimes \rho^B_\lambda\right)\right] \\ &= \sum_\lambda p(\lambda)\, p(a|x,\lambda)\, \rho^B_\lambda,\end{aligned} \tag{5.5}$$

where $p(a|x,\lambda) = \mathrm{tr}\left[M_a^{VA}(\omega_x^V \otimes \rho^A_\lambda)\right]$. This is the most general classical teleportation scheme which can be realized if Alice and Bob have access only to classical randomness λ and the ability to locally prepare quantum states in their labs. In what follows we will consider teleportation instruments arising from a separable shared state as "free", and We will denote the set of all such instruments by $\mathcal{F}_{\mathrm{T}}$, in analogy

with the set of free objects studied in the context of resource theories (see Chapter 2). If the teleportation data $\{\sigma^{\mathrm{B}}_{a|\omega_x}\}$ cannot be explained as coming from a classical teleportation instrument, we will refer to the associated teleportation instrument as "quantum" and denote the set of all such instruments with $\mathcal{R}_{\mathrm{T}}$.

5.2.2 Benchmarking teleportation performance

In the standard approach the quality of a given teleportation instrument is assessed using the average fidelity of teleportation (Popescu, 1994), which in the present context is given by:

$$\langle F\rangle = \max_{\{U_a\}} \frac{1}{n}\sum_{a,x} p(a|x)\langle\omega_x|U_a\rho^{B}_{a|\omega_x}U_a^{\dagger}|\omega_x\rangle, \tag{5.6}$$

where the maximisation is over all correcting unitaries for Bob $\{U_a\}_a$. Unfortunately, this single quantity does not utilize all the data produced in the teleportation experiment.

An alternative method for quantifying how 'close' a set of data is to that which could arise from a classical teleportation instrument is to solve the following convex optimization problem:

$$\begin{aligned}\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}) := \min_{\{\Lambda^c_a\},\{\Lambda'_a\},\eta}\ & r, \\ \text{s.t.}\quad & \frac{1}{1+r}\Lambda_a + \frac{r}{1+r}\Lambda'_a = \Lambda^c_a, \\ & \sum_a \Lambda'_a \in \mathbf{CPTP}, \qquad \{\Lambda^c_a\}\in\mathcal{F}_{\mathrm{T}}, \qquad \{\Lambda'_a\}\in\mathcal{R}_{\mathrm{T}}.\end{aligned} \tag{5.7}$$

where Λ'_a describes the "noise" which comes from some other teleportation instrument $\mathbb{\Lambda}'$ and which has to be added to the teleportation data $\sigma_{a|\omega_x} = \Lambda_a[\omega_x]$ for there to exist an explanation in terms of classical data $\Lambda^c_a(\omega_x)$. This noise is allowed to arise from any teleportation instrument, not necessarily classical one.

The quantity $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ is the (generalized) robustness of teleportation (RoT) and was introduced in (Cavalcanti, Skrzypczyk, and Supic, 2017). We highlight that for complete teleportation experiments the RoT is a function of the teleportation instrument $\mathbb{\Lambda}$ alone, and is independent of the specific set of states used $\{\omega_x\}$, and the data they produce $\{\sigma^{\mathrm{B}}_{a|\omega_x}\}$. We prove this important fact in Sec. 5.5.

Basic properties of the Robustness of Teleportation

Similarly to other robustness-based measures that can be found in the literature (Vidal and Tarrach, 1999; Napoli et al., 2016; Skrzypczyk and Linden, 2019; Uola, Kraft, and Abbott, 2020; Ducuara and Skrzypczyk, 2019), the RoT has a number of useful properties which can be easily deduced from (5.7). Leaving the technical proofs to Sec. 5.5, here we state the most important ones.
(*i*) It is *faithful*, meaning that it vanishes if and only if teleportation instrument is classical, i.e:

$$\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}) = 0 \iff \mathbb{\Lambda}\in\mathcal{F}_{\mathrm{T}}. \tag{5.8}$$

(*ii*) It is *convex*, meaning that having access to teleportation instruments $\mathbb{\Lambda}_1$ and $\mathbb{\Lambda}_2$ one cannot obtain a better one by using them probabilistically, i.e for $\mathbb{\Lambda}' = p\,\mathbb{\Lambda}_1 + (1-p)\,\mathbb{\Lambda}_2$

with $0 \leq p \leq 1$, we have:

$$\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}') \leq p\, \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_1) + (1-p)\, \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_2). \tag{5.9}$$

(*iii*) It is *monotonic* (non-increasing) under quantum and classical simulations that we define below. That is if $\mathbb{\Lambda}'$ can be simulated by $\mathbb{\Lambda}$ using a quantum or a classical simulation then

$$\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}') \leq \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}). \tag{5.10}$$

A quantum simulation is one whereby there exist probability distributions $p(\lambda)$, $p(b|a,\lambda)$ and channels Θ_λ and Ω_λ such that:

$$\Lambda'_b = \sum_{a,\lambda} p(\lambda)\, p(b|a,\lambda)\, \Theta_\lambda \circ \Lambda_a \circ \Omega_\lambda, \tag{5.11}$$

holds for all b. We denote the order induced by this type of simulation by $\mathbb{\Lambda}' \prec_q \mathbb{\Lambda}$. A classical simulation is one whereby there exist probability distributions $p(b|a)$ such that:

$$\Lambda'_b = \sum_{a} p(b|a)\, \Lambda_a, \tag{5.12}$$

holds for all b and is similarly denoted by $\mathbb{\Lambda}' \prec_c \mathbb{\Lambda}$. In the resource-theoretic approach one can think about these maps as free operations of the framework. These two notions of simulation will each be seen to be relevant for one of operational tasks introduced in the next section. Therefore this robustness measure admits all the desirable properties of a resource measure that we highlighted in Chapter 2.

5.3 Results

In this section we show that the robustness of teleportation can be viewed as the maximal achievable advantage when using quantum over classical resources in two unrelated operational tasks. Often it is illustrative to phrase such tasks in terms of games played between parties according to a pre-defined set of rules and scores. We follow this approach here and describe two operational tasks using such games.

5.3.1 Teleportation of quantum correlations

Consider a game played between a Verifier and a collaborative party, Alice and Bob, who try to convince the Verifier about their ability to transfer correlations between spatially separated labs. More explicitly, we consider the following scenario:

1. The Verifier prepares an arbitrary bipartite state $\psi^{V'V}$ and shares one part of this state with Alice.

2. Alice inputs the state she received to the teleportation instrument $\mathbb{\Lambda}' = \{{\Lambda'_b}^{V\to B}\}$ that connects her local lab with Bob's, and which they can locally simulate using $\mathbb{\Lambda}$, obtaining measurement outcome b and changing the joint state into $\rho^{V'B}_{b|\psi} = (\mathrm{id}^{V'} \otimes {\Lambda'_b}^{V\to B})\psi^{V'V}$.

3. Conditioned on the value of b, Bob applies locally an arbitrary correction $\mathcal{E}^B_b \in \mathsf{CPTP}$ to his share of the state and returns the output state and outcome of the measurement to the Verifier.

4. The Verifier assesses the quality of the teleportation instrument by checking the overlap between the joint state after correction $(\mathrm{id}^{V'} \otimes \mathcal{E}_b^B)\rho_{b|\psi}^{V'B}$ and a predefined set of target states $\{\xi_b^{V'B}\}$. If the teleported state is the same as the target state, then Bob receives a score $f(b) \geq 0$.

The game is fully specified by a tuple $\mathcal{G} = \{\psi, \{\xi_b\}, f(b)\}$. The average score using the teleportation instrument $\mathbb{\Lambda}$ is given by:

$$q(\mathcal{G}, \mathbb{\Lambda}) = \max_{\substack{\mathbb{\Lambda}' \prec_q \mathbb{\Lambda} \\ \mathcal{E}_b \in \mathrm{CPTP}}} \sum_b f(b) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_b \otimes \Lambda_b')\psi \cdot \xi_b\right], \tag{5.13}$$

where the optimization ranges over all corrections $\{\mathcal{E}_b\}$ and all teleportation instruments $\mathbb{\Lambda}'$ which can be quantum-simulated using $\mathbb{\Lambda}$, via (5.11).

In Sec. 5.5 we show that the maximal advantage which Bob can achieve using a teleportation instrument $\mathbb{\Lambda} \in \mathcal{R}_\mathrm{T}$ over any classical instrument $\mathbb{\Lambda}^c \in \mathcal{F}_\mathrm{T}$ is fully specified by the robustness of teleportation. More formally, we have the following.

Nonclassical teleportation as a resource for teleporting quantum correlations

Result 5.1. Let $\mathbb{\Lambda}$ be a teleportation instrument and $\mathcal{G}$ describe the task of teleporting quantum correlations. Then:

$$\max_{\mathcal{G}} \frac{q(\mathcal{G}, \mathbb{\Lambda})}{q^c(\mathcal{G})} = 1 + \mathbf{R}_\mathrm{T}(\mathbb{\Lambda}), \tag{5.14}$$

where the classical score, that is the maximal score which can be achieved using classical resources in the same game, is given by:

$$q^c(\mathcal{G}) = \max_{\mathbb{\Lambda}^c \in \mathcal{F}_\mathrm{T}} q(\mathcal{G}, \mathbb{\Lambda}^c) \leq 1/d_V \tag{5.15}$$

The proof technique is to (i) use (5.7) to show that $1 + \mathbf{R}_\mathrm{T}(\mathbb{\Lambda})$ is an upper bound on the advantage for all games $\mathcal{G}$; (ii) use duality theory of convex optimisation to find the dual form of (5.7) and construct a game $\mathcal{G}^*$ from the optimal dual variables that saturate the bound. See Sec. 5.5 for details.

Connection with the average fidelity of teleportation

It is interesting that the average fidelity $\langle F \rangle$ can be viewed as the average score in this type of task for a particular game $\mathcal{G}$. To see this, consider a setting in which the verifier provides a uniform classical-quantum (CQ) state

$$\psi^* = \frac{1}{n} \sum_x |x\rangle\langle x| \otimes |\omega_x\rangle\langle\omega_x| \tag{5.16}$$

and demands that the state returned by Bob is exactly the same for all b, that is $\xi_b^* = \psi^*$. For each transmission the Verifier will give Bob the same score $f^*(b) = n$. This defines a game $\mathcal{G}^* = \{\psi^*, \xi_b^*, f^*(b)\}$, whose average score is:

$$q(\mathcal{G}^*, \mathbb{\Lambda}) = \max_{\mathcal{E}_a \in \mathrm{CPTP}} \frac{1}{n} \sum_{a,x} p(a|x)\langle\omega_x|\mathcal{E}_a(\rho_{b|\omega_x}^B)|\omega_x\rangle. \tag{5.17}$$

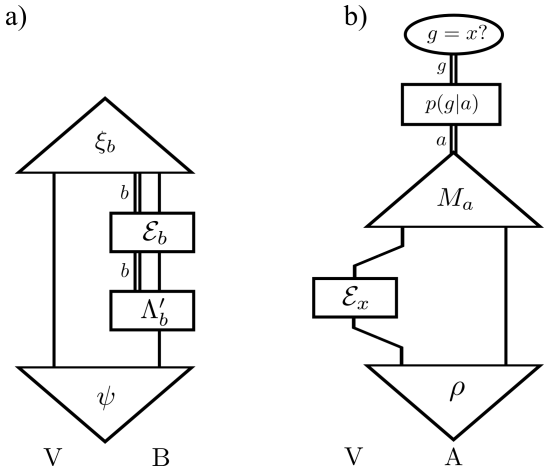

FIGURE 5.1: The two operational tasks. Fig. (a) presents teleportation of quantum correlations specified by $\mathcal{G} = \{\psi, \xi_b, f(b)\}$, where ψ is the input state, $\{\xi_b\}_b$ are the target states and $f(b)$ is the score given when a correct target state is obtained. Bob is allowed to perform any local quantum simulation of his teleportation instrument $\mathbb{\Lambda}$, i.e. he has access to Λ'_b of the form (5.11). Fig. (b) shows the entanglement-assisted subchannel discrimination task which involves a set of subchannels to discriminate $\mathbb{E} = \{\mathcal{E}_x\}_x$ and uses quantum resources of the teleportation experiment (bipartite state and measurement).

This is the ordinary average fidelity (5.6), except that Bob is allowed to use an arbitrary correction $\mathcal{E}_b$, instead of a unitary one. An interesting feature of this game is that Bob doesn't need to tell the Verifier which measurement result occurred.

This provides insight into why not all entangled states are 'useful' for teleportation. Since the average fidelity of teleportation corresponds to a game in which the Verifier asks Bob to transfer classical correlations, the fact that $\langle F \rangle$ cannot surpass the classical threshold for some entangled states only means that they cannot be used to transfer classical correlations better than the optimal classical state. However, if the verifier poses a more difficult task where the correlations to be transferred are genuinely quantum, then all entangled states can outperform classical states for a specific choice of target states. Alternatively, one can view this task as generalising from teleportation to entanglement swapping. More specifically, the Verifier and Bob initially start uncorrelated. Then they use a fixed teleportation instrument shared between Alice and Bob, and a fixed entangled state between the Referee and Alice to achieve a given ensemble of correlated states between the Referee and Bob.

5.3.2 Entanglement-assisted subchannel discrimination with fixed measurement

Let us now consider the task of sub-channel discrimination, where the player is allowed to use a fixed entangled state to assist them, and only has the ability to perform a fixed entangled measurement. The task is specified by a collection of subchannels, $\mathbb{E} = \{\mathcal{E}_x\}$, which form an instrument, i.e. $\sum_x \mathcal{E}_x = \mathcal{E} \in \mathsf{CPTP}$. The resources of the player will be specified by $\mathbb{A} = \{\{M_a\}, \rho\}$, where $\{M_a\} \in \mathrm{POVM}$ is a bipartite measurement and ρ is a bipartite state. We use a similar letter to emphasize that these resources are exactly the same as in the case of nonclassical teleportation, the only difference being how they are formed to play the game. We consider the following game set-up:

1. Alice sends one half of the state ρ^{VA} to the verifier.

2. The verifier applies a subchannel $\mathcal{E}_x^V$ from the instrument $\mathbb{E}$ to their share of ρ^{VA}, which prepares the state $\rho_x^{VA} = (\mathcal{E}_x^V \otimes \mathrm{id}^A)\rho^{VA}$ with probability $p(x|\rho) = \mathrm{tr}\,\rho_x^{VA}$. The verifier then passes their share back to Alice.

3. Alice uses the measurements $\{M_a^{VA}\}$ to identify which subchannel $\mathcal{E}_x^V$ was applied. Based on her measurement outcome a she produces a guess of x, denoted g, according to a strategy $p(g|a)$.

The average probability of guessing which subchannel was applied when having access to a pre-shared state ρ and bipartite measurement $\{M_a\}$, optimized over all post-processings $p(g|a)$ is given by:

$$p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}) = \max_{p(g|a)} \sum_{x,a,g} p(g|a)\,\mathrm{tr}\left[M_a \cdot (\mathcal{E}_x \otimes \mathrm{id})\rho\right] \delta_{g,x}. \tag{5.18}$$

We will compare this success probability to the best success probability Alice could achieve if she had access to only classical resources. In particular, if the state used by Alice is separable, then we will say that she uses a classical strategy (or an LOSR-free strategy, in the terminology of Ref. (Schmid, Rosset, and Buscemi, 2020)). The (maximal) average guessing probability for such a classical strategy is given by

$$p_{\mathrm{succ}}^c(\mathbb{E}) = \max_{\mathbb{A}^c \in \mathcal{F}_{\mathrm{T}}} p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}^c), \tag{5.19}$$

where $\mathbb{A}^c$ denotes a classical resource (separable state with arbitrary measurement). It can be shown (see Sec. 5.5) that the maximal classical probability of guessing can be equivalently written as:

$$p_{\mathrm{succ}}^c(\mathbb{E}) = \max_{\sigma} \max_{x} \ \mathrm{tr}\,\mathcal{E}_x(\sigma). \tag{5.20}$$

With this in mind we can now state the second main result of this chapter. In Sec. 5.5 we prove in detail the following fact.

Nonclassical teleportation as a resource for subchannel discrimination

Result 5.2. Let $\mathbb{\Lambda}$ be a teleportation instrument formed by a measurement $\{M_a\}$ and a state ρ. Then the maximal advantage offered by the strategy $\mathbb{A} = \{\{M_a\}_a, \rho\}$ over the best classical strategy in entanglement-assisted subchannel discrimination with fixed measurement is given by

$$\max_{\mathbb{E}} \frac{p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A})}{p_{\mathrm{succ}}^c(\mathbb{E})} = 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}), \tag{5.21}$$

Thus, the maximal advantage is constant among all strategies $\mathbb{A}$ that lead to the same teleportation instrument $\mathbb{\Lambda}$ that depends only on ρ and $\{M_a\}_a$.

In Sec. 5.5 we show furthermore that $p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A})$ in fact only depends on $\mathbb{A}$ through $\mathbb{\Lambda}$. The above reveals that the robustness of teleportation fits into the program of robustness-based quantifiers and discrimination tasks, where the specific restrictions are on the resource state and resource measurement used to play the game.

5.3.3 Complete sets of monotones for teleportation simulation

The average score (5.13) and average guessing probability (5.18) are also important as they provide complete characterisations for the two notions of teleportation simulation introduced in (5.11) and (5.12). In particular, in Sec. 5.5 we show the following:

Complete sets of monotones for quantum and classical simulation

Result 5.3. Let $\mathbb{\Lambda}$ and $\mathbb{\Lambda}'$ be two teleportation instruments. Then $\mathbb{\Lambda}$ can quantum-simulate $\mathbb{\Lambda}'$, $\mathbb{\Lambda} \succ_q \mathbb{\Lambda}'$, if and only if

$$q(\mathcal{G}, \mathbb{\Lambda}) \geq q(\mathcal{G}, \mathbb{\Lambda}') \text{ for all games } \mathcal{G} \tag{5.22}$$

Similarly, $\mathbb{\Lambda}$ can classically simulate $\mathbb{\Lambda}'$, $\mathbb{\Lambda} \succ_c \mathbb{\Lambda}'$ if and only if:

$$p_{\text{succ}}(\mathbb{E}, \mathbb{\Lambda}) \geq p_{\text{succ}}(\mathbb{E}, \mathbb{\Lambda}') \text{ for all games } \mathbb{E}. \tag{5.23}$$

This means that both $q(\mathcal{G}, \mathbb{\Lambda})$ and $p_{\text{succ}}(\mathbb{E}, \mathbb{\Lambda})$ constitute "complete set of monotones", the former for the partial order of quantum-simulation, and the latter for classical-simulation. Both of these complete sets of monotones can be viewed as "second laws" of the resource theory of LOSR when specified to teleportation instruments.

5.4 Discussion and open problems

We have analysed a robustness-based quantifier of teleportation and shown that it has operational significance in two unrelated directions. On the one hand it quantifies the advantage that a given teleportation instrument offers for the task of teleporting quantum correlations. On the other hand, it also quantifies the advantage offered by a fixed entangled state and fixed entangled measurement in the task of subchannel discrimination with side information.

We showed that the first task is a natural generalisation of the standard task used for benchmarking the quality of a teleportation set-up (the average fidelity of teleportation), and thus provides an answer to the question of in what sense is every state useful for teleportation: Every state has the ability to teleport quantum correlations strictly better than can be achieved by any classical teleportation scheme.

We finally showed that the two tasks which give operational meaning to the robustness of teleportation also form complete sets of monotones, which fully characterise two natural notions of simulation that arise for teleportation, one purely classical, and the other quantum.

We believe that these results highlight an interesting new direction of studying composite resource theories, that is QRTs where the main object of interest consists of resourceful objects of differing types. This could give rise to interesting relationships and trade-offs between not only different types of resources, but also due to differing types of objects. In this work we have seen that an object composed of a bipartite measurement and bipartite state can be seen as a resource for entanglement-assisted subchannel discrimination. It would be interesting to see if one can characterise operationally such composite resources in full generality. Some work in this direction has been already done e.g. in (Schmid, Rosset, and Buscemi, 2020) and (Ducuara, Lipka-Bartosik, and Skrzypczyk, 2020b).

5.5 Proofs

5.5.1 Reformulating robustness of teleportation

Let us start with the primal formulation of the optimization problem (5.7). By multiplying both sides of the first constraint by $1+r$, labelling $\eta_{a|\omega_x} = r\,\Lambda'_a(\omega_x)$, $\xi_{a|\omega_x} = (1+r)\,\Lambda^c_a(\omega_x)$ and $\widetilde{\eta} = r\,\eta$ we can turn the original problem into:

$$\begin{aligned} \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}, \{\omega_x\}) = \quad \min \quad & r \\ & \sigma_{a|\omega_x} + \eta_{a|\omega_x} = \xi_{a|\omega_x} \\ & \sum_a \eta_{a|\omega_x} = \widetilde{\eta}, \quad \operatorname{tr}\widetilde{\eta} = r \\ & \xi_{a|\omega_x} = (1+r)\,\Lambda^c_a(\omega_x), \quad \Lambda^c_a \in \mathcal{F}_{\mathrm{T}} \\ & \eta_{a|\omega_x} = r\,\Lambda'_a(\omega_x), \quad \Lambda'_a \in \mathcal{R}_{\mathrm{T}} \end{aligned} \tag{5.24}$$

Let us now characterise the set of unnormalized states $\xi_{a|\omega_x}$. Writing explicitly we have:

$$\xi_{a|\omega_x} = \operatorname{tr}_{VA}\left[(O^{VA}_a \otimes \mathbb{1}^B)\cdot(\mathbb{1}^V \otimes \widetilde{\xi}^{AB})\right] \tag{5.25}$$

$$= \operatorname{tr}_V\left[O^{VB}_a \cdot (\omega^V_x \otimes \mathbb{1}^B)\right], \tag{5.26}$$

where $O^{VB}_a = \operatorname{tr}_A[(O^{VA}_a \otimes \mathbb{1}^B)\cdot(\mathbb{1}^V \otimes \widetilde{\xi}^{AB})]$ is a positive operator and $\widetilde{\xi}^{AB} = (1+r)\,\xi^{AB}$ for some state ξ. Our goal is to ultimately rewrite the optimization problem (5.24) using operators of this type. Since the data (5.25) corresponds to a classical teleportation instrument, the state $\xi^{AB} = \sum_\lambda p(\lambda)\xi^A_\lambda \otimes \xi^B_\lambda$ is separable. This means that we can rewrite O^{VB}_a as:

$$O^{VB}_a = (1+r)\sum_\lambda p(\lambda)\operatorname{tr}_A[(O^{VA}_a \otimes \mathbb{1}^B)\cdot(\mathbb{1}^V \otimes \xi^A_\lambda \otimes \xi^B_\lambda)] \tag{5.27}$$

$$= (1+r)\sum_\lambda p(\lambda)\, O^V_{a|\lambda} \otimes \xi^B_\lambda, \tag{5.28}$$

where $O^V_{a|\lambda} = \operatorname{tr}_A[O^{VA}_a(\mathbb{1}^V \otimes \xi^A_\lambda)]$. This means that the operators O^{VB}_a have the form $O^{VB}_a = \sum_\lambda X_\lambda \otimes Y_\lambda$ for some positive operators X_λ, Y_λ and thus the unnormalized states $\xi_{a|\omega_x}$ are also separable. We now focus on the structure of $\eta_{a|\omega_x}$. We have:

$$\eta_{a|\omega_x} = \operatorname{tr}_{VA}\left[(N^{VA}_a \otimes \mathbb{1}^B)\cdot(\mathbb{1}^V \otimes \widetilde{\eta}^{AB})\right] \tag{5.29}$$

$$= \operatorname{tr}_V\left[N^{VB}_a \cdot (\omega^V_x \otimes \mathbb{1}^B)\right] \tag{5.30}$$

where $N^{VB}_a = \operatorname{tr}_A[(N^{VA}_a \otimes \mathbb{1}^B)\cdot(\mathbb{1}^V \otimes \widetilde{\eta}^{AB})]$, N^{VA}_a is an arbitrary bipartite measurement and $\widetilde{\eta}^{AB}$ is an arbitrary unnormalized state. Denoting partial transpose with respect to system V with T_V we can verify that N^{VB}_a have a positive partial transpose:

$$(N^{VB}_a)^{T_V} = \operatorname{tr}_A\left[(N^{VA}_a \otimes \mathbb{1}^B)(\mathbb{1}^V \otimes \widetilde{\eta}^{AB})\right]^{T_V} \tag{5.31}$$

$$= d_V \cdot \operatorname{tr}_{VA}\left[(\mathbb{1}^{V'} \otimes N^{VA}_a \otimes \mathbb{1}^B)(\phi^{V'V}_+ \otimes \widetilde{\eta}^{AB})\right] \geq 0, \tag{5.32}$$

as this is a product of positive operators. This is not only a necessary, but also a sufficient condition, i.e. any operator which has a positive partial transpose can be written in the form of N^{VB}_a. To see this, consider an arbitrary X^{VB} such that

$(X^{VB})^{T_V} \geq 0$ and take $N_a^{VA} = \phi_+^{VA}$ and $\widetilde{\eta}_a^{AB} = d_V \cdot (X^{AB})^{T_A}$ in the definition of N_a^{VB} below (5.30). It is straightforward to verify that:

$$N_a^{VB} = d_V \cdot \mathrm{tr}_A[(\phi_+^{VA} \otimes \mathbb{1}^B)(\mathbb{1}^V \otimes (X^{AB})^{T_A})] = X^{VB}, \tag{5.33}$$

where we used the identity: $\mathrm{tr}_A[(\phi_+^{VA} \otimes \mathbb{1}^B)(\mathbb{1}^V \otimes \widetilde{\eta}^{AB})] = d_V^{-1}(\widetilde{\eta}^{VB})^{T_V}$. Having characterized operators O_a^{VB} and N_a^{VB} we can rewrite the observed data $\sigma_{a|\omega_x}$ in an analogous form, i.e. $\sigma_{a|\omega_x}^B = \mathrm{tr}_V[M_a^{VB}(\omega_x^V \otimes \mathbb{1}^B)]$, where $M_a^{VB} = \mathrm{tr}_A[(M_a^{VA} \otimes \mathbb{1}^B)(\mathbb{1}^V \otimes \rho^{AB})]$. We also assume that the states $\{\omega_x\}$ form a tomographically-complete set, i.e. they form an operator basis. This implies that the constraint $\mathrm{tr}_V[(M_a^{VB} + N_a^{VB} - O_a^{VB})(\omega_x^V \otimes \mathbb{1}^B)] = 0$ becomes equivalent to $M_a^{VB} + N_a^{VB} - O_a^{VB} = 0$ and makes the optimization problem independent of $\{\omega_x\}$. To emphasize this we write $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ instead of $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}, \{\omega_x\})$. Combining this with our previous realizations allows (5.7) to be written in the equivalent form, which is now manifestly a semi-definite program:

$$\begin{aligned} \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}) = \min_{\{N_a^{VB}\}_a, \{O_a^{VB}\}_a, \widetilde{\eta}^B} \quad & r \\ & M_a^{VB} + N_a^{VB} = O_a^{VB} \\ & \sum_a N_a^{VB} = \mathbb{1}^V \otimes \widetilde{\eta}^B \\ & \mathrm{tr}\,\widetilde{\eta}^B = r, \quad O_a^{VB} \in \mathsf{SEP}(A:B), \quad \left(N_a^{VB}\right)^{T_V} \geq 0. \end{aligned} \tag{5.34}$$

This can be further simplified if we notice the following identity:

$$\begin{aligned} (M_a^{VB})^{T_V} &= \mathrm{tr}_A\left[((M_a^{VA})^{T_V} \otimes \mathbb{1}^B)(\mathbb{1}^V \otimes \rho^{AB})\right] && (5.35)\\ &= d_V \cdot \mathrm{tr}_{VA}\left[(\mathbb{1}^{V'} \otimes M_a^{VA} \otimes \mathbb{1}^B)(\phi_+^{V'V} \otimes \rho^{AB})\right] && (5.36)\\ &= d_V \cdot J_a^{VB} && (5.37) \end{aligned}$$

where J_a is a Choi matrix associated to the subchannel $\Lambda_a^{V' \to B}[\cdot]$:

$$J_a^{VB} = (\mathrm{id}^V \otimes \Lambda_a^{V' \to B})\,\phi_+^{VV'}. \tag{5.38}$$

Since J_a are positive operators, we can introduce a family of unnormalized states $\widetilde{\rho}_a^{VB} := J_a^{VB}$ such that $\rho_a^{VB} = J_a^{VB}/p_{\mathrm{T}}(a)$ with $p_{\mathrm{T}}(a) := \mathrm{tr}\left[J_a^{VB}\right]$ are states. Notice also that by the definition of teleportation instrument $\{\Lambda_a^{V' \to B}\}$ (see also (5.3)) we have

$$\sum_a J_a^{VB} = d_V^{-1} \cdot \mathbb{1}^V \otimes \rho^B. \tag{5.39}$$

Taking partial transpose with respect to subsystem (V) of the first constraint in (5.34) and using the fact that this preserves separability allows to obtain a simplified form

of the primal problem:

$$\mathbf{R}_{\mathrm{T}}(\Bbbk) = \min_{\{O_a^{VB}\}, Z^B} \quad \operatorname{tr} Z^B - 1 \tag{5.40}$$

$$O_a^{VB} \geq d_V \cdot \widetilde{\rho}_a^{VB} \qquad A_a^{VB} \tag{5.41}$$

$$\sum_a O_a^{VB} = \mathbb{1}^V \otimes Z^B, \quad O_a^{VB} \in \mathsf{SEP}(A:B), \qquad B^{VB},\ W_a^{VB}. \tag{5.42}$$

Notice that in our case strong duality holds since we can always find feasible $O_a^{VB} = \alpha \cdot \mathbb{1}^{VB}$ and $Z^B = \alpha \cdot \sum_a \mathbb{1}^B$ for some $\alpha \geq 0$. We now look at the dual formulation of the above problem. To do so we first write the associated Lagrangian using the dual variables associated with each set of constraints: $\{A_a^{VB}\}_a$ with $A_a^{VB} \geq 0$, B^{VB} and $\{W_a^{VB}\}_a$ with $\{W_a^{VB}\} \in \mathcal{R}_{\mathrm{T}}^*$ (displayed above on the right-hand side in grey), that is:

$$\mathcal{L} = \operatorname{tr} Z^B - 1 - \sum_a \operatorname{tr} A_a^{VB} \left[O_a^{VB} - d_V \cdot J_a^{VB}\right] + \operatorname{tr} B^{VB} \left[\sum_a O_a^{VB} - \mathbb{1}^V \otimes Z^B\right] \tag{5.43}$$

$$- \sum_a \operatorname{tr}\left[W_a^{VB} O_a^{VB}\right] \tag{5.44}$$

$$= \sum_a \operatorname{tr} O_a^{VB} \left[-A_a^{VB} + B^{VB} - W_a^{VB}\right] + \operatorname{tr} Z^B \left[\mathbb{1}^B - B^B\right] \tag{5.45}$$

$$+ d_V \sum_a \operatorname{tr}\left[A_a^{VB} J_a^{VB}\right] - 1. \tag{5.46}$$

By demanding that the terms in the square brackets which appear with the primal variables in the last line vanish we can ensure $\mathcal{L} \leq r$. This leads to the following (dual) semi-definite program:

$$\mathbf{R}_{\mathrm{T}}(\Bbbk) = \max_{\{A_a^{VB}\}_a, B^{VB}} \quad d_V \cdot \sum_a \operatorname{tr}\left[A_a^{VB} J_a^{VB}\right] - 1 \tag{5.47}$$

$$W_a = B^{VB} - A_a^{VB}, \{W_a\} \in \mathcal{R}_{\mathrm{T}}^*$$

$$B^B = \mathbb{1}^B, \quad A_a^{VB} \geq 0.$$

Let us now return to the primal formulation of the robustness problem (5.40) and let $O_a^{VB} = O_a^*$ and $Z^B = Z^*$ be the optimal choice of primal variables. Notice that $1 + \mathbf{R}_{\mathrm{T}}(\Bbbk) = d_V^{-1} \cdot \sum_a \operatorname{tr} O_a^*$, where O_a^* is a separable operator. Denoting $\sigma_a^* := O_a^* / \operatorname{tr} O_a^*$ and $p_S(a) := \left[d_V\,(1 + \mathbf{R}_{\mathrm{T}})\right]^{-1} \cdot \operatorname{tr} O_a^*$ we can write:

$$\widetilde{\rho}_a \leq \frac{1}{d_V} O_a^* = \frac{1}{d_V} \operatorname{tr} E_a^* \cdot \sigma_a^* = (1 + \mathbf{R}_{\mathrm{T}})\, p_S(a) \cdot \sigma_a^*, \tag{5.48}$$

where σ_a^* is some separable state and $p_S(a)$ forms a probability distribution. We will use both the dual (5.47) and primal (5.40) forms to prove some of the results described in the main text.

5.5.2 Properties of the robustness of teleportation

In this section we prove the basic three properties of robustness of teleportation which we used in the previous sections.

Faithfulness If a teleportation instrument is classical, that is $\mathbb{\Lambda} \in \mathcal{F}_{\mathrm{T}}$, then we can always choose a feasible $r = 0$ in the defining optimization problem (5.7). Since $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ is non-negative, then $r = 0$ is also optimal.

Convexity Let $\{\eta^1_{a|\omega_x}, \xi^1_{a|\omega_x}\}$ be the optimal primal variables for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_1)$ and similarly for $\{\eta^2_{a|\omega_x}, \xi^2_{a|\omega_x}\}$ and $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_2)$. Let $\mathbb{\Lambda}' = \{\Lambda'_a\}_a$ be a convex mixture of the two teleportation instruments, that is $\Lambda'_a(\cdot) = p\,\Lambda^1_a(\cdot) + (1-p)\,\Lambda^2_a(\cdot)$ for each a. We can construct (potentially sub-optimal) solutions for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')$ using: $\eta'_{a|\omega_x} = p\,\eta^1_{a|\omega_x} + (1-p)\,\eta^2_{a|\omega_x}$ and $\xi'_{a|\omega_x} = p\,\xi^1_{a|\omega_x} + (1-p)\,\xi^2_{a|\omega_x}$. Substituting $\eta'_{a|\omega_x}$ and $\xi'_{a|\omega_x}$ into the constraints of problem (5.7) for $\mathbb{\Lambda}'$ shows that this choice is feasible. This leads to the upper bound on $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')$:

$$\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}') \leq \mathrm{tr} \sum_a \eta'_{a|\omega_x} = p\ \mathrm{tr} \sum_a \eta^1_{a|\omega_x} + (1-p)\ \mathrm{tr} \sum_a \eta^2_{a|\omega_x} \tag{5.49}$$

$$= p\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_1) + (1-p)\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}_2). \tag{5.50}$$

Monotonicity Let us start with quantum simulation. Assume that $\mathbb{\Lambda}$ can simulate $\mathbb{\Lambda}'$, i.e. $\mathbb{\Lambda} \succ_q \mathbb{\Lambda}'$. This means that there exists a collection of channels Θ_λ, Ω_λ and probability distributions $p(\lambda)$ and $p(b|a,\lambda)$ such that for all b:

$$\Lambda'_b(\cdot) = \sum_{a,\lambda} p(\lambda) p(b|a,\lambda) \Theta_\lambda \circ \Lambda'_b \circ \Omega_\lambda(\cdot) \tag{5.51}$$

Suppose now that we solved the dual problem (5.47) for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')$ using optimal dual variables B' and A'_b. Using these we can construct an educated guess for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')$ in the following way:

$$B^* = \sum_{b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \left[(\Omega^T_\lambda)^\dagger \otimes \Theta^\dagger_\lambda\right] B, \tag{5.52}$$

$$A^*_a = \sum_{b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \left[(\Omega^T_\lambda)^\dagger \otimes \Theta^\dagger_\lambda\right] A'_b. \tag{5.53}$$

Using these we can find the following lower bound:

$$
\begin{aligned}
1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}) &\geq d_V \cdot \sum_a \mathrm{tr}\left[\widetilde{\rho}_a A_a^*\right] && (5.54)\\
&= \sum_{a,b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[(\mathrm{id}\otimes\Lambda_a)\,\phi_+ \cdot \left[(\Omega_\lambda^T)^\dagger \otimes \Theta_\lambda^\dagger\right] A_b'\right] && (5.55)\\
&= \sum_{a,b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[(\Omega_\lambda^T\otimes\Lambda_a)\,\phi_+ \cdot (\mathrm{id}\otimes\Theta_\lambda^\dagger)\, A_b'\right] && (5.56)\\
&= \sum_{a,b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[(\mathrm{id}\otimes\Lambda_a\circ\Omega_\lambda)\phi_+ \cdot (\mathrm{id}\otimes\Theta_\lambda^\dagger)\, A_b'\right] && (5.57)\\
&= \sum_b \mathrm{tr}\left[(\mathrm{id}\otimes\Lambda_b')\,\phi_+ \cdot A_b'\right] && (5.58)\\
&= 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}'). && (5.59)
\end{aligned}
$$

Let us now show that the choice (5.52) is feasible. By construction we have $B^* \geq 0$, $A_a^* \geq 0$ and $\mathrm{tr}_V\, B^* = \mathbb{1}$, since:

$$
\begin{aligned}
\mathrm{tr}_V\left[\left((\Omega_\lambda^T)^\dagger\otimes\Theta_\lambda^\dagger\right) B\right] &= \mathrm{tr}_V\left[\left((\Omega_\lambda^T)^\dagger\otimes\Theta_\lambda^\dagger\circ\mathcal{B}^\dagger\right)\phi_+\right] && (5.60)\\
&= \mathrm{tr}_V\left[\left(\mathrm{id}\otimes\Theta_\lambda^\dagger\circ\mathcal{B}^\dagger\circ\Omega_\lambda^\dagger\right)\phi_+\right] && (5.61)\\
&= \Theta_\lambda^\dagger\circ\mathcal{B}^\dagger\circ\Omega_\lambda^\dagger(\mathbb{1}) && (5.62)\\
&= \mathbb{1}, && (5.63)
\end{aligned}
$$

where in the first line we used the Choi-Jamiołkowski isomorphism $B = (\mathrm{id}\otimes\mathcal{B})\phi_+$ for some map $\mathcal{B} \in \mathsf{CPTP}$ and in the third line we used the fact that the adjoint of a CPTP map is unital. It remains to show that $B^* - A_a^*$ is an entanglement witness. Let ρ_S be an arbitrary separable state. We have:

$$
\begin{aligned}
\mathrm{tr}\left[(B^* - A_a^*)\rho_S\right] &= \sum_{b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[\left((\Omega_\lambda^T)^\dagger\otimes\Theta_\lambda^\dagger\right)(B - A_b')\cdot\rho_S\right] && (5.64)\\
&= \sum_{b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[(B - A_b')\cdot\left(\Omega_\lambda^T\otimes\Theta_\lambda\right)\rho_S\right] && (5.65)\\
&= \sum_{b,\lambda} p(\lambda)\, p(b|a,\lambda)\ \mathrm{tr}\left[(B - A_b')\cdot\rho_\lambda'\right] && (5.66)\\
&\geq 0, && (5.67)
\end{aligned}
$$

where we used the fact that $W_b = B - A_b'$ is by assumption an entanglement witness and $\rho_\lambda' = \left(\Omega_\lambda^T\otimes\Theta_\lambda\right)\rho_S$ is a separable operator. To show analogous statement about classical simulation is simple, as this is just a special case of quantum simulation resulting from choosing $p(\lambda) = 1/o_\lambda$, where o_λ size of the alphabet associated with λ, $\Theta_\lambda = \Omega_\lambda = \mathrm{id}$.

5.5.3 RoT as an advantage in the teleportation of quantum correlations

Here we prove that the robustness of teleportation $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ can be viewed as the best advantage in the task of teleporting quantum correlations using a fixed quantum teleportation instrument $\mathbb{\Lambda}$ over any classical teleportation instrument. We start by constructing a particular game $\mathcal{G}^*$ using the dual formulation of the RoT and then show that $1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ gives a meaningful lower bound on the advantage. We then use

primal formulation (5.40) and show that $1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})$ also bounds the advantage from above.

Suppose we have solved the dual problem for the RoT as given by (5.47) using dual variables B and A_a. We can construct a (potentially sub-optimal) task $\mathcal{G}^* = \{\psi^*, \xi_a^*, f^*(a)\}$ using these optimal variables in the following way:

$$\psi^* = \phi_+, \quad \xi_a^* = \frac{A_a}{\operatorname{tr} A_a}, \quad f^*(a) = \operatorname{tr} A_a. \tag{5.68}$$

The maximal average score which can be achieved using classical teleportation instruments $\mathbb{\Lambda}^c$ in game $\mathcal{G}^*$ can be bounded by:

$$\begin{aligned}
q^c(\mathcal{G}^*) = \max_{\mathbb{\Lambda}^c \in \mathcal{F}_{\mathrm{T}}} q(\mathcal{G}^*, \mathbb{\Lambda}^c) &= \max_{\mathbb{\Lambda}^c \in \mathcal{F}_{\mathrm{T}}} \max_{\mathcal{E}_a \in \mathsf{CPTP}} \sum_a f^*(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda_a^c)\, \psi^* \cdot \xi_a^*\right] && (5.69)\\
&= \max_{\mathbb{\Lambda}^c \in \mathcal{F}_{\mathrm{T}}} \max_{\mathcal{E}_a \in \mathsf{CPTP}} \sum_a \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda_a^c)\, \phi_+ \cdot A_a\right] && (5.70)\\
&= \max_{\widetilde{\rho}_a \in \mathsf{SEP}(A:B)} \max_{\mathcal{E}_a \in \mathsf{CPTP}} \sum_a \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a)\widetilde{\rho}_a \cdot (B - W_a)\right] && (5.71)\\
&\leq \max_{\widetilde{\rho}_a \in \mathsf{SEP}(A:B)} \sum_a \operatorname{tr}\left[\widetilde{\rho}_a \cdot B\right] && (5.72)\\
&= \frac{1}{d_V} \max_{\rho^B} \operatorname{tr}\left[\rho^B \cdot B^B\right] && (5.73)\\
&\leq \frac{1}{d_V}, && (5.74)
\end{aligned}$$

where in the second line we used $\widetilde{\rho}_a = (\mathrm{id} \otimes \Lambda_a^c)\, \phi_+ \in \mathsf{SEP}(A:B)$, in the third line we used the constraint from the dual: $B - A_a = W_a \in \mathsf{SEP}(A:B)^*$, in the fourth line we employed the fact that W_a is an entanglement witness and finally we used $\sum_a \widetilde{\rho}_a = d_V^{-1} \cdot \mathbb{1}^V \otimes \rho^B$. Notice now that for an arbitrary teleportation instrument $\mathbb{\Lambda}$ we can write:

$$\max_{\mathcal{G}} \frac{q(\mathcal{G}, \mathbb{\Lambda})}{q^c(\mathcal{G})} \geq \frac{q(\mathcal{G}^*, \mathbb{\Lambda})}{q^c(\mathcal{G}^*)} = d_V \sum_a \operatorname{tr}\left[\widetilde{\rho}_a \cdot A_a\right] = 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}). \tag{5.75}$$

To prove the reverse direction let us look at the average score for an arbitrary game $\mathcal{G} = \{\psi, \xi_a, f(a)\}$:

$$\begin{aligned}
q(\mathcal{G}, \mathbb{\Lambda}) &= \max_{\substack{\mathbb{\Lambda}' \leq \mathbb{\Lambda} \\ \mathcal{E}_a \in \mathsf{CPTP}}} \sum_a f(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda_a')\, \psi \cdot \xi_a\right] && (5.76)\\
&= \max_{\substack{\mathbb{\Lambda}' \leq \mathbb{\Lambda} \\ \mathcal{E}_a \in \mathsf{CPTP}}} \sum_a f(a) \operatorname{tr}\left[(\mathcal{N} \otimes \mathcal{E}_a)\ \widetilde{\rho}_a' \cdot \xi_a\right] && (5.77)\\
&\leq \max_{\mathbb{\Lambda}' \leq \mathbb{\Lambda}} \left[1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')\right] \max_{\sigma_a' \in \mathsf{SEP}(A:B)} \sum_a f(a) \operatorname{tr}\left[(\mathcal{N} \otimes \mathrm{id})\ \sigma_a' \cdot \xi_a\right] && (5.78)\\
&\leq \max_{\mathbb{\Lambda}' \leq \mathbb{\Lambda}} \left[1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}')\right]\, q^c(\mathcal{G}) && (5.79)\\
&\leq \left[1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})\right] q^c(\mathcal{G}), && (5.80)
\end{aligned}$$

where in the second line we used $\psi = (\mathcal{N} \otimes \mathrm{id})\, \phi_+$ for some (possibly trace non-increasing) map $\mathcal{N} \in \mathtt{CPTNI}$ and in the third line we used (5.48). The last inequality follows from monotonicity of the RoT. Note that the above reasoning is valid for any

game $\mathcal{G}$ and thus by taking the maximum over all $\mathcal{G}$ we obtain:

$$\max_{\mathcal{G}} \frac{q(\mathcal{G}, \mathbb{A})}{q^c(\mathcal{G})} \leq 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A}). \tag{5.81}$$

Combined with the lower bound, this proves the equality.

5.5.4 RoT as an advantage in entanglement-assisted subchannel discrimination

Let $\mathbb{E} = \{\mathcal{E}_x\}$ be an instrument, such that $\sum_x \mathcal{E}_x(\cdot) = \mathcal{E}(\cdot)$ is a channel, and let $\mathbb{A} = \{\{M_a\}_a, \rho\}$ be a resource used in the game, consisting of a bipartite measurement $\{M_a\}_a \in \mathrm{POVM}$ and a bipartite state ρ. The average probability of guessing which subchannel from the instrument $\mathbb{E}$ was applied to the state ρ is given by:

$$p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}) = \max_{p(g|a)} \sum_{x,a,g} p(g|a)\ \mathrm{tr}\left[M_a \cdot (\mathcal{E}_x \otimes \mathrm{id})\rho\right]\ \delta_{g,x}. \tag{5.82}$$

In what follows we will use the following operator identity:

$$d_V^2\ \mathrm{tr}\left[X^{VB}\,\phi_+^{VB}\right] = \mathrm{tr}\left[M^{AB} \cdot (\mathcal{E}^A \otimes \mathrm{id}^B)\rho^{AB})\right], \tag{5.83}$$

where $X^{VB} = \mathrm{tr}_{V'A}\left[\left(\mathbb{1}^{VB} \otimes M^{V'A}\right)\left((\mathrm{id}^V \otimes \mathcal{E}^{V'})\,\phi_+^{VV'} \otimes \rho^{AB}\right)\right]$ and $\mathcal{E}$ is an arbitrary map. The above identity can be proven by direct substitution. Using identity (5.83) and the fact that $\widetilde{\rho}_a^{VB} = (\mathrm{id}^V \otimes \Lambda_a^{V'\to B})\,\phi_+^{VV'}$ leads to:

$$\begin{aligned}
&d_V^2\,\mathrm{tr}\left[(\mathrm{id}^V \otimes \mathcal{E}_x^B)\widetilde{\rho}_a^{VB} \cdot \phi_+^{VB}\right] && (5.84)\\
&= \mathrm{tr}\left[\left(\mathrm{id}^{VV'A} \otimes \mathcal{E}_x^B\right)\left(\mathbb{1}^{VB} \otimes M_a^{V'A}\right)\left(\phi_+^{VV'} \otimes \rho^{AB}\right) \cdot \left(\mathbb{1}^{VA} \otimes \phi_+^{V'B}\right)\right] && (5.85)\\
&= \mathrm{tr}\left[\left(\mathbb{1}^{VB} \otimes M_a^{V'A}\right)\left((\mathrm{id}^V \otimes \mathcal{E}_x^{V'})\phi_+^{VV'} \otimes \rho^{AB}\right)\left(\mathbb{1}^{V'A} \otimes \phi_+^{VB}\right)\right] && (5.86)\\
&= \mathrm{tr}\left[M_a^{AB} \cdot (\mathcal{E}_x^A \otimes \mathrm{id}^B)\hat{\rho}^{AB}\right] && (5.87)
\end{aligned}$$

In this way we can rewrite (5.82) as:

$$\begin{aligned}
p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}) &= d_V^2 \cdot \max_{p(g|a)} \sum_{x,a,g} p(g|a)\ \mathrm{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\widetilde{\rho}_a \cdot \phi_+\right]\ \delta_{g,x} && (5.88)\\
&= d_V^2 \cdot \max_{p(x|a)} \sum_{x,a} p(x|a)\ \mathrm{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\widetilde{\rho}_a \cdot \phi_+\right] && (5.89)
\end{aligned}$$

Suppose now that we solved the dual problem for the RoT as given by (5.47) using dual variables B^* and A_x^*. We will now construct a sequence of games $\mathbb{E}^* = \{\mathcal{E}_x^*\}$, parametrized with N, using these optimal variables. This proof technique is inspired by the methods used in (Piani and Watrous, 2015). Let us define a set of subchannels via their duals, i.e:

$$(\mathcal{E}_x^*)^\dagger(\rho) = \begin{cases} \alpha\ \mathrm{tr}_V\left[(\rho^T \otimes \mathbb{1})\,A_x^*\right] & \text{for} \quad 1 \leq x \leq o_a, \\ \dfrac{1}{N \cdot d_V}\left[\mathbb{1} - \alpha \displaystyle\sum_{x'=1}^{o_a} A_{x'}^B\right]\mathrm{tr}(\rho) & \text{for} \quad o_a + 1 \leq x \leq o_a + N. \end{cases} \tag{5.90}$$

In the above $\alpha = \left\| \sum_{x'=1}^{o_a} A^B_{x'} \right\|_\infty^{-1}$ is a real parameter chosen such that the map defined above is completely positive. Notice that the constraints of the dual problem (5.47) imply that $0 \leq A_{x'} \leq \mathbb{1}$. To verify that $\mathcal{E}^* = \sum_x \mathcal{E}^*_x$ defines a channel recall that $\mathcal{E} \in \mathsf{CPTP}$ if and only if its dual map $\mathcal{E}^\dagger$ is unital. By construction we have:

$$\sum_{x=1}^{o_a+N} (\mathcal{E}^*_x)^\dagger(\mathbb{1}) = \alpha \sum_{x=1}^{o_a} \mathrm{tr}_V \left[A^*_x\right] + \mathbb{1} - \alpha \sum_{x=1}^{o_a} \mathrm{tr}_V \left[A^*_x\right] = \mathbb{1}. \tag{5.91}$$

Notice that by our particular definition of the instrument $\mathcal{G}^*$ we also have the following relation:

$$\left[\mathrm{id} \otimes (\mathcal{E}^*_x)^\dagger\right] \phi_+ = \begin{cases} \dfrac{\alpha}{d_V} A_x & \text{for} \quad 1 \leq x \leq o_a, \\ \dfrac{1}{N \cdot d_V^2} \mathbb{1} \otimes \left(\mathbb{1} - \alpha \displaystyle\sum_{x'=1}^{o_a} A^B_x \right) & \text{for} \quad o_a + 1 \leq x \leq o_a + N. \end{cases} \tag{5.92}$$

Let us now upper bound the maximal probability of guessing in a game specified by $\mathcal{G}^*$ and when having access only to classical resources. This is specified by $p^c_{\mathrm{succ}}(\mathbb{E}^*) = \max_{\mathbb{A}^c \in \mathcal{F}_\mathrm{T}} p_{\mathrm{succ}}(\mathbb{E}^*, \mathbb{A}^c)$, where the optimization is performed over all $\mathbb{A}^c = \{\{M_a\}_a, \sigma\}$ with $\sigma \in \mathsf{SEP}(A:B)$ and arbitrary measurements $\{M_a\}_a$. Using (5.89) this becomes:

$$p^c_{\mathrm{succ}}(\mathbb{E}^*) = d_V^2 \max_{\sigma_a \in \mathsf{SEP}(A:B),\, p_\mathrm{T}(a)} \max_{p(x|a)} \sum_{x,a} p(x|a)\, p_\mathrm{T}(a)\, \mathrm{tr}\left[(\mathrm{id} \otimes \mathcal{E}^*_x)\sigma_a \cdot \phi_+\right] \tag{5.93}$$

$$= d_V^2 \max_{\sigma_a \in \mathsf{SEP}(A:B), p_\mathrm{T}(a)} \max_{p(x|a)} \sum_a p_\mathrm{T}(a) \left[\frac{\alpha}{d_V} \sum_{x=1}^{o_a} p(x|a)\, \mathrm{tr}[\sigma_a A_x] \right. \tag{5.94}$$

$$\left. + \frac{1}{N d_V^2} \sum_{x=o_a+1}^{o_a+N} p(x|a)\, \mathrm{tr}\left[\sigma_a - \alpha \left(\mathbb{1} \otimes \sum_{x'=1}^{o_a} A^B_{x'} \right) \sigma_a \right] \right] \tag{5.95}$$

$$\leq d_V \max_{\sigma_a \in \mathsf{SEP}(A:B), p_\mathrm{T}(a)} \max_{p(x|a)} \sum_a p_\mathrm{T}(a) \left[\alpha \sum_{x=1}^{o_a} p(x|a) \alpha\, \mathrm{tr}[\sigma_a A_x] \right. \tag{5.96}$$

$$\left. + \frac{1}{N \cdot d_V} \sum_{x=o_a+1}^{o_a+N} p(x|a)\, \mathrm{tr}[\sigma_a] \right] \tag{5.97}$$

$$\leq d_V \max_{\sigma_a \in \mathsf{SEP}(A:B), p_\mathrm{T}(a)} \max_{p(x|a)} \sum_a p_\mathrm{T}(a) \left[\alpha \sum_{x=1}^{o_a} p(x|a)\, \mathrm{tr}[\sigma_a A_x] \right] + \frac{1}{N}.$$

In the first line we labelled $p_\mathrm{T}(a) = \mathrm{tr}\, \widetilde{\sigma}_a$ to be the probability of an outcome a in the (classical) teleportation instrument and in the third line we used the fact that subchannels corresponding to fictitious outcomes $o_a + 1 \leq x \leq o_a + N$ are positive. Recall that the operators A_x must satisfy certain constraints in order to be feasible solutions of the dual problem (5.47). In particular, $A_x = B - W_x$, where B is a positive matrix with $\mathrm{tr}_V B = B^B = \mathbb{1}$ and $W_x \in \mathsf{SEP}(A:B)^*$ is an entanglement witness. This

allows for the following bound to be obtained:

$$\sum_a p_{\mathrm{T}}(a) \sum_{x=1}^{o_a} p(x|a) \operatorname{tr}[\sigma_a A_x] = \sum_a p_{\mathrm{T}}(a) \sum_{x=1}^{o_a} p(x|a) \operatorname{tr}[\sigma_a (B - W_x)] \tag{5.98}$$

$$\leq \sum_a p_{\mathrm{T}}(a) \sum_{x=1}^{o_a} p(x|a) \operatorname{tr}[\sigma_a B] \tag{5.99}$$

$$\leq \sum_a p_{\mathrm{T}}(a) \operatorname{tr}[\sigma_a B] \tag{5.100}$$

$$= \frac{1}{d_V} \operatorname{tr}\left[(\mathbb{1} \otimes \sigma^B) B\right] \tag{5.101}$$

$$= \frac{1}{d_V} \operatorname{tr}\left[\sigma^B B^B\right] \tag{5.102}$$

$$= \frac{1}{d_V}. \tag{5.103}$$

This in turn leads to a bound on the classical probability of guessing (5.93):

$$p^c_{\mathrm{succ}}(\mathbb{E}^*) \leq \alpha + \frac{1}{N}. \tag{5.104}$$

Let us now bound the average probability of guessing in game $\mathbb{E}^*$ when having access to a resource $\mathbb{A}$. We have:

$$p_{\mathrm{succ}}(\mathbb{E}^*, \mathbb{A}) = d_V^2 \cdot \max_{p(x|a)} \sum_{x,a} p(x|a)\, p_{\mathrm{T}}(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}^*_x)\rho_a \cdot \phi_+\right] \tag{5.105}$$

$$\geq \alpha\, d_V \cdot \sum_a p_{\mathrm{T}}(a) \operatorname{tr}\left[\rho_a A^*_a\right] \tag{5.106}$$

$$= \alpha \cdot [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})]\,. \tag{5.107}$$

In the second line we chose a strategy which does not use the fictitious outcomes, i.e. $p(x|a) = \delta_{x,a}$ and used the identity: $[\mathrm{id} \otimes (\mathcal{E}^*_x)^\dagger]\phi_+ = \frac{\alpha}{d_V} A_x$. Combining the bounds (5.104) and (5.107) we find that the maximal advantage optimized over all games is lower bounded by:

$$\max_{\mathbb{E}^*} \frac{p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A})}{p^c_{\mathrm{succ}}(\mathbb{E})} \geq \frac{p_{\mathrm{succ}}(\mathbb{E}^*, \mathbb{A})}{p^c_{\mathrm{succ}}(\mathbb{E}^*)} \geq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})] \cdot \frac{1}{1 + \frac{1}{\alpha N}}, \tag{5.108}$$

where $\mathbb{A}$ is a teleportation instrument constructed from $\mathbb{A}$. Since we are free to choose N as big as we like, in the limit $N \to \infty$ the advantage is lower-bounded by $1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})$. To prove the reverse direction we look at the probability of guessing for an arbitrary game $\mathbb{E}$:

$$p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}) = d_V^2 \cdot \max_{p(g|a)} \sum_{x,a,g} p(g|a)\ \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\widetilde{\rho}_a \cdot \phi_+\right] \delta_{g,x} \tag{5.109}$$

$$\leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})]\, d_V^2 \cdot \max_{p(g|a)} \sum_{x,a,g} p(g|a)\, p_{\mathrm{T}}(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\, \sigma_a \cdot \phi_+\right] \delta_{g,x} \tag{5.110}$$

$$\leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})]\, d_V^2 \cdot \max_{\sigma_a \in \mathsf{SEP}(A:B)} \max_{p(g|a)} \sum_{x,a,g} p(g|a)\, p_{\mathrm{T}}(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\, \sigma_a \cdot \phi_+\right] \delta_{g,x} \tag{5.111}$$

$$= [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{A})]\, p^c_{\mathrm{succ}}(\mathcal{G}), \tag{5.112}$$

where the first inequality follows from (5.48), that is $\widetilde{\rho}_a \leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda})]\, p_{\mathrm{T}}(a)\, \sigma_a$ for a probability distribution $p_{\mathrm{T}}(a)$ and some separable state σ_a. Since this holds for any game $\mathbb{E}$ we can write equivalently:

$$\max_{\mathbb{E}} \frac{p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A})}{p^c_{\mathrm{succ}}(\mathbb{E})} \leq 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}). \tag{5.113}$$

Combining the bounds (5.108) and (5.113) we arrive at:

$$\max_{\mathbb{E}} \frac{p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A})}{p^c_{\mathrm{succ}}(\mathbb{E})} = 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}). \tag{5.114}$$

Notice that so far our choice for α was arbitrary. Consider now the maximal classical probability of guessing:

$$p^c_{\mathrm{succ}}(\mathbb{E}) = \max_{\mathbb{A}^c \in \mathcal{F}_{\mathrm{T}}} \max_{p(g|a)} \sum_{x,a,g} p(g|a)\ \mathrm{tr}\big[M_a \cdot (\mathcal{E}_x \otimes \mathrm{id})\rho\big]\ \delta_{g,x}, \tag{5.115}$$

where $\mathbb{A}^c = \{\{M_a\},\ \rho\}$ and optimization is performed over all separable states ρ and arbitrary bipartite measurements $\{M_a\}$. Notice that, by convexity, we can assume that the optimal separable state is of the form $\rho = \sigma \otimes \sigma'$. This allows us to write:

$$\begin{aligned}
p^c_{\mathrm{succ}}(\mathbb{E}) &= \max_{\sigma,\sigma'} \max_{\{M_a\}_a} \max_{p(x|a)} \sum_{x,a} p(x|a)\ \mathrm{tr}[M_a \cdot (\mathcal{E}_x(\sigma) \otimes \sigma')] && (5.116)\\
&= \max_{\sigma} \max_{\{M'_a\}_a} \max_{p(x|a)} \sum_{x,a} p(x|a)\ \mathrm{tr}[M'_a\, \mathcal{E}_x(\sigma)] && (5.117)\\
&= \max_{\sigma} \max_{\{M''_x\}_x} \sum_{x} \mathrm{tr}[M''_x\, \mathcal{E}_x(\sigma)] && (5.118)\\
&= \max_{\sigma} \max_{x} \mathrm{tr}[\mathcal{E}_x(\sigma)] && (5.119)
\end{aligned}$$

where in the second line we defined a new measurement $M'_a = \mathrm{tr}_2[M_a(\mathbb{1} \otimes \sigma')]$ and in the third line we defined $M''_x = \sum_a p(x|a) M'_a$. The last equality follows from the fact that the optimal measurement M''_x is the one which chooses the most likely outcome, i.e. $M''_x = \delta_{x,x^*}\mathbb{1}$ for $x = x^*$ such that $p(x|\sigma) = \mathrm{tr}[\mathcal{E}_x\sigma]$ is maximal.

5.5.5 Complete set of monotones for channel simulation

In this section we show that $q(\mathcal{G}, \mathbb{\Lambda})$ which we defined in (5.13), provide a complete set of monotones for quantum simulation, i.e. all local pre- and post-processings of the the teleportation instrument $\mathbb{\Lambda}$, and that $p_{\mathrm{succ}}(\mathbb{E}, \mathbb{\Lambda})$ which we defined in (5.18) provides a complete set of monotones for classical simulation.

Let us start by focusing on $q(\mathcal{G}, \mathbb{\Lambda})$ and assuming that $\mathbb{\Lambda}$ can be used to simulate $\mathbb{\Lambda}^*$, that is $\mathbb{\Lambda} \geq \mathbb{\Lambda}^*$. We have:

$$\begin{aligned}
q(\mathcal{G}, \mathbb{\Lambda}) &= \max_{\substack{\mathbb{\Lambda}' \leq \mathbb{\Lambda} \\ \mathcal{E}_a \in \mathrm{CPTP}}} \sum_a f(a)\, \mathrm{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda'_a)\, \psi \cdot \xi_a\right] && (5.120)\\
&\geq \max_{\substack{\mathbb{\Lambda}' \leq \mathbb{\Lambda}^* \\ \mathcal{E}_a \in \mathrm{CPTP}}} \sum_a f(a)\, \mathrm{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda'_a)\, \psi \cdot \xi_a\right] && (5.121)\\
&= q(\mathcal{G}, \mathbb{\Lambda}^*), && (5.122)
\end{aligned}$$

since the set $\{\mathbb{A}'|\mathbb{A}' \leq \mathbb{A}^*\}$ is a subset of $\{\mathbb{A}'|\mathbb{A}' \leq \mathbb{A}\}$. We will now assume that $q(\mathcal{G}, \mathbb{A}) \geq q(\mathcal{G}, \mathbb{A}^*)$ holds for all games $\mathcal{G} = \{\psi, \xi_a, f(a)\}$ and show that there always exist a subroutine which allows for the simulation of $\mathbb{A}^*$ by $\mathbb{A}$. Let us start by noting that the following must hold:

$$\forall\, \mathcal{G} \qquad \max_{\substack{\mathbb{A}' \leq \mathbb{A} \\ \mathcal{E}_a \in \text{CPTP}}} \sum_a f(a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda'_a)\, \psi \cdot \xi_a\right] \tag{5.123}$$

$$- \max_{\substack{\mathbb{A}'' \leq \mathbb{A}^* \\ \mathcal{E}'_b \in \text{CPTP}}} \sum_b f(b) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}'_b \circ \Lambda''_b)\, \psi \cdot \xi_b\right] \geq 0. \tag{5.124}$$

Since $\mathbb{A}'' \leq \mathbb{A}^*$ we can write $\Lambda''_b = \sum_{a,\lambda} p(\lambda)\, p(b|a,\lambda)\, \Theta_\lambda \circ \Lambda^*_a \circ \Omega_\lambda$. If we now make a particular (and possibly sub-optimal) choice of $p(\lambda) = \delta_{0,\lambda}$, $p(b|a,\lambda) = \delta_{b,a}$ and $\Theta_\lambda = \Omega_\lambda = \mathrm{id}$ for all λ and also choose $\mathcal{E}'_b = \mathrm{id}$ for all b, then (5.123) implies:

$$\forall\, \mathcal{G} \qquad \max_{\substack{\mathbb{A}' \leq \mathbb{A} \\ \mathcal{E}_a \in \text{CPTP}}} \left[\sum_a f(a) \operatorname{tr}\left[\xi_a \cdot \left((\mathrm{id} \otimes \mathcal{E}_a \circ \Lambda'_a)\psi - (\mathrm{id} \otimes \circ \Lambda^*_a)\psi\right)\right] \right] \geq 0. \tag{5.125}$$

We will now claim that (5.125) can only hold if $\mathbb{A}$ can be used to simulate $\mathbb{A}^*$. Let ω be an arbitrary quantum state and define an operator $\Delta_a := \Lambda'_a(\omega) - \Lambda^*_a(\omega) = \sum_{x,\lambda} p(\lambda) p(b|x,\lambda) \Theta_\lambda \circ \Lambda_x \circ \Omega_\lambda(\omega) - \Lambda^*_a(\omega)$. Notice that using (5.3) we can write:

$$\sum_a \Delta_a = \sum_{a,x,\lambda} p(\lambda) p(a|x,\lambda)\, \Theta_\lambda \circ \Lambda_x \circ \Omega_\lambda(\omega) - \sum_a \Lambda^*_a(\omega) \tag{5.126}$$

$$= \sum_{a,\lambda} p(\lambda) \Theta_\lambda(\rho) - \rho^*. \tag{5.127}$$

which holds for any choice of state ω. Notice that since ρ and ρ^* are density operators and Θ_λ is a CPTP map, the trace of the operator $\widetilde{\Delta} = \sum_a \Delta_a$ vanishes. This means that $\widetilde{\Delta}$ must either have (i) positive and negative eigenvalues or (ii) all eigenvalues equal to zero. However, if (i) holds then we can always choose channels $\mathcal{E}_a$ in (5.125) such that $\mathcal{E}_a = \mathcal{E}$ for all a and $\mathcal{E}$ projects all input states onto the eigenspace corresponding to the negative eigenvalue of Δ, thus violating (5.125) and our initial assumption. This means that (i) cannot be true and the only possibilty is that (ii) holds, i.e. all eigenvalues of Δ are zero and $\Delta = \sum_a \Delta_a = 0$.

Knowing this we will now claim that (5.123) necessarily implies that the operators Δ_b are all identically zero. First, notice that if at least one of them has at least one negative eigenvalue, then (5.123) leads to a contradiction. In particular, let us denote this eigenvalue with $\lambda_* < 0$ and the associated eigenvector with $|\lambda_*\rangle$. Similarily as before, we can always choose $\mathcal{E}_a$ to be channels projecting onto the eigenspace spanned by $|\lambda_*\rangle$, that is $\mathcal{E}_a(\cdot) = \mathrm{tr}[\cdot] \cdot |\lambda_*\rangle\langle\lambda_*|$ for all a. This means that the left-hand side of (5.123) is negative, which is a contradiction. Since $\sum_a \Delta_a = 0$, the only possibility is that for all a we have $\Delta_a = 0$, which then implies that:

$$\Lambda^*_a = \sum_{x,\lambda} p(\lambda) p(a|x,\lambda) \Theta_\lambda \circ \Lambda_x \circ \Omega_\lambda, \tag{5.128}$$

meaning that $\mathbb{A}^*$ can be simulated using $\mathbb{A}$.

Let us now move onto $p_{\text{succ}}(\mathbb{E}, \mathbb{A})$ which we defined in (5.18), and show that it provides a complete set of monotones for classical simulation. We proceed analogously as in the case of $q(\mathcal{G}, \mathbb{A})$. To prove one direction, assume that $p_{\text{succ}}(\mathbb{E}, \mathbb{A}) \geq p_{\text{succ}}(\mathbb{E}, \mathbb{A}^*)$

holds for all $\mathbb{E}$. This and the identity (5.89) implies:

$$\forall\, \mathbb{E} \qquad \max_{p(x|a)} \sum_{a,x} p(x|a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\widetilde{\rho}_a \cdot \phi_+\right] \tag{5.129}$$

$$- \max_{p'(x|b)} \sum_{b,x} p'(x|b) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\widetilde{\rho}^*_b \cdot \phi_+\right] \geq 0, \tag{5.130}$$

where we denoted $\widetilde{\rho}_a = (\mathrm{id} \otimes \Lambda_a)\phi_+$ and $\widetilde{\rho}^*_b = (\mathrm{id} \otimes \Lambda^*_b)\phi_+$. If we now make a particular choice of $p'(x|b) = \delta_{x,b}$ for all b, then (5.129) implies:

$$\forall\, \mathbb{E} \qquad \max_{p(x|a)} \sum_{x} \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x)\left(\sum_a p(x|a)\widetilde{\rho}_a - \widetilde{\rho}^*_x\right) \cdot \phi_+\right] \geq 0. \tag{5.131}$$

We will now claim that (5.131) can only hold if $\mathbb{A}$ can be used to classically simulate $\mathbb{A}^*$. To do so, we can define an operator $\Delta_x := \sum_a p(x|a)\widetilde{\rho}_a - \widetilde{\rho}^*_x = \sum_a p(x|a)(\mathrm{id} \otimes \Lambda_a)\phi_+ - (\mathrm{id} \otimes \Lambda^*_x)\phi_+$. Using analogous arguments as we used below (5.126) we can infer that (5.131) necessarily implies that $\Delta_x = 0$ for all x, or equivalently:

$$\forall\, x \qquad \Lambda^*_x = \sum_a p(x|a)\Lambda_a, \tag{5.132}$$

which means that $\mathbb{A}$ can be used to classicaly simulate $\mathbb{A}^*$ or equivalently $\mathbb{A} \succ_c \mathbb{A}^*$. To prove the reverse direction we assume $\mathbb{A} \succ_c \mathbb{A}^*$ which implies that there exist $p(b|a)$ such that $\Lambda^*_b = \sum_a p(b|a)\Lambda_a$ for all b. For all games $\mathbb{E}$ we then have:

$$p_{\mathrm{succ}}(\mathbb{E}, \mathbb{A}^*) = \max_{p'(x|b)} \sum_{b,x} p'(x|b) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x \circ \Lambda^*_b)\phi_+ \cdot \phi_+\right] \tag{5.133}$$

$$= \max_{p'(x|b)} \sum_{a,b,x} p'(x|b)p(b|a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x \circ \Lambda_a)\phi_+ \cdot \phi_+\right] \tag{5.134}$$

$$\leq \max_{p'(x|a)} \sum_{a,x} p'(x|a) \operatorname{tr}\left[(\mathrm{id} \otimes \mathcal{E}_x \circ \Lambda_a)\phi_+ \cdot \phi_+\right], \tag{5.135}$$

where in the last line we defined a new probability distribution $p'(x|a) = \sum_b p'(x|b)p(b|a)$ and inequality follows since this may be not the most general conditional probability distribution.

Chapter 6

Operational significance of Buscemi nonlocality

6.1 Introduction

6.1.1 Background

Bell nonlocality is a phenomenon occuring when correlations shared between spatially separated parties cannot be explained as arising from any mechanism that operates according to the laws of classical mechanics. The concept of Bell nonlocality is perhaps best understood in terms of a Bell experiment, which is sometimes also called a "no-signalling game". In such a game, a referee distributes two physical systems to two spatially separated players, Alice (A) and Bob (B). Upon receiving their systems, each player is asked a question from a pre-arranged set of questions, labelled x for Alice and y for Bob. Depending on which of the questions was asked, Alice measures her system locally and obtains an outcome a. Similarly, based on his own question, Bob measures his share of the system and obtains b. The data produced from the experiment can be described using a conditional probability distribution $p(a, b|x, y)$, that is the probability of producing outcomes a and b given the choice of measurements labelled by x and y. More details about this approach to quantifying Bell nonlocality can be found in Chapter 2.

6.1.2 Motivation

Importantly, not all entangled states can display Bell non-locality (Werner, 1989; Barrett, 2002; Augusiak, Demianowicz, and Acín, 2014). Quantum states can actually demonstrate other forms of nonlocality which are not accessible in a Bell experiment but which may become apparent in different experimental settings. In a seminal work (Buscemi, 2012) Buscemi generalized Bell's original experiment by allowing the referee to ask "quantum questions". This amounts to replacing the original set of classical (and therefore mutually orthogonal) questions $\{|x\rangle\}$ with a set of quantum states $\{|\omega_x\rangle\}$ which need not be orthogonal. The correlation data $p(a, b|\omega_x, \omega_y)$ obtained in this modified experiment, dubbed *semi-quantum non-signalling games*, differs significantly from its archetypical counterpart. Perhaps the most striking consequence is that the new experiment is powerful enough to reveal the nonlocality[1] of any entangled quantum state, even the nonlocality which would be hidden under a standard Bell test (ibid.). This semi-quantum approach, also called measurement-device-independent (MDI), has been a fruitful line of investigation during the last

[1] We use the term "nonlocality" whenever there does not exist a local model for the observed correlations

decade (Branciard et al., 2013; Cavalcanti, Hall, and Wiseman, 2013; Li et al., 2020; Zhang and Zhao, 2020; Supic et al., 2020).

In this Chapter we propose interpreting the correlation data obtained in a semi-quantum non-signalling game as an indicator of this type of nonlocality which we refer to as *Buscemi nonlocality*. Still, in order to make this realisation well-grounded we should be able to associate Buscemi nonlocality with an operationally meaningful information processing task, in a manner similar to standard Bell nonlocality. Whether such association exists is the main problem addressed in this Chapter. More specifically, we study the following question.

Main question

What is the operational significance of Buscemi nonlocality?

In what follows we will answer this question using the framework of Quantum Resource Theories (QRTs) and the duality of convex optimisation, both of which were described in more detail in Chapter 2.

6.1.3 Contribution

In this Chapter, we focus on the quantum resource theory of Buscemi nonlocality, which is an instance of the resource theory from (Schmid et al., 2020). The natural object relevant for this theory is a generalized measurement (POVM) performed by spatially-separated parties that do not communicate (distributed measurement). We investigate a geometric measure that quantifies the amount of Buscemi nonlocality contained within a given distributed measurement termed Robustness of Buscemi Nonlocality (RoBN). We then address Buscemi nonlocality as a property of states, by considering the maximal amount of Buscemi nonlocality that can be obtained using a given state by any local set of measurements on Alice's and Bob's side.

As our first and main result we show that Buscemi nonlocality has operational significance, by finding an operational task for which Buscemi nonlocality is a natural resource. This can be seen as akin to several seminal results in the field of quantum information which showed the operational character of coherence (Napoli et al., 2016), entanglement (Takagi and Zhuang, 2018), steering (Piani and Watrous, 2015) or Bell nonlocality (Acin, Gisin, and Masanes, 2006) in terms of experimentally relevant information-processing tasks. Moreover, our task gives rise to a complete family of monotones for this resource theory, i.e provides a sufficient and necessary characterisation of Buscemi nonlocality contained in a distributed measurement. Consequently, the average probability of guessing in these family of tasks can be interpreted as a simple and complete set of "Buscemi inequalities" which characterise nonlocality of distributed measurements, in analogy with the celebrated Bell inequalities characterising nonlocality of states.

Our second main result concerns how Buscemi nonlocality relates to other types of nonclassical phenomena studied in the literature: nonclassical teleportation and entanglement. We show that the maximal value of RoBN which can be achieved when Bob (Alice) is allowed to use any measurement is precisely the so-called robustness of teleportation (RoT) of a teleportation channel from Alice (Bob) to Bob (Alice). On the other hand, optimising RoBN over all local measurements for both parties leads to the robustness of entanglement of the state shared by Alice and Bob. This result, despite its clarifying character being of independent interest, leads to new operational tasks for which both nonclassical teleportation and entanglement are natural

resources. These quantitative relationships further expand the results presented in (Cavalcanti and Skrzypczyk, 2016), (Takagi and Zhuang, 2018) and (Lipka-Bartosik and Skrzypczyk, 2020) by proposing new discrimination tasks for which both entanglement and nonclassical teleportation provide advantage over their classical (i.e. separable) counterparts.

As our third and final main result we interpret Buscemi nonlocality from the perspective of single-shot quantum information theory. We show that Buscemi nonlocality, when viewed as a property of a communication channel between the sender (the Referee) and receiver (Alice and Bob), quantifies the maximal amount of information that can be sent reliably when the channel is used only once (the so-called single-shot capacity of a quantum channel). This establishes an important link between Buscemi nonlocality and quantum communication.

6.1.4 Structure

The Chapter is organized as follows. In Sec. 6.2 we cover the relevant formalism, remind the idea of characterizing nonlocality in terms of non-signalling games and recall the robustness quantifier of Buscemi nonlocality (RoBN). In Sec. 6.3.1 we find its operational interpretation in terms of the advantage in the task of distributed state discrimination (DSD). In Sec. 6.3.2 we explore the relationship between Buscemi nonlocality and the concepts of nonclassical teleportation and entanglement. Finally, in Sec. 6.3.3 we describe a tangential view on RoBN from the perspective of single-shot information theory. We conclude with Sec. 6.4 where we summarize our findings and highlight several open problems.

6.2 Buscemi nonlocality

In what follows we will denote a local bipartite measurement on Alice's side (system AA') with $\mathbb{M}^A = \{M_a^{AA'}\}$, where each $M_a^{AA'}$ is a positive semi-definite operator that adds up to the identity (POVM). Similarly we will use $\mathbb{M}^{AB}$ to indicate that the measurement is non-local, i.e. we will treat systems labelled with different letters, e.g. A and B , as two spatially separated parties. We are interested in the most general type of measurement that can be performed in this bipartite scenario without the aid of classical or quantum communication. This can be realized by (i) allowing Alice and Bob to apply arbitrary bipartite measurements in their labs, denoted respectively $\mathbb{M}^A = \{M_a^{AA'}\}$ and $\mathbb{M}^B = \{M_b^{B'B}\}$, where $a \in \{1, \ldots, o_A\}$ and $b \in \{1, \ldots, o_B\}$ denote Alice's and Bob's outcomes and (ii) allowing the two parties to share a quantum state $\rho^{A'B'}$. In this way Alice and Bob can store and share all types of classical information (e.g. classical memory or measurement strategy), as well as quantum information (i.e. shared entanglement). We denote such a measurement with $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$, where the corresponding POVM elements are of the following general form:

$$M_{ab}^{AB} = \mathrm{tr}_{A'B'}\left[\left(M_a^{AA'} \otimes M_b^{B'B}\right)\left(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B\right)\right]. \tag{6.1}$$

Since the sets of all quantum states and quantum measurements are both convex sets, it follows that the set of measurements of the form (6.1) is also a convex set. We will refer to measurements of the form (6.1) as *distributed measurements* and denote the set of all such measurements with $\mathcal{R}_{\mathrm{BN}}$. These measurements are the main (*resourceful*) objects of the resource theory we consider here. Whenever the elements of measurement $\mathbb{M}^{AB}$ can be written as in (6.1) for some choice of shared state and local measurements we will write $\mathbb{M}^{AB} \in \mathcal{R}_{\mathrm{BN}}$. Later in Sec. 6.3.1 we will formally define

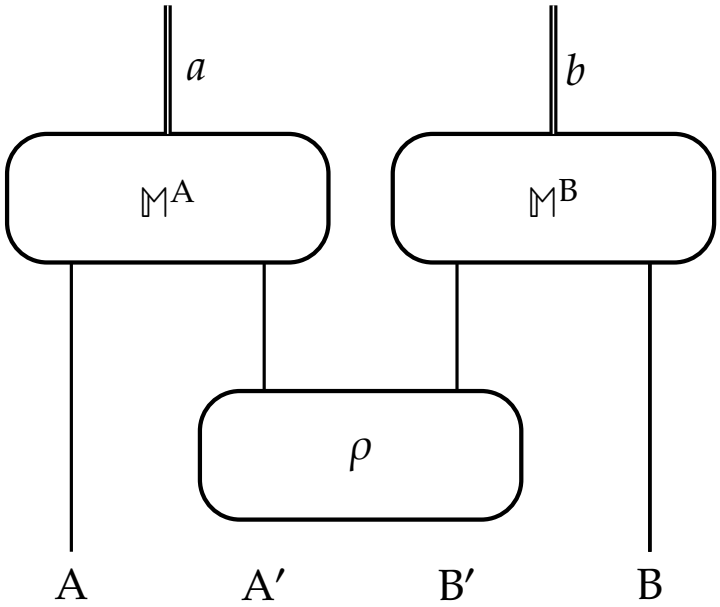

FIGURE 6.1: A schematic diagram of a distributed measurement $\mathbb{M}^{AB}$ composed of local measurements for Alice $\mathbb{M}^{A} = \{M_a^{AA'}\}$, for Bob $\mathbb{M}^{B} = \{M_b^{B'B}\}$ and a state $\rho^{A'B'}$ shared between them. This is the most general type of measurement which Alice and Bob can perform in a distributed scenario which does not allow for communication.

the set of *free measurements* of this resource theory, which turn out to be distributed measurements with a separable shared state. Figure 6.1 illustrates a distributed measurement and describes the relationship between different subsystems. This type of objects appear naturally in a wide range of contexts when studying non-local effects in an MDI setting (Buscemi, 2012).

We now specify the most general class of operations that the separated parties in A and B can perform, without communicating, to improve the properties of their distributed measurement $\mathbb{M}^{AB} = \{M_{ij}^{AB}\}$, where indices $i \in \{1, \dots o_A\}$ and $j \in \{1, \dots o_B\}$ describe measurements outcomes. The free operations for the QRT of Buscemi nonlocality are the so-called Local Operations and Shared Randomness (LOSR). There, Alice and Bob are allowed to share any amount of classical memory described by a random variable λ. Formally this is specified by providing a probability distribution $p(\lambda)$ which is available to both parties. Moreover, before measuring their systems both parties are allowed to locally perform any completely positive and trace-preserving map, potentially conditioned on the value of the shared memory, i.e. we allow for applying $\mathcal{E}_\lambda$ on Alice's and $\mathcal{N}_\lambda$ on Bob's side. Finally, the parties are allowed to post-process their measurement outcomes using arbitrary classical channels $p(a|i,\lambda)$ and $p(b|j,\lambda)$ to produce their final guesses. This procedure leads to the most general type of LOSR operation that can be performed on a measurement of the form (6.1). In what follows we will refer to this as *quantum simulation*:

Definition 6.1 (Quantum simulation)
A quantum simulation of a bipartite measurement $\mathbb{M} = \{M_{ij}\}$ with a subroutine:

$$\mathcal{S} = \{p(\lambda), p(a|i,\lambda), p(b|j,\lambda), \mathcal{E}_\lambda, \mathcal{N}_\lambda\} \tag{6.2}$$

is a transformation which maps the POVM elements of $\mathbb{M}$ into:

$$M'_{ab} = \sum_{i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(\mathcal{E}^\dagger_\lambda \otimes \mathcal{N}^\dagger_\lambda)[M_{ij}], \tag{6.3}$$

where $\mathcal{E}^\dagger$ denotes the (unique) dual map to $\mathcal{E}$. The dual map is formally defined as the adjoint relative to Hilbert-Schmidt inner product.

In other words, any action that can be performed by Alice and Bob in their labs without access to communication can be described by some quantum simulation subroutine.

Quantum simulation induces a natural preorder on the set of all bipartite measurements. Formally, a preorder is an ordering relation that is reflexive ($a \succ a$) and transitive ($a \succ b$ and $b \succ c$ implies $a \succ c$). Here the preorder induced by quantum simulation will be denoted with $\succ_{\mathrm{q}}$, i.e. $\mathbb{M} \succ_{\mathrm{q}} \mathbb{M}'$ if and only if there there exists a subroutine $\mathcal{S}$ which allows $\mathbb{M}$ to simulate $\mathbb{M}'$, i.e. for the two measurements $\mathbb{M}$ and $\mathbb{M}'$, condition (6.3) in Definition 6.1 holds. The notion of simulation will turn out to be relevant for the operational tasks introduced later on.

6.2.1 Buscemi nonlocality and no-signalling games

One of the ways to understand nonlocality is using the concept of no-signalling games that was described in Chapter 2. Let us briefly remind this setting and discuss its relationship with Buscemi nonlocality.

The standard scenario of a no-signalling game involves two cooperating players (Alice and Bob) who play the game against a third party, the referee. The referee chooses a *classical* question $x \in \mathcal{X}$ for Alice and $y \in \mathcal{Y}$ for Bob according to some probability distribution $p(x,y) : \mathcal{X} \times \mathcal{Y} \to [0,1]$, where $\mathcal{X}$ and $\mathcal{Y}$ denote finite sets of questions. Without communicating, and therefore, without knowing what question the other player was asked, Alice (Bob) returns an answer $a \in \mathcal{A}$ ($b \in \mathcal{B}$) from a finite set of possible answers $\mathcal{A}$ ($\mathcal{B}$). Based on the questions asked and the received answers, the referee determines whether the players win or lose the game, according to a pre-arranged set of rules. Such rules are typically expressed using a function $V : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to [0,1]$, where $V(a,b,x,y) = 1$ if and only if Alice and Bob win the game by answering a and b for questions x and y. The average guessing probability in this task can be written as

$$p^V_{\mathrm{guess}}(\mathcal{G}, \mathbb{M}^{AB}) = \sum_{a,b,x,y} p(x,y)p(a,b|x,y)V(a,b,x,y), \tag{6.4}$$

where $\mathcal{G} = \{p(x,y), V\}$ defines the no-signalling game and the conditional probabilities $p(a,b|x,y)$ are related to the local measurements $\{M^{\mathrm{A}}_{a|x}\}$ for Alice and $\{M^{\mathrm{B}}_{b|y}\}$ for Bob, via the Born rule:

$$p(a,b|x,y) = \mathrm{tr}\left[\left(M^{\mathrm{A}'}_{a|x} \otimes M^{\mathrm{B}'}_{b|y}\right)\rho^{\mathrm{A}'\mathrm{B}'}\right]. \tag{6.5}$$

Before going into the details, let us note that the process of "asking classical questions" can be mathematically modelled by sending states from a collection of orthogonal states from a fixed basis, e.g. $\{|x\rangle\}$ such that $\sum_x |x\rangle\langle x| = \mathbb{1}$ and $\langle x|x'\rangle = \delta_{x,x'}$ and similarly for $\{|y\rangle\}$. Such states are perfectly distinguishable and hence Alice and Bob, after receiving their questions, may choose their measurements unambiguously. This can be viewed as giving Alice and Bob the ability to perform controlled bipartite

measurements $\mathbb{M}^{AA'} = \{M_a^{AA'}\}$ and $\mathbb{M}^{B'B} = \{M_b^{B'B}\}$ with the POVM elements:

$$M_a^{AA'} = \sum_x |x\rangle\langle x|^A \otimes M_{a|x}^{A'}, \tag{6.6}$$

$$M_b^{B'B} = \sum_y M_{b|y}^{B'} \otimes |y\rangle\langle y|^B . \tag{6.7}$$

If Alice and Bob share a quantum state $\rho^{A'B'}$ then effectively they have access to a distributed measurement $\mathbb{M}^{AB}$ of the form (6.1). This measurement is then applied to the "questions" they receive, which we denote here with $\omega_x^A = |x\rangle\langle x|^A$ for Alice and $\omega_y^B = |y\rangle\langle y|^B$ for Bob. Therefore their behavior $p(a,b|\omega_x,\omega_y)$ can be written as:

$$p(a,b|\omega_x,\omega_y) := \mathrm{tr}\left[M_{ab}^{AB}\left(\omega_x^A \otimes \omega_y^B\right)\right], \tag{6.8}$$

$$= \mathrm{tr}\left[M_{ab}^{AB}\left(|x\rangle\langle x|^A \otimes |y\rangle\langle y|^B\right)\right], \tag{6.9}$$

$$= \mathrm{tr}\left[\left(M_{a|x}^{A'} \otimes M_{b|y}^{B'}\right)\rho^{A'B'}\right], \tag{6.10}$$

$$= p(a,b|x,y). \tag{6.11}$$

With this in mind we can now formalize the process of asking "quantum questions". This happens precisely when the states sent by the referee are chosen from an arbitrary collection of states $\{\omega_x\}$. Crucially, these states need not be distinguishable and so each of them can be in a superposition of different orthogonal states.

Notice, however, that using quantum states as inputs to the distributed measurement $\mathbb{M}^{AB}$ with local measurements of the form (6.6) and (6.7) can only lead to a probabilistic version of the standard no-signalling game, i.e. Alice and Bob randomize their choices of measurements according to the respective overlaps $p(x'|x) = \langle x'|\omega_x|x'\rangle$ and $p(y'|y) = \langle y'|\omega_y|y'\rangle$. Thus, in order to use the power of asking genuinely quantum questions, one needs to allow for arbitrary bipartite local measurements on both sides. This leads to the general form of a distributed measurement (6.1) with the local POVM elements $\{M_a^{AA'}\}$ and $\{M_b^{BB'}\}$ being now fully general bipartite measurements, and therefore a Buscemi behaviour is of the form:

$$p(a,b|\omega_x,\omega_y) = \mathrm{tr}\left[(M_a^{AA'} \otimes M_b^{B'B})(\omega_x^A \otimes \rho^{A'B'} \otimes \omega_y^B)\right]. \tag{6.12}$$

The above extension of a no-signalling game leads to a novel type of nonlocality which was noticed for the first time in (Buscemi, 2012). Here we will refer to this type of nonclassical correlations as *Buscemi nonlocality*. In this language the main result of (ibid.) states that all entangled states are Buscemi nonlocal.

In what follows we present a consistent way of quantifying Buscemi nonlocality. First we define a proxy quantity called the robustness of Buscemi nonlocality (RoBN) that quantifies how much Buscemi nonlocality can be evidenced using a fixed distributed measurement. This provides a natural quantifier for the resource theory of Buscemi nonlocality of distributed measurements, which is our main focus here. Optimizing the quantity over all choices of local measurements for Alice and Bob gives rise to a quantity which measures the maximal degree of Buscemi nonlocality which can ever be obtained using a given quantum state.

6.2.2 Quantitative measure of Buscemi nonlocality

The fact that Alice and Bob may share entanglement in (6.1) and use it to perform a measurement means that the measurement is inherently nonlocal and can lead to interesting correlations, even when measured on completely independent systems. Our central question then is how to quantify this nonlocality present in a bipartite measurement. To build a valid reference point we first consider the case when a given distributed measurement cannot produce quantum correlations. This means that the behavior $p(a,b|\omega_x,\omega_y) = \mathrm{tr}\left[M_{ab}^{AB}(\omega_x^A \otimes \omega_y^B)\right]$ results from the measurement $\{M_{ab}^{AB}\}$ formed using a separable shared state $\rho^{A'B'} \in \mathsf{SEP}(A:B)$, where $\mathsf{SEP}(A:B)$ denotes the set of all separable operators. Alternatively, in the language of (Schmid, Rosset, and Buscemi, 2020), we can also say that the measurement is free relative to LOSR. Any separable state can be written as:

$$\rho^{A'B'} = \sum_\lambda p(\lambda)\, \rho_\lambda^{A'} \otimes \rho_\lambda^{B'}, \tag{6.13}$$

where $p(\lambda)$ is a classical probability distribution corresponding to a shared random variable λ and $\{\rho_\lambda^{A'}\}$ and $\{\rho_\lambda^{B'}\}$ are collections of local quantum states. The associated distributed measurement from Eq. (6.1) takes the form:

$$M_{ab}^{AB} = \sum_\lambda p(\lambda)\, M_{a|\lambda}^{A} \otimes M_{b|\lambda}^{B}, \tag{6.14}$$

where we denoted $M_{a|\lambda}^{A} := \mathrm{tr}_{A'}[M_a^{AA'}(\mathbb{1}^A \otimes \rho_\lambda^{A'})]$ for Alice and $M_{b|\lambda}^{B} := \mathrm{tr}_{B'}[M_b^{B'B}(\rho_\lambda^{B'} \otimes \mathbb{1}^B)]$ for Bob. This is the most general classical measurement scheme which can be realized if Alice and Bob have access only to classical randomness λ and the ability to locally prepare quantum states in their labs. The set of all measurements that can be written as in (6.14) will be denoted by $\mathcal{F}_{\mathrm{BN}}$. These measurements are the most natural candidates for free objects in the resource theory of Buscemi nonlocality. Notice that measurements from this set have POVM elements that are all separable ($\mathsf{SEP}(A:B)$) and admit a quantum realization ($\mathcal{R}_{\mathrm{BN}}$), i.e can be written as in (6.1) for some choice of local measurements and shared state. Such measurements can never demonstrate Buscemi nonlocality, regardless of the state being measured.

In order to better understand the difference between the sets $\mathcal{R}_{\mathrm{BN}}$ (all distributed measurements) and $\mathcal{F}_{\mathrm{BN}}$ (free distributed measurements), let us consider the following simple example.

Example 1: Distributed measurements

Let Alice and Bob share a two-qubit Werner state:

$$\rho^{A'B'} = p\, \phi_+^{A'B'} + (1-p)\frac{\mathbb{1}^{A'B'}}{4}, \tag{6.15}$$

where $p \in [0,1]$, the state $\phi_+ = |\phi_+\rangle\langle\phi_+|$ and $|\phi_+\rangle := \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$ is a maximally-entangled state. It is widely known that the Werner state (6.15) is separable for all $p \leq 1/3$. Let $\{U_a\}$ for $a = 1,\ldots,4$, be a set of Pauli operators. Consider a measurement $\mathbb{M}^A = \{M_a^{A'A}\}$ with elements:

$$M_a^{A'A} = (U_a^{A'} \otimes \mathbb{1}^A)\phi_+^{A'A}(U_a^{A'} \otimes \mathbb{1}^A)^\dagger, \tag{6.16}$$

Defining an analogous measurement for Bob $\mathbb{M}^B = \{M_b^{BB'}\}$ and using the definition (6.1) allows us to write the distributed measurement $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$ for Alice and Bob as:

$$M_{ab}^{AB} = (U_a^{A'} \otimes U_b^{B'})\rho^{A'B'}(U_a^{A'} \otimes U_b^{B'})^\dagger \tag{6.17}$$

$$= p\,\phi_{ab}^{A'B'} + (1-p)\frac{\mathbb{1}^{A'B'}}{4}, \tag{6.18}$$

where we labelled $\phi_{ab}^{A'B'} := (U_a^{A'} \otimes U_b^{B'\dagger})\phi_+^{A'B'}(U_a^{A'} \otimes U_b^{B'\dagger})^\dagger$. Clearly, $\phi_{ab}^{A'B'}$ is again a maximally-entangled state and therefore each POVM element of $\mathbb{M}^{AB}$ is a Werner state, up to local unitaries. Since entanglement is preserved under local unitary operations, all elements of the distributed measurement $\mathbb{M}^{AB}$ are entangled operators for $p > 1/3$.
Therefore we can conclude that for $p \leq 1/3$ the distributed measurement $\mathbb{M}^{AB}$ can be written as in (6.14), which by definition means that $\mathbb{M}^{AB} \in \mathcal{F}_{\text{BN}}$. Moreover, for $p > 1/3$ we know that each $M_{ab}^{AB} \notin \mathsf{SEP}(A:B)$ and therefore $\mathbb{M}^{AB} \notin \mathcal{F}_{\text{BN}}$. This implies that this distributed measurement is a resourceful measurement in the resource theory of Buscemi nonlocality.

A natural question at this point is: given an arbitrary bipartite measurement $\mathbb{M}^{AB} \in \mathcal{R}_{\text{BN}}$, how can its nonlocal properties be quantified, in particular its ability to generate Buscemi nonlocality? For this purpose it is useful to define the following quantity (Supic, Skrzypczyk, and Cavalcanti, 2017).

Definition 6.2 (Robustness of Buscemi nonlocality)
The robustness of Buscemi nonlocality (RoBN) of a distributed measurement $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$ is the solution to the following optimization problem:

$$\begin{aligned} \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB}) = \min \quad & r \\ s.t. \quad & M_{ab}^{AB} + rN_{ab}^{AB} = (1+r)O_{ab}^{AB} \quad \forall\, a,b, \\ & \{O_{ab}^{AB}\} \in \mathcal{F}_{\text{BN}}, \quad \{N_{ab}^{AB}\} \in \mathcal{R}_{\text{BN}}. \end{aligned} \tag{6.19}$$

Although this may not seem obvious at first sight, the above is a convex optimization problem and hence can be efficiently solved numerically (Boyd and Vandenberghe, 2004; Watrous, 2018) (see Sec. 6.5 further details). Moreover, due to the duality of convex optimization problems the dual formulation of the above has several nice properties which will be useful for our purposes. Robustness-based quantifiers were introduced in (Vidal and Tarrach, 1999; Steiner, 2003) as entanglement quantifiers and since then successfully applied in a wide range of QRTs. The above variant is closely related to the MDI-nonlocality robustness introduced in (Supic, Skrzypczyk, and Cavalcanti, 2017) at the level of probabilities (6.12). In particular, the two quantities are equivalent when the sets of input states $\{\omega_x\}$ and $\{\omega_y\}$ are tomographically-complete. It is also worth mentioning that the quantity defined in Def. 6.2 is not a particular case of the robustness defined for general convex resource theories of measurements (Oszmaniec and Biswas, 2019; Takagi et al., 2019). In particular, in Def. 6.2 the optimization is over all measurements $\{N_{ab}^{AB}\}$ and $\{O_{ab}^{AB}\}$ which have a quantum realization in the no-signalling scenario, whereas the quantifiers considered in (Oszmaniec and Biswas, 2019) allow for arbitrary measurements (in particular also those which require communication). In other words, the above general approach is valid only for measurements performed in a single location,

whereas here we are explicitly interested in a distributed, multipartite scenario. Hence our robustness measure is a genuinely different quantity than the generalized robustness of measurements studied in the above papers.

In Sec. 6.5 we will derive the dual formulation of the RoBN, which can be used to study its operational characterisation. Furthermore, we note that RoBN possesses three natural properties which one would expect from a reasonable measure of nonlocality, i.e:

(*i*) It is *faithful*, meaning that it vanishes if and only if the measurement is classical, i.e:

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = 0 \iff \mathbb{M}^{AB} \in \mathcal{F}_{\mathrm{BN}}. \tag{6.20}$$

(*ii*) It is *convex*, meaning that having access to two distributed measurements $\mathbb{M}_1^{AB}$ and $\mathbb{M}_2^{AB}$ one cannot obtain a better one by using them probabilistically, i.e for $\mathbb{M}^{AB} = p\,\mathbb{M}_1 + (1-p)\,\mathbb{M}_2$ with $0 \leq p \leq 1$, we have:

$$\mathbf{R}_{\mathrm{BN}}\left(\mathbb{M}^{AB}\right) \leq p\,\mathbf{R}_{\mathrm{BN}}(\mathbb{M}_1^{AB}) + (1-p)\,\mathbf{R}_{\mathrm{BN}}(\mathbb{M}_2^{AB}). \tag{6.21}$$

(*iii*) It is *monotonic* (non-increasing) under all quantum simulations. That is, if $\mathbb{N}^{AB}$ can be simulated by $\mathbb{M}^{AB}$ using some quantum simulation strategy (6.2) then

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{N}^{AB}) \leq \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}). \tag{6.22}$$

These properties were proven in (Rosset, Schmid, and Buscemi, 2019) for a more general class of objects. For completeness, we give an independent proof in Sec. 6.5.

Finally, we introduce a quantity which measures how much Buscemi nonlocality can be generated by using a fixed shared state. This leads to the following definition.

Definition 6.3 (Robustness of Buscemi nonlocality of a state)
The robustness of Buscemi nonlocality of a state ρ_{AB} is the solution of the following optimisation problem:

$$\mathbf{R}_{\mathrm{BN}}(\rho_{AB}) := \max_{\mathbb{M}^A,\,\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}), \tag{6.23}$$

where the optimization ranges over all local measurements on Alice's and Bob's side, $\mathbb{M}^{AB}$ is a distributed measurement of the form (6.1) and $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})$ is the robustness quantifier defined in (6.19).

In this way the quantity from Eq. (6.23) is only a function of the *shared state*, rather than the whole *distributed measurement*. It quantifies the maximal "amount" of nonlocality of the corresponding behavior $\{p(a,b|\omega_x,\omega_y)\}$ that can be generated using a fixed $\rho^{A'B'}$, arbitrary local measurements $\mathbb{M}^A$, $\mathbb{M}^B$ and arbitrary input states $\{\omega_x\}$, $\{\omega_y\}$.

6.3 Results

6.3.1 Operational characterisation of RoBN

In the previous section we introduced a measure of Buscemi nonlocality quantifying how "close" a given measurement is to that which would arise from using only local

measurements and shared randomness, i.e. a measurement of the form (6.14). In what follows we will show that RoBN quantifies the advantage offered by a fixed distributed measurement over all classical measurements in a special type of a state discrimination task relevant in the distributed scenario.

Let us now consider a task which is a special case of the no-signalling game described in Sec. 2.4.1. In this case we choose the function $V(a,b,x,y) = \delta_{ax}\,\delta_{by}$. This means that Alice and Bob win if they both manage to guess the values of $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ which were supplied to them by the referee. This is a variation of the standard state discrimination task in which a single player has to guess the realization of a single random variable x. Interestingly, due to the assumption that the players cannot communicate, distributed state discrimination cannot be reduced to the standard state discrimination task.

Task 6.1 (Distributed state discrimination (DSD))**.** The task consists of the following steps:

1. The referee chooses a bipartite state from the ensemble $\{p(x,y), \sigma_{xy}\}$ according to $p(x,y)$ and distributes it among parties by sending one part of it to Alice and the other part to Bob.

2. After receiving their systems, Alice and Bob can preprocess them using arbitrary channels $\{\mathcal{E}^A_\lambda\}$ and $\{\mathcal{N}^B_\lambda\}$, potentially conditioned on a shared randomness λ (so-called LOSR processing).

3. Alice and Bob apply *fixed* local measurements $\mathbb{M}^{AA'} = \{M^{AA'}_i\}$ and $\mathbb{M}^{B'B} = \{M^{AA'}_j\}$ to their shares of the state σ_{xy} and a part of the shared state $\rho^{A'B'}$. They obtain outcomes i and j respectively, which they can postprocess to produce their guesses a and b.

4. Alice and Bob communicate their guesses a and b to the referee and win the game if they *both* correctly guess, i.e. when $a = x$ and $b = y$.

Notice that the second and the third step can be also formulated as allowing Alice and Bob apply any *quantum simulation* (6.3) to their distributed measurement $\mathbb{M}^{AB} \in \mathcal{R}_{\mathrm{BN}}$. Hence the two players are effectively simulating a distributed measurement, denoted by $\mathbb{N}^{AB} \prec \mathbb{M}^{AB}$[2]. The average probability of discriminating states in this discrimination game as specified by $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ can be expressed as

$$p^{DSD}_{\mathrm{guess}}(\mathcal{G}, \mathbb{M}^{AB}) = \max_{\mathbb{N}^{AB} \prec_q \mathbb{M}^{AB}} \sum_{a,b,x,y} p(x,y)\,\mathrm{tr}\left[N_{ab}\sigma_{xy}\right]\delta_{xa}\,\delta_{yb}, \tag{6.24}$$

where the optimization ranges over all measurements $\mathbb{N}^{AB} = \{N_{ab}\}$ which can be quantum-simulated using $\mathbb{M}^{AB}$.

Let us now consider two different situations: (i) a classical scenario in which the distributed measurement performed by Alice and Bob is classical, i.e. $\mathbb{M}^{AB} \in \mathcal{F}_{\mathrm{BN}}$, and (ii) a quantum scenario in which the measurement performed by Alice and Bob is genuinely quantum, i.e it cannot be written as in (6.14).

In the classical case (i) the optimal average probability of guessing which state from the ensemble $\{p(x,y), \sigma_{xy}\}$ was provided can be expressed as

$$p^{DSD}_{\mathrm{guess}}(\mathcal{G}) = \max_{\mathbb{N}^{AB} \in \mathcal{F}_{\mathrm{BN}}} p^{DSD}_{\mathrm{guess}}(\mathcal{G}, \mathbb{N}^{AB}), \tag{6.25}$$

[2]It appears that in the literature the more often used synonim for this type of operations is *LOSR processing*

Note that the above optimization has to be performed over the convex set of measurements of the form (6.14), which is a subset of all separable measurements.

In the quantum case (ii) the above score can be further improved by exploiting Buscemi nonlocality contained in an entangled state which forms the distributed measurement $\mathbb{M}^{AB}$. The maximal amount by which quantum score outperforms classical one can be quantified by studying the ratio

$$\max_{\mathcal{G}} \frac{p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^{AB})}{p^{DSD}_{\text{guess}}(\mathcal{G})}. \tag{6.26}$$

In Sec. 6.5 we show that the maximal advantage which Alice and Bob can achieve when using $\mathbb{M}^{AB} \in \mathcal{R}_{\text{BN}}$ over the best classical distributed measurement is precisely equal to the robustness of Buscemi nonlocality defined in (6.19). Formally, we have the following relation.

Operational character of Buscemi nonlocality

Result 6.1. Let $\mathbb{M}^{AB} = \{M^{AB}_{ab}\}$ be a distributed measurement and $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ be an ensemble of bipartite states. Then :

$$\max_{\mathcal{G}} \frac{p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^{AB})}{p^{DSD}_{\text{guess}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB}). \tag{6.27}$$

This provides a direct operational meaning for Buscemi nonlocality. The proof of Result 6.1 consists of three parts. First we use the primal formulation of the problem (6.19) to show that the advantage from (6.27) is always upper-bounded by the RoBN. Secondly, we identify a set of properties which characterize all distributed measurements and add them to the optimization problem (6.19) as superfluous constraints. Finally, using this characterization we obtain a dual formulation of the problem which, after some simplifications, allows us to extract the optimal ensemble of states $\{p(x,y), \sigma_{xy}\}$ which achieves the optimum in (6.27). The full proof of this result is presented in Sec. 6.5.

The task of distributed state discrimination is a particular instance of a no-signalling game. In this respect we can further consider an advantage (6.26), with the average score $p_{succ}(\mathcal{G}, \mathbb{M}^{AB})$ given by (2.48), and optimize it over all ensembles $\mathcal{G}$ *and* scoring functions $V(a,b,x,y)$. This would allow us to find the largest possible advantage which can be achieved in any possible nonsignalling game. In this way Result 6.1 naturally leads to the following corollary:

Corollary 6.1. Let $\mathbb{M}^{AB}$ and $\mathcal{G}$ be defined as above and let $V(a,b,x,y) : \mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y} \to [0,1]$. Then:

$$\max_{V,\mathcal{G}} \frac{p^{V}_{\text{guess}}(\mathcal{G}, \mathbb{M}^{AB})}{\max\limits_{\substack{\mathbb{N}^{AB}\ s.t. \\ \rho^{A'B'} \in \mathsf{SEP}(A:B)}} p^{V}_{\text{guess}}(\mathcal{G}, \mathbb{N}^{AB})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB}). \tag{6.28}$$

In this way we can also interpret RoBN as a quantifier of the Buscemi nonlocality contained within a given distributed measurement.

6.3.2 Connecting Buscemi nonlocality with other notions of nonclassicality

In this section we show that Buscemi nonlocality can be viewed as a type of nonlocality which is strictly stronger than two other well-known notions of nonlocal correlations: entanglement and nonclassical teleportation.

It is worth mentioning that the authors of (Schmid, Rosset, and Buscemi, 2020) also studied the relationship between Buscemi nonlocality, nonclassical teleportation and entanglement by studying a partial order between objects representing these resources: distributed measurements for Buscemi nonlocality. teleportation instruments for nonclassical teleportation and bipartite states for entanglement. Here we address an analogous problem using a more direct approach: we relate robustness quantifiers of these resource theories and find a direct and simple relationship between them.

Recall that a distributed measurement is composed of two local bipartite measurements and a shared state. This setting is very similar to the teleportation protocol in which Alice locally measures an input state provided by the referee and a part of an entangled state which she shares with Bob. Since the resource used in the teleportation task is effectively "contained" in the resource which is used in the task of distributed state discrimination, it is natural to ask if we can see some connection between these two tasks. In particular, how is the ability of performing nonclassical teleportation related to the ability of demonstrating Buscemi nonlocality? Furthermore, since teleportation is intrinsically related with entanglement (Cavalcanti, Skrzypczyk, and Supic, 2017), also Buscemi nonlocality should be quantitatively related to the entanglement content of a state. In the next section we will show that in fact these three notions of nonclassical correlations are inherently connected and all describe different types of nonlocality.

Buscemi nonlocality and nonclassical teleportation

We have already seen in Chapter 5 that nonclassical teleportation can be viewed as a way of testing nonlocality of a pair of objects: a state and measurement. In particular, the "teleportation resource" in that case is the teleportation channel or, more precisely, a collection of subchannels which form a *teleportation instrument* constructed using the shared state and Alice's measurement. In order to relate nonclassical teleportation with Buscemi nonlocality we first recall the notion of a teleportation instrument.

Definition 6.4 (Teleportation instrument)
A teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ from Alice to Bob is a collection of subchannels $\{\Lambda_a^{A\to B'}\}$ defined as:

$$\Lambda_a^{A\to B'}[\omega^A] = \mathrm{tr}_{AA'}\left[(M_a^{AA'}\otimes \mathbb{1}^{B'})\left(\omega^A\otimes\rho^{A'B'}\right)\right]. \tag{6.29}$$

The above notion fully captures the type of channel obtained during the generalized teleportation experiment. For some applications it may be easier to work with states rather than subchannels. In that case for a collection of input states $\{\omega_x^A\}$ one can consider the so-called teleportation assemblages (teleportages) $\{\tau_{a|x}^{B'}\}$, where the elements of the assemblage are given by $\tau_{a|x}^{B'} := \Lambda_a^{A\to B'}[\omega_x^A]$.

Let us recall that a teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ is said to be classical (or free) if it describes a teleportation experiment performed using a separable shared

state. We can find a general form of a classical teleportation instrument by taking $\rho^{A'B'} = \sum_\lambda p(\lambda)\, \rho_\lambda^{A'} \otimes \rho_\lambda^{B'}$. The associated (classical) teleportation instrument reads:

$$\begin{aligned}\Lambda_a^c(\omega_x) &= \sum_\lambda p(\lambda)\, \mathrm{tr}_{AA'}\left[\left(M_a^{AA'} \otimes \mathbb{1}^B\right)\left(\omega_x^A \otimes \rho_\lambda^{A'} \otimes \rho_\lambda^{B'}\right)\right] \\ &= \sum_\lambda p(\lambda)\, p(a|x,\lambda)\, \rho_\lambda^{B'},\end{aligned} \tag{6.30}$$

where $p(a|x,\lambda) = \mathrm{tr}\left[M_a^{AA'}(\omega_x^A \otimes \rho_\lambda^{A'})\right]$. This is the most general classical teleportation scheme which can be realized if Alice and Bob have access only to classical randomness λ and the ability to locally prepare quantum states in their labs. In what follows we will denote the set of all instruments which can be written as in (6.30) by $\mathcal{F}_\mathrm{T}$. If a teleportation instrument cannot be written in this way, we will refer to it as "nonclassical" and denote the set of all such instruments with $\mathcal{R}_\mathrm{T}$. The quantity which quantitatively measures the amount of nonclassicality associated with a given teleportation instrument is called Robustness of Teleportation (RoT) (ibid.). For a teleportation instrument $\mathbb{\Lambda}^{A\to B'} = \{\Lambda_a^{A\to B'}\}$ it is defined as:

$$\begin{aligned}\mathbf{R}_\mathrm{T}(\mathbb{\Lambda}^{A\to B'}) = \min_{r,\,\{\Gamma_a^{A\to B'}\},\{\Omega_a^{A\to B'}\}} \quad & r \\ \text{s.t.} \quad & \Lambda_a^{A\to B'} + r\,\Omega_a^{A\to B'} = (1+r)\,\Gamma_a^{A\to B'} \;\; \forall\, a, \\ & \{\Gamma_a^{A\to B'}\} \in \mathcal{F}_\mathrm{T}, \quad \{\Omega_a^{A\to B'}\} \in \mathcal{R}_\mathrm{T}.\end{aligned} \tag{6.31}$$

It turns out that the above is also a convex optimization problem which can be seen by formulating the constraints using the Choi-Jamiołkowski isomorphism (see Appendix D for details). With the above notation we can now address our next result which relates Buscemi nonlocality with nonclassical teleportation.

Relation between Buscemi nonlocality and nonclassical teleportation

Result 6.2. Let $\mathbb{M}^{AB}$ be a distributed measurement composed of local bipartite measurements $\mathbb{M}^A$ and $\mathbb{M}^B$ and a shared state $\rho^{A'B'}$. Then:

$$\max_{\mathbb{M}^B} \quad \mathbf{R}_\mathrm{BN}(\mathbb{M}^{AB}) = \mathbf{R}_\mathrm{T}(\mathbb{\Lambda}^{A\to B'}), \tag{6.32}$$

where the optimization is over all local measurements $\mathbb{M}^B = \{M_b^{B'B}\}$ for Bob and $\mathbb{\Lambda}^{A\to B'}$ is defined as in Eq. (6.31). An analogous result holds for a teleportation instrument $\mathbb{\Lambda}^{B\to A'}$ if we instead optimize the l.h.s of Eq. (6.32) over all local measurements for Alice.

The proof of this result is in Sec. 6.5. Let us now use this result to show a new operational interpretation of the above teleportation quantifier.

Consider a task involving two players, Alice and Bob, who have access to a teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ connecting their labs. Let the referee be in possession of an ensemble of bipartite quantum states $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$. Just as before, the players may discuss their strategy before the game begins. This means that they may use a shared classical memory λ with a corresponding distribution $p(\lambda)$ and conditioning on it Alice may apply one of the channels $\{\mathcal{E}_\lambda^A\}$ to the input of the teleportation instrument and Bob may apply $\{\mathcal{N}_\lambda^{B'}\}$ to the output. The crucial difference here between the standard teleportation protocol is that Bob *does not* know Alice's

measurement outcome and so his correction cannot depend on it. The task posed between Alice and Bob is the following:

Task 6.2 (Teleportation-assisted state discrimination (TSD))**.** The task consists of the following steps:

1. The referee chooses a bipartite state from the ensemble $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ according to $p(x,y)$ and distributes it among parties by sending one part of it to Alice and the other part to Bob.

2. Alice sends her part of the state to Bob using a teleportation instrument $\mathbb{\Lambda}^{A\to B'}$. She is also allowed to pre-process her part of the state conditioned on the classical randomness λ using a collection of channels $\{\mathcal{N}^A_\lambda\}$. Based on the outcome of the teleportation instrument i and potentially λ she produces a guess a via $p(a|i,\lambda)$.

3. Bob applies a correction $\{\mathcal{E}^{B'}_\lambda\}$ conditioned on the value of a shared random variable λ to the teleported state he received from Alice. He then measures both parts of the system using an *arbitrary* measurement $\mathbb{M}^B = \{M^{BB'}_b\}$ and produces a guess b.

4. Alice and Bob win the game if they both simultaneously guess correctly, i.e. if $a = x$ and $b = y$.

The average probability of guessing in the above discrimination task can be expressed as:

$$p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{\Lambda}^{A\to B'}) = \max_{\mathbb{M}^B} \max_{\Phi \prec_q \mathbb{\Lambda}} \sum_{a,b,x,y} p(x,y)\, \text{tr}\left[M^{B'B}_b (\Phi^{A\to B'}_a \otimes \text{id}^B)\, \sigma^{AB}_{xy}\right] \delta_{xa}\delta_{yb}, \quad (6.33)$$

where the optimization ranges over all measurements $\mathbb{M}^B = \{M^{B'B}_b\}$ on Bob's side and all teleportation instruments $\Phi^{A\to B'} = \{\Phi^{A\to B'}_a\}$ which can be quantum-simulated using the instrument $\mathbb{\Lambda}^{A\to B'} = \{\Lambda^{A\to B'}_i\}$. The elements of such a simulated instrument are of the form

$$\Phi^{A\to B'}_a[\cdot] = \sum_{i,\lambda} p(\lambda) p(a|i,\lambda) \mathcal{N}^{B'}_\lambda \circ \Lambda^{A\to B'}_i \circ \mathcal{E}^A_\lambda[\cdot] \quad (6.34)$$

for some choice of local channels $\{\mathcal{E}^A_\lambda\}$, $\{\mathcal{N}^{B'}_\lambda\}$ and probabilities $p(a|i,\lambda)$ and $p(\lambda)$.

The optimal average probability of guessing that can be achieved using only classical resources (i.e. a separable shared state, meaning that the teleportation instrument is classical) can be written as

$$p^{TSD}_{\text{guess}}(\mathcal{G}) = \max_{\mathbb{F}^{A\to B'} \in \mathcal{F}_{\text{T}}} p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{F}^{A\to B'}), \quad (6.35)$$

where $\mathbb{F}^{A\to B'}$ stands for a classical teleportation instrument from Alice to Bob. The maximal advantage which can be offered by any resourceful teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ in the task of TSD is precisely equal to the quantifier of nonclassical teleportation defined in (6.31). This is captured by the following result.

Operational character of nonclassical teleportation

Result 6.3. Let $\mathbb{\Lambda}^{A\to B'} = \{\Lambda_a^{A\to B'}\}$ be a teleportation instrument from Alice to Bob and let $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ be an ensemble of bipartite states. Then the following holds:

$$\max_{\mathcal{G}} \frac{p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{\Lambda}^{\mathbb{A}\to\mathbb{B}})}{p^{TSD}_{\text{guess}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{T}}(\mathbb{\Lambda}^{A\to B}). \tag{6.36}$$

Proof. Consider maximizing both sides of Eq. (6.27) over all measurements $\mathbb{M}^B$ on Bob's side. Due to the Result 6.2, the right-hand side of Eq. (6.27) is equal to $1 + \mathbf{R}_{\text{T}}(\mathbb{\Lambda}^{A\to B'})$. On the other hand, notice that we can interchange maximisation over $\mathcal{G}$ with maximisation over $\mathbb{M}^B$. Since $p^{DSD}_{\text{guess}}(\mathcal{G})$ does not depend on $\mathbb{M}^B$, the left-hand side of Eq. (6.27) becomes:

$$\max_{\mathcal{G}} \frac{\max_{\mathbb{M}^B} p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^{AB})}{p^{DSD}_{\text{guess}}(\mathcal{G})} = \max_{\mathcal{G}} \frac{p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{\Lambda}^{A\to B'})}{p^{DSD}_{\text{guess}}(\mathcal{G})} \tag{6.37}$$

$$= \max_{\mathcal{G}} \frac{p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{\Lambda}^{A\to B'})}{p^{TSD}_{\text{guess}}(\mathcal{G})}, \tag{6.38}$$

where the last equality follows since:

$$p^{DSD}_{\text{guess}}(\mathcal{G}) = \max_{\mathbb{F}^{AB}\in\mathcal{F}_{BN}} p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{F}^{AB}) \tag{6.39}$$

$$= \max_{\mathbb{F}^{A\to B'}\in\mathcal{F}_T} \max_{\mathbb{M}^B} p^{TSD}_{\text{guess}}(\mathcal{G}, \mathbb{F}^{A\to B'}) \tag{6.40}$$

$$= p^{TSD}_{\text{guess}}(\mathcal{G}). \tag{6.41}$$

This completes the proof. □

Buscemi nonlocality and entanglement

Let us now explore the link between Buscemi nonlocality, which we defined as a property of a bipartite state and local measurements, and entanglement (a property of the state only). Among the large variety of known entanglement quantifiers, we are going to choose the one which most naturally relates to the RoBN — the so-called generalized Robustness of Entanglement (RoE), denoted here with $\mathbf{R}_{\text{E}}(\rho)$. This entanglement quantifier was considered for the first time in (Vidal and Tarrach, 1999) and generalized in (Steiner, 2003) and since then proved to be useful in several different contexts, e.g. in proving that all entangled states can demonstrate nonclassical teleportation (Cavalcanti and Skrzypczyk, 2016), in exploring the connection between entanglement and permutation symmetry (Rajagopal and Rendell, 2002) or in studying the effects of local decoherence on multi-party entanglement (Simon and Kempe, 2002). This quantifier also has two interesting operational interpretations: it quantifies the maximal advantage that can be achieved in a bipartite subchannel discrimination task (Takagi et al., 2019) and the maximal advantage in the task of local subchannel discrimination with a quantum memory (Lipka-Bartosik and Skrzypczyk, 2020). It is defined in terms of the following convex optimization

problem:

$$\begin{aligned} \mathbf{R}_{\mathrm{E}}(\rho^{AB}) = \min_{r,\eta^{AB},\sigma^{AB}} \quad & r \\ \text{s.t.} \quad & \rho^{AB} + r\,\eta^{AB} = (1+r)\sigma^{AB} \\ & \eta^{AB} \geq 0, \quad \operatorname{tr}\eta^{AB} = 1 \\ & \sigma^{AB} \in \mathsf{SEP}(A:B), \quad \operatorname{tr}\sigma^{AB} = 1. \end{aligned} \tag{6.42}$$

Using this definition we can now address our next result which relates Buscemi nonlocality with entanglement.

Relation between Buscemi nonlocality and entanglement

Result 6.4. Let $\mathbb{M}^{AB}$ be a distributed measurement composed of local measurements $\mathbb{M}^{A}$ and $\mathbb{M}^{B}$ and a shared state $\rho^{A'B'}$. Then:

$$\max_{\mathbb{M}^{A},\mathbb{M}^{B}} \quad \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \mathbf{R}_{\mathrm{E}}(\rho^{A'B'}), \tag{6.43}$$

where the optimization is over all local measurements for Alice $\mathbb{M}^{A} = \{M_a^{AA'}\}$ and for Bob $\mathbb{M}^{B} = \{M_b^{B'B}\}$.

The proof of this result can be found in Sec. 6.5. Notice that the above relationship allows us to directly infer that the maximal amount of Buscemi nonlocality that can ever be generated using a given state, defined in (6.23), is precisely equal to its entanglement content. Therefore we may write:

$$\mathbf{R}_{\mathrm{BN}}(\rho_{AB}) = \mathbf{R}_{\mathrm{E}}(\rho^{AB}). \tag{6.44}$$

The above fact can be also understood by noting that distributed measurements can "encode" the nonclassicality of every other type of bipartite nosignalling resource (Schmid, Rosset, and Buscemi, 2020). The relationship (6.43) along with Result 6.1 also allows to find a new operational interpretation of the RoE. Consider again the task of DSD with the relaxation that Alice and Bob may now apply arbitrary local measurements in their labs. The goal for Alice and Bob remains the same: to guess which state from the ensemble $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ was prepared, under the assumption that no communication is allowed. In this way the task posed between Alice and Bob is the following:

Task 6.3. (Entanglement-assisted state discrimination (ESD)) The task consists of the following steps:

1. The referee chooses a bipartite state from the ensemble $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ according to $p(x,y)$ and distributes it among parties by sending one part of it to Alice and the other part to Bob.

2. Alice and Bob apply *arbitrary* local measurements $\mathbb{M}^{A}$ and $\mathbb{M}^{B}$ to the states they received and their part of the shared state $\rho^{A'B'}$ and receive outcomes a and b, respectively.

3. Alice and Bob win the game if they both guess which state was provided, i.e. guess both x and y.

The average probability of guessing in this task can be expressed as:

$$p^{ESD}_{\text{guess}}(\mathcal{G}, \rho^{A'B'}) = \max_{\mathbb{M}^A, \mathbb{M}^B} \sum_{a,b,x,y} p(x,y) \operatorname{tr}\left[M^{AB}_{ab} \sigma^{AB}_{xy}\right] \delta_{xa}\delta_{yb}, \tag{6.45}$$

where the optimization ranges over all measurements $\mathbb{M}^A = \{M^{AA'}_a\}$ on Alice's and $\mathbb{M}^B = \{M^{B'B}_b\}$ on Bob's side with measurement M^{AB}_{ab} of the form (6.1).

The best average probability of guessing in the classical scenario (i.e. when the shared state is separable) is given by:

$$\begin{aligned} p^{ESD}_{\text{guess}}(\mathcal{G}) &= \max_{\sigma^{A'B'} \in \mathsf{SEP}(A:B)} p^{ESD}_{\text{guess}}(\mathcal{G}, \sigma^{A'B'}) \\ &= \max_{\mathbb{N}^{AB} \in \mathcal{F}_{\text{BN}}} p^{ESD}_{\text{guess}}(\mathcal{G}, \mathbb{N}^{AB}) \\ &= p^{DSD}_{\text{guess}}(\mathcal{G}). \end{aligned} \tag{6.46}$$

The maximal advantage which can be offered by an entangled state $\rho^{A'B'}$ in the ESD task can be quantified using the RoE. This is the content of our next result.

Operational character of entanglement

Result 6.5. Let $\rho^{A'B'}$ be a bipartite state shared between Alice and Bob and let $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ be an ensemble of bipartite states. Then the following holds:

$$\max_{\mathcal{G}} \frac{p^{ESD}_{\text{guess}}(\mathcal{G}, \rho^{A'B'})}{p^{ESD}_{\text{guess}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{E}}(\rho^{A'B'}). \tag{6.47}$$

Proof. The proof of Result 6.5 proceeds similarly to the case of nonclassical teleportation. Let us maximise both sides of (6.27) over all measurements on Alice's and Bob's side, i.e. over all $\mathbb{M}^A$ and $\mathbb{M}^B$. Due to Result 6.4 , the right-hand side of (6.27) is equal to $1 + \mathbf{R}_{\text{E}}(\rho^{A'B'})$. On the other hand, due to (6.46) we can write the left-hand side of (6.27) as:

$$\max_{\mathcal{G}} \frac{\max_{\mathbb{M}^A, \mathbb{M}^B} p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^{AB})}{p^{DSD}_{\text{guess}}(\mathcal{G})} = \max_{\mathcal{G}} \frac{p^{ESD}_{\text{guess}}(\mathcal{G}, \rho^{A'B'})}{p^{ESD}_{\text{guess}}(\mathcal{G})}. \tag{6.48}$$

This completes the proof.

Finally, let us note that entanglement-assisted state discrimination is a particular instance of a no-signalling game in which we fix $V(a,b,x,y) = \delta_{xa}\delta_{by}$ and allow for optimising over local measurements. This exactly corresponds to the average score studied in Ref. (Buscemi, 2012). Using this realisation we can now consider the maximal advantage in the task of entanglement-assisted state discrimination (6.47) and optimise it not only over ensembles $\mathcal{G}$, but also over all predicates $V(a,b,x,y)$, in a manner exactly similar as in the case of Corollary 6.1. This therefore yields the largest possible advantage that can be achieved in *any* no-signalling game. In this way Result 6.5 naturally leads to the following corollary.

Corollary 6.2. Let $\mathbb{M}^{AB}$ and $\mathcal{G}$ be defined as before and let $V(a,b,x,y) : \mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y} \to [0,1]$. Then:

$$\max_{V,\mathcal{G}} \frac{\max\limits_{\mathbb{M}^A,\mathbb{M}^B} p^V_{\text{guess}}(\mathcal{G},\mathbb{M}^{AB})}{\max\limits_{\sigma\in\text{SEP}(A:B)} \max\limits_{\mathbb{N}^A,\mathbb{N}^B} p^V_{\text{guess}}(\mathcal{G},\mathbb{N}^{AB})} = 1 + \mathbf{R}_{\text{E}}(\rho^{A'B'}), \tag{6.49}$$

where $\mathbb{N}^{AB} = \{N^{AB}_{ab}\}$ with the POVM elements defined as:

$$N_{ab} := \text{tr}_{A'B'}\left[\left(N^{AA'}_a \otimes N^{B'B}_b\right)\left(\mathbb{1}^A \otimes \sigma^{A'B'} \otimes \mathbb{1}^B\right)\right]. \tag{6.50}$$

In this way we can now interpret RoE as a quantifier of the Buscemi nonlocality contained within a given *state*. This not only re-derives the main result of Ref. (Buscemi, 2012), but also makes it significantly stronger; the RoE can now be seen as the quantifier of the maximal advantage in any no-signalling game, therefore providing a completely new interpretation for this well-known entanglement quantifier. □

Complete sets of monotones for quantum simulation

We finish this section by showing that the average guessing probability in the task of DSD completely describes the preorder induced by quantum simulation on distributed measurements $\mathbb{M}^{AB}$. Formally this means that the average guessing probability $p^{DSD}_{\text{guess}}(\mathcal{G},\mathbb{M}^{AB})$ when viewed as a function of $\mathcal{G}$ forming a complete set of monotones for quantum simulation of $\mathbb{M}^{AB}$. This is captured by the following result:

Complete set of monotones for quantum simulation

Result 6.6. Any distributed measurement $\mathbb{M}^{AB}$ can quantum-simulate another measurement $\mathbb{N}^{AB}$ if and only if for all ensembles $\mathcal{G} = \{p(x,y),\sigma_{xy}\}$ the following holds:

$$p^{DSD}_{\text{guess}}(\mathcal{G},\mathbb{M}^{AB}) \geq p^{DSD}_{\text{guess}}(\mathcal{G},\mathbb{N}^{AB}). \tag{6.51}$$

This infinite family of inequalitites (one for each game $\mathcal{G}$) we interpret as "Buscemi inequalities" mentioned in the introduction.

In other words, quantum simulation (or LOSR) can never improve the discrimination ability of any distributed measurement. The proof of this result is in Sec. 6.5

6.3.3 RoBN as a quantifier in single-shot information theory

We now address another way of interpreting RoBN from the point of view of single-shot quantum information theory. In particular, in Sec. 6.5 we show that RoBN also quantifies the entanglement-assisted min-accessible information of a quantum-to-classical bipartite channel (i.e. a channel with quantum inputs and classical outputs). This connection parallels analogous results from the literature which correspond to single party quantum-to-classical channels (Skrzypczyk and Linden, 2019; Takagi et al., 2019). We start by noticing that any distributed measurement $\mathbb{M}^{AB}$ can be seen

as an entanglement-assisted quantum-to-classical channel:

$$\mathcal{N}^{AB\to XY}[\omega_x^A \otimes \omega_y^B] = \sum_{a,b} p(a,b|\omega_x,\omega_y)\,|a\rangle\langle a|^X \otimes |b\rangle\langle b|^Y\,, \tag{6.52}$$

with $p(a,b|\omega_x,\omega_y)$ as in (6.12). In quantum information theory the standard quantifier of the maximal amount of classical information that can be reliably sent through a quantum channel is the accessible information which is defined for an arbitrary quantum channel $\mathcal{R}$ as:

$$I^{\text{acc}}(\mathcal{R}) = \max_{\mathcal{E},\mathcal{D}} I(X:G), \tag{6.53}$$

where $\mathcal{E} = \{p(x),\sigma_x\}$ is an ensemble of states which encode classical random variable X distributed according to $p(x)$, $\mathcal{D} = \{D_g\}$ is the decoding POVM which produces an outcome g with probability $p(g|x) := \text{tr}\big[D_g \cdot \mathcal{R}[\sigma_x]\big]$ and $I(X;G) = H(X) - H(X|G)$ is the mutual information of the distribution $p(x,g) := p(x)p(g|x)$. In the single-shot case a more relevant quantity is the *min-accessible information* $I^{\text{acc}}_{\min}(\mathcal{R})$ which is defined as (Ciganovi, Beaudry, and Renner, 2014):

$$I^{\text{acc}}_{\min}(\mathcal{R}) = \max_{\mathcal{E},\mathcal{D}} \big[H_{\min}(X) - H_{\min}(X|G)\big]\,, \tag{6.54}$$

where the optimization ranges over the same encodings and decodings as before and single-shot entropies are given by (Renner, 2005):

$$H_{\min}(X) = -\log \max_x p(x), \tag{6.55}$$

$$H_{\min}(X|G) = -\log\left[\sum_g \max_x p(x,g)\right], \tag{6.56}$$

Let us now consider an encoding of a bipartite random variable $X \times Y$, i.e $\mathcal{E} = \{p(x,y),\sigma_{xy}\}$ and the associated decoding $\mathcal{D} = \{D_g\}$ for $g = 1,\ldots,|X|\cdot|Y|$. In Appendix G we show that for this particular setting RoBN quantifies the min-accessible information of the channel $\mathcal{N}^{AB\to XY}$. Formally, we have the following result:

Information-theoretic interpretation of Buscemi nonlocality

Result 6.7. Let $\mathcal{N}^{AB\to XY}$ be a quantum-to-classical channel of the form (6.52). Then the following holds:

$$I^{\text{acc}}_{\min}(\mathcal{N}^{AB\to XY}) = \log[1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB})] \tag{6.57}$$

The proof of this result is in Sec. 6.5. The above result provides an alternative way of interpreting RoBN as the maximal amount of min-mutual information that can be obtained between the input and output of the channel (6.52) when using it only once.

6.4 Discussion and open problems

In this work we have studied the notion of Buscemi nonlocality when it is formalized as a quantum resource theory of distributed measurements. This formulation

allowed us to establish a direct operational interpretation of Buscemi nonlocality in terms of a practical information-theoretic task called distributed state discrimination (Result 6.1). We have shown that the average guessing probability in this task provides a complete set of monotones for the partial order of distributed measurements induced by quantum simulation (Result 6.6). This result can be also interpreted as a *complete* family of "Buscemi inequalities" which quantify nonlocal properties of distributed measurements.

This operational link was derived using a geometric quantity measuring the strength of nonlocal correlations generated using a given distributed measurement (namely RoBN). By connecting this quantifier with other measures of nonlocality we inferred a quantitative relationship between distributed measurements, nonclassical teleportation and quantum entanglement, a realisation which we believe to be of an independent interest. In particular, we have shown that the robustness of Buscemi nonlocality optimised over all local measurements for one party is equal to the robustness of nonclassical teleportation (Result 6.2). Similarily, optimising RoBN over local measurements for both parties gives the robustness of entanglement (Result 6.4). This naturally leads to new operational interpretations for both of these quantifiers, in terms of appropriately tailored state discrimination tasks of: teleportation-assisted state discrimination (Result 6.3), and entanglement-assisted state discrimination (Result 6.5).

We have also shown that the maximal amount of nonlocality that can ever be generated using a fixed bipartite state, is directly proportional to its entanglement content. The entanglement content in this case is characterised by the robustness of entanglement, a widely-known entanglement quantifier with direct operational significance. Importantly, this not only re-derives the main result of Ref. (Buscemi, 2012), but also makes it significantly stronger; the generalised robustness of entanglement can now be seen as the quantifier of the maximal advantage in any no-signalling game (Corollary 6.1 and Corollary 6.2)

As our last result we have interpreted Buscemi nonlocality from the perspective of single-shot quantum information theory (Result 6.7). In particular, we have shown that Buscemi nonlocality, when viewed as a property of a communication channel between the sender (the Referee) and receiver (Alice and Bob), quantifies the maximal amount of information that can be sent reliably when the channel is used only once (the so-called single-shot capacity of a quantum channel). We have shown that the RoBN can be viewed as the maximal single-shot capacity offered by a bipartite quantum-to-classical channel. This establishes an important link between Buscemi nonlocality and the single-shot theory of quantum communication.

Finally, we emphasize that while we focused exclusively on quantifying Buscemi nonlocality using a robustness-based measure, our results can be easily extended to address the so-called *weight-based* resource quantifiers (Elitzur, Popescu, and Rohrlich, 1992; Lewenstein and Sanpera, 1998). These geometric measures find their operational meaning in the so-called exclusion tasks (Ducuara and Skrzypczyk, 2019; Uola, Kraft, and Abbott, 2020). Consequently, the resource quantifiers of: weight of Buscemi nonlocality, weight of nonclassical teleportation, and the weight of entanglement, are quantifiers characterising: distributed state exclusion (DSE), teleportation-assisted state exclusion (TSE), and entanglement-assisted state exclusion (ESE), respectively.

We believe that the results presented in this work will shed new light on the complex structure of different types of nonclassical effects observed in Nature, as well as on their practical relevance for physically-motivated tasks.

This work also provides an example of a multiobject quantum resource theory which cannot be reduced to a theory of either measurements, states, channels, or state-measurement pairs (Ducuara, Lipka-Bartosik, and Skrzypczyk, 2020a). This also means that the composite objects we study here constitute genuine multiobject quantum resources (i.e. those which demonstrate a nontrivial "interaction" between different component resources). In our case such a nontrivial interaction appears when having a resourceful state (entangled) and a resourceful measurement can still lead to a free composite object (i.e. LOSR-free distributed measurement). It is an interesting open question to see if one can find additional examples of multiobject resource theories which address such irreducible resources. This is in sharp contrast to a recently introduced multiobject resource theory of state-measurement pairs, where the resources independently contribute to the benefit of the operational task of discrimination and exclusion of subchannels (ibid.).

One of the standard questions addressed by quantum resource theories is determining when and at what rate a large number of copies of one resource can be converted into another. The fact that multiobject QRTs cannot be seen as resource theories of constituent objects leads a natural question of whether this can be used to improve the existing asymptotic protocols. For example, in the resource theory of nonclassical teleportation one can ask whether n uses of teleportation instrument can lead to a better teleportation than using n copies of the shared state. Similarly we can ask whether access to n uses of a distributed measurement can be in advantageous over using bipartite measurements and n copies of the shared state.

6.5 Proofs

6.5.1 Equivalent formulation for RoBN

By definition RoBN is a conic program. This means that we can use the tools of convex optimization theory to find its dual and from that obtain useful information about the primal problem. We will assume a knowledge of the tools of conic programming, and direct the interested reader to (Boyd and Vandenberghe, 2004). Let us start from the formulation given in the main text and substitute $\widetilde{N}^{AB}_{ab} = rN^{AB}_{ab}$ and $\widetilde{O}^{AB}_{ab} = (1+r)O^{AB}_{ab}$. After this substitution the primal problem can be written as:

$$\begin{aligned} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \quad & \min \quad r && (6.58)\\ & s.t. \quad M^{AB}_{ab} + \widetilde{N}^{AB}_{ab} = \widetilde{O}^{AB}_{ab} \qquad \forall\, a, b, && (6.59)\\ & \{\widetilde{O}^{AB}_{ab}\} \in \mathcal{F}_{\mathrm{BN}}, \quad \{\widetilde{N}^{AB}_{ab}\} \in \mathcal{R}_{\mathrm{BN}}, && (6.60) \end{aligned}$$

where the optimization is performed over r, $\{\widetilde{N}^{AB}_{ab}\}$ and $\{\widetilde{O}^{AB}_{ab}\}$. Notice that any collection of operators inside $\mathcal{R}_{\mathrm{BN}}$ or $\mathcal{F}_{\mathrm{BN}} = \mathcal{F}_{\mathrm{SEP}} \cap \mathcal{R}_{\mathrm{BN}}$ satisfies its own "no-signalling" constraint which can be easily deduced from the definition of the set $\mathcal{R}_{\mathrm{BN}}$. Moreover, any operator in $\mathcal{F}_{\mathrm{BN}}$ is separable. In this way for any $\{X^{AB}_{ab}\} \in \mathcal{F}_{\mathrm{BN}}$ we can write:

$$\sum_a X^{AB}_{ab} = \mathbb{1}^A \otimes X^B_b \quad \forall\, b, \qquad \sum_b X^{AB}_{ab} = X^A_a \otimes \mathbb{1}^B \quad \forall\, a, \qquad X^{AB}_{ab} \in \mathsf{SEP}(A:B), \tag{6.61}$$

$$\sum_b X^B_b = \mathbb{1}^B, \qquad \sum_a X^A_a = \mathbb{1}^A. \tag{6.62}$$

Now we are going to add a family of such redundant constraints to our optimization problem. Note that we can always do that since adding constraints which are automatically satisfied by any operator in the feasible set does not change the optimal value of the program. Moreover, we can also relax the constraint (6.59) to an inequality $M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} \leq \widetilde{O}^{AB}$ without changing the optimal value of the conic program. To see why this is the case suppose we have solved the relaxed problem using variables r^{rel}, $\{\widetilde{N}_{ab}^{AB,rel}\}$, $\{\widetilde{O}_{ab}^{AB,rel}\}$ and $X_{ab}^{AB,rel} \geq 0$ and such that for all a and b we have: $M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,rel} = \widetilde{O}_{ab}^{AB,rel} - X_{ab}^{AB,rel}$. Then the optimal value of the relaxed program becomes:

$$\begin{aligned}
\mathbf{R}_{\mathrm{BN}}^{\mathrm{rel}}(\mathbb{M}^{AB}) &= -1 + \frac{1}{d^2}\sum_{ab} \operatorname{tr} \widetilde{O}_a^{AB,rel} && (6.63)\\
&= -1 + \frac{1}{d^2}\sum_{ab} \operatorname{tr}\left[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,rel} + X_{ab}^{AB,rel}\right] && (6.64)\\
&\geq -1 + \frac{1}{d^2}\sum_{ab} \operatorname{tr}\left[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,rel}\right] && (6.65)\\
&\geq -1 + \frac{1}{d^2}\sum_{ab} \operatorname{tr}\left[M_{ab}^{AB} + \widetilde{N}'_{ab}\right] && (6.66)\\
&\geq -1 + \frac{1}{d^2}\sum_{ab} \operatorname{tr}\left[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB}\right] = \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}). && (6.67)
\end{aligned}$$

where $\{\widetilde{N}'_{ab}\}$ is a set of dual variables feasible for our initial problem (6.58). In this way the conic program defining RoBN becomes:

$$\begin{aligned}
\mathbf{R}_{\mathrm{BN}}(\mathbb{M}) = \quad & \min \quad r && (6.68)\\
& s.t. \quad M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} \leq \widetilde{O}_{ab}^{AB} \quad \forall\, a, b, && (6.69)\\
& \sum_a \widetilde{O}_{ab}^{AB} = \mathbb{1}^A \otimes \widetilde{O}_b^B \quad \forall\, b, \qquad \sum_b \widetilde{O}_b^B = (1+r)\mathbb{1}^B, && (6.70)\\
& \sum_b \widetilde{O}_{ab}^{AB} = \widetilde{O}_a^A \otimes \mathbb{1}^B \quad \forall\, a, \qquad \sum_a \widetilde{O}_a^A = (1+r)\mathbb{1}^A, && (6.71)\\
& \{\widetilde{O}_{ab}^{AB}\} \in \mathcal{F}_{\mathrm{BN}} \qquad \forall\, a, b, \qquad O_{ab}^{AB} \in \mathsf{SEP}(A:B) \quad \forall\, a, b, && (6.72)\\
& \{\widetilde{N}_{ab}^{AB}\} \in \mathcal{R}_{\mathrm{BN}} \quad \forall\, a, b, && (6.73)
\end{aligned}$$

where the minimization is performed over r, $\{\widetilde{O}_{ab}^{AB}\}$, $\{\widetilde{O}_a^A\}$, $\{\widetilde{O}_b^B\}$ and $\{\widetilde{N}_{ab}^{AB}\}$.

In what follows we will denote a dual cone to $\mathcal{R}$ using $\mathcal{R}^*$, that is $\mathcal{R}^* := \{X \mid \operatorname{tr} XQ \geq 0 \text{ for all } Q \in \mathcal{R}\}$. We will now write the dual formulation of the above problem. To do so we first write the associated Lagrangian using dual Hermitian variables associated with a corresponding set of constraints: $\{A_{ab}^{AB}\}$ such that $A_{ab}^{AB} \geq 0$ for all a, b, $\{B_b^{AB}\}$, $\{C_a^{AB}\}$, $D^A \geq 0$, $E^B \geq 0$, $\{F_{ab}^{AB}\} \in \mathcal{F}_{\mathrm{BN}}^*$ meaning that $\sum_{ab} \operatorname{tr}\left[F_{ab}^{AB} X_{ab}^{AB}\right] \geq 0$ for all $\{X_{ab}^{AB}\} \in \mathcal{F}_{\mathrm{BN}}$, $G_{ab}^{AB} \in \mathcal{F}_{\mathrm{SEP}}^*$ for all a, b, meaning that $\operatorname{tr}\left[G_{ab}^{AB} X^{AB}\right] \geq 0$ for all a, b and all separable operators $X^{AB} \in \mathcal{F}_{\mathrm{SEP}}$ and, finally, $\{H_{ab}^{AB}\} \in \mathcal{R}_{\mathrm{BN}}^*$. With this the

Lagrangian function of the conic program (6.68—6.72) becomes:

$$\mathcal{L} = r + \sum_{ab} \operatorname{tr} A_{ab}^{AB} \left[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} - \widetilde{O}_{ab}^{AB} \right] + \sum_{b} \operatorname{tr} B_{b}^{AB} \left[\sum_{a} \widetilde{O}_{ab}^{AB} - \mathbb{1}^{A} \otimes \widetilde{O}_{b}^{B} \right] \tag{6.74}$$

$$+ \sum_{a} \operatorname{tr} C_{a}^{AB} \left[\sum_{b} \widetilde{O}_{ab}^{AB} - \widetilde{O}_{a}^{A} \otimes \mathbb{1}^{B} \right] + \operatorname{tr} D^{A} \left[\sum_{a} \widetilde{O}_{a}^{A} - (1+r)\mathbb{1}^{A} \right] \tag{6.75}$$

$$+ \operatorname{tr} E^{B} \left[\sum_{b} \widetilde{O}_{b}^{B} - (1+r)\mathbb{1}^{B} \right] - \sum_{a,b} \operatorname{tr} \left[F_{ab}^{AB} \widetilde{O}_{ab}^{AB} \right] - \sum_{a,b} \operatorname{tr} \left[G_{ab}^{AB} \widetilde{O}_{ab}^{AB} \right] \tag{6.76}$$

$$- \sum_{a,b} \operatorname{tr} \left[H_{ab}^{AB} \widetilde{N}_{ab}^{AB} \right] \tag{6.77}$$

$$= r \cdot \left[1 - \operatorname{tr} D^{A} - \operatorname{tr} E^{B} \right] + \sum_{a,b} \operatorname{tr} \widetilde{N}_{ab} \left[A_{ab}^{AB} - H_{ab}^{AB} \right] + \sum_{a} \operatorname{tr} O_{a}^{A} \left[D^{A} - C_{a}^{A} \right] \tag{6.78}$$

$$+ \sum_{a,b} \operatorname{tr} \widetilde{O}_{ab} \left[-A_{ab}^{AB} + B_{b}^{AB} + C_{a}^{AB} - F_{ab}^{AB} - G_{ab}^{AB} \right] + \sum_{b} \operatorname{tr} O_{b}^{B} \left[E^{B} - B_{b}^{B} \right] \tag{6.79}$$

$$+ \sum_{ab} \operatorname{tr} \left[A_{ab}^{AB} M_{ab}^{AB} \right] - \operatorname{tr} D^{A} - \operatorname{tr} E^{B}. \tag{6.80}$$

By demanding that the terms in the square brackets which appear along with the dual variables vanish we can ensure $\mathcal{L} \leq r$. This leads to the following (dual) conic program:

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \quad \max \quad \sum_{ab} \operatorname{tr} \left[A_{ab}^{AB} M_{ab}^{AB} \right] - 1 \tag{6.81}$$

$$s.t. \quad C_{a}^{AB} + B_{b}^{AB} = A_{ab}^{AB} + F_{ab}^{AB} + G_{ab}^{AB} \quad \forall\, a, b, \tag{6.82}$$

$$A_{ab}^{AB} = H_{ab}^{AB} \quad \forall\, a, b, \qquad C_{a}^{A} = D^{A} \quad \forall\, a, \tag{6.83}$$

$$B_{b}^{B} = E^{B} \quad \forall\, b, \qquad A_{ab}^{AB} \geq 0 \quad \forall\, a, b, \tag{6.84}$$

$$\{H_{ab}^{AB}\} \in \mathcal{R}_{\mathrm{BN}}^{*}, \quad \{F_{ab}^{AB}\} \in \mathcal{F}_{\mathrm{BN}}^{*}, \quad \operatorname{tr} D^{A} + \operatorname{tr} E^{B} = 1.$$

Notice now that the set $\mathcal{F}_{\mathrm{BN}} \in \mathcal{F}_{\mathrm{SEP}}$, which implies that the dual sets satisfy $\mathcal{F}_{\mathrm{SEP}}^{*} \in \mathcal{F}_{\mathrm{BN}}^{*}$. Hence without loss of generality we can assume $G_{ab}^{AB} = 0$ for all a and b. In this way we can express the above program in the following way:

$$1 + \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \quad \max \quad \sum_{ab} \operatorname{tr} \left[A_{ab}^{AB} M_{ab}^{AB} \right] \tag{6.85}$$

$$s.t. \quad C_{a}^{AB} + B_{b}^{AB} - A_{ab}^{AB} = F_{ab}^{AB} \in \mathcal{F}_{\mathrm{BN}}^{*} \quad \forall\, a, b, \tag{6.86}$$

$$C_{a}^{A} = D^{A} \quad \forall\, a, \qquad C_{a}^{A}, D^{A} \geq 0 \quad \forall\, a, \tag{6.87}$$

$$A_{ab}^{AB} \geq 0 \quad \forall\, a, b, \quad \operatorname{tr} D^{A} + \operatorname{tr} E^{B} = 1. \tag{6.88}$$

Using both primal (6.58) and dual (6.85) formulations we can now describe some basic properties of the RoBN.

6.5.2 Basic properties of RoBN

Here we prove the three basic properties of RoBN highlighted in the main text.

Faithfulness If $\mathbb{M}^{AB} \in \mathcal{F}_{\text{BN}}$ then we can always choose a feasible $r = 0$ in the primal form (6.58). Since the solution is always non-negative, $r = 0$ is also optimal.

Convexity Let $\{N^1_{ab}, O^1_{ab}\}$ be optimal primal variables for $\mathbf{R}_{\text{BN}}(\mathbb{M}^1)$ and similarly let $\{N^2_{ab}, O^2_{ab}\}$ be primal-optimal for $\mathbf{R}_{\text{BN}}(\mathbb{M}^2)$. Define $\mathbb{M}' = \{M'_{ab}\}$ as a convex combination of the two measurements, that is $M'_{ab} = p\, M^1_{ab} + (1-p)\, M^2_{ab}$ for each a and b. We can construct a set of feasible variables for $\mathbf{R}_{\text{BN}}(\mathbb{M}')$ in the following way: $N'_{ab} = p\, N^1_{ab} + (1-p)\, N^2_{ab}$ and $O'_{ab} = p\, O^1_{ab} + (1-p)\, O^2_{ab}$. Substituting N'_{ab} and O'_{ab} into the constraints of the primal form for $\mathbf{R}_{\text{BN}}(\mathbb{M})$ shows that this choice is feasible. In this way we obtain an upper bound on $\mathbf{R}_{\text{BN}}(\mathbb{M}')$:

$$\mathbf{R}_{\text{BN}}(\mathbb{M}') \leq \text{tr} \sum_{a,b} N'_{ab} \tag{6.89}$$

$$= p\ \text{tr} \sum_{a,b} N^1_{ab} + (1-p)\ \text{tr} \sum_{a,b} N^2_{ab} \tag{6.90}$$

$$= p\mathcal{R}(\mathbb{M}^1) + (1-p)\mathcal{R}(\mathbb{M}^2). \tag{6.91}$$

Monotonicity Let us start with the assumption that there is a subroutine:

$$\mathcal{S} = \{p(\lambda), p(a|i,\lambda), p(b|j,\lambda), \mathcal{E}_\lambda, \mathcal{N}_\lambda\}$$

which allows to simulate $\mathbb{M}'$ using $\mathbb{M}$, i.e. $\mathbb{M} \succ_q \mathbb{M}'$. This means that the POVM elements $\{M_{ab}\}$ of $\mathbb{M}$ can be mapped into:

$$M'_{ab} = \sum_{i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(\mathcal{E}^\dagger_\lambda \otimes \mathcal{N}^\dagger_\lambda)[M_{ij}].$$

Suppose now that we solved the dual problem for $\mathbf{R}_{\text{BN}}(\mathbb{M}')$ using the optimal dual variables $\{A'_{ab}\}$, $\{B'_b\}$,$\{C'_a\}$, D', E' and $\{F'_{ab}\}$. Using these we construct an educated guess for $\mathbf{R}_{\text{BN}}(\mathbb{M})$ in the following way:

$$A^*_{ij} = \sum_{a,b,\lambda} p(\lambda)\, p(a|i,\lambda)p(b|j,\lambda)\ (\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)\,[A'_{ab}], \tag{6.92}$$

$$B^*_j = \sum_{b,\lambda} p(\lambda)\, p(b|j,\lambda)\ (\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)\,[B'_b], \tag{6.93}$$

$$C^*_i = \sum_{a,\lambda} p(\lambda)\, p(a|i,\lambda)\ (\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)\,[C'_a], \tag{6.94}$$

$$D^* = \sum_{\lambda} p(\lambda)\, \mathcal{E}_\lambda[D'], \tag{6.95}$$

$$E^* = \sum_{\lambda} p(\lambda)\, \mathcal{N}_\lambda[E'], \tag{6.96}$$

$$F^*_{ij} = \sum_{a,b,\lambda} p(\lambda)\, p(a|i,\lambda)p(b|j,\lambda)\ (\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)\,[F'_{ab}]. \tag{6.97}$$

It can be verified that the above choice of variables is feasible for the dual problem (6.85). In particular, notice that by construction we have $C^*_i + B^*_j - A^*_{ij} = F^*_{ij}$ for all i, j since the primed dual variables satisfy the constraints of (6.85). Furthermore, since $\text{tr}_B(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[X^{AB}] = \mathcal{E}_\lambda[X^A]$ we can infer that $\text{tr}_B\, C^*_i = D^*$ and $\text{tr}_A\, B^*_j = E^*$. Moreover, as separable maps preserve both positivity and separability we also have

that $A^*_{ij} \geq 0$ for all i, j and $\{F^*_{ij}\} \in \mathcal{F}^*_{\text{BN}}$. Using the proposed set of dual variables we find the following lower bound:

$$1 + \mathbf{R}_{\text{BN}}(\mathbb{M}) \geq \sum_{i,j} \text{tr}\Big[M_{ij} A^*_{ij}\Big] \tag{6.98}$$

$$= \sum_{a,b,i,j,\lambda} p(\lambda)\, p(a|i,\lambda) p(b|j,\lambda)\ \text{tr}\big[M_{ij} \cdot (\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)\,[A_{ab}]\big] \tag{6.99}$$

$$= \sum_{a,b,i,j,\lambda} p(\lambda)\, p(a|i,\lambda) p(b|j,\lambda)\ \text{tr}\big[\big(\mathcal{E}^\dagger_\lambda \otimes \mathcal{N}^\dagger_\lambda\big)\,[M_{ij}] \cdot A_{ab}\big] \tag{6.100}$$

$$= \sum_{a,b} \text{tr}[M_{ab} A_{ab}] \tag{6.101}$$

$$= 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}'). \tag{6.102}$$

This proves that RoBN is monotonic under quantum simulation.

6.5.3 Proof of Result 1

In this section we prove that RoBN can be seen as a quantifier of the advantage a given distributed measurement provides in the task of distributed state discrimination. To simplify notation in this section we shall omit subsystem labels whenever it is clear from the context. Let us recall that the average guessing probability in the task of distributed state discrimination using a distributed measurement $\mathbb{M}$ can be expressed as:

$$p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}) = \max_{\mathbb{N} \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[N_{ab}\, \sigma_{xy}\right] \delta_{xa} \delta_{yb}, \tag{6.103}$$

where the optimization ranges over all measurements $\mathbb{N} = \{N_{ab}\}$ which can be quantum-simulated using $\mathbb{M} = \{M_{ij}\}$, where

$$M_{ij} = \text{tr}_{AB}\left[(M^{A'A}_i \otimes M^{BB'}_j)(\mathbb{1}^{A'} \otimes \rho^{AB} \otimes \mathbb{1}^{B'})\right] \tag{6.104}$$

is a distributed measurement and $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ is an ensemble of bipartite states. Suppose that we have solved the dual problem for RoBN (6.85) using the set of dual variables $\{A_{ab}\}$, $\{C_a\}$, $\{B_b\}$, D, E and $\{G_{ab}\}$. Notice also that due to the constraints in (6.85) the matrix A_{ab} is positive semi-definite for all values of a and b. Let us now consider a particular game setting $\mathcal{G}^* = \{p^*(x,y), \sigma^*_{xy}\}$ defined in the following way:

$$C = \sum_{x,y} \text{tr}\, A_{xy}, \qquad p^*(x,y) = \frac{\text{tr}\, A_{xy}}{C}, \qquad \sigma^*_{xy} = \frac{A_{xy}}{\text{tr}\, A_{xy}}, \tag{6.105}$$

where $x = 1, \ldots, o_A$, $y = 1, \ldots, o_B$ and o_A, o_B are the numbers of outcomes of local measurements performed by A and B. The best average guessing probability which

can be achieved in the game $\mathcal{G}^*$ using a distributed measurement $\mathbb{M}$ is given by:

$$p^{DSD}_{\text{guess}}(\mathcal{G}^*, \mathbb{M}) = \max_{\mathbb{N} \prec_q \mathbb{M}} \sum_{a,b,x,y} p^*(x,y)\ \text{tr}\left[N_{ab}\,\sigma^*_{xy}\right]\,\delta_{xa}\delta_{yb} \tag{6.106}$$

$$\geq \sum_{x,y} \frac{\text{tr}\,A_{xy}}{C} \cdot \text{tr}\left[M_{xy}\,\frac{A_{xy}}{\text{tr}\,A_{xy}}\right] \tag{6.107}$$

$$= \frac{1}{C}\sum_{x,y} \text{tr}\left[M_{xy}A_{xy}\right] \tag{6.108}$$

$$= \frac{1}{C}\left[1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB})\right], \tag{6.109}$$

where the inequality in the second line we follows from choosing a particular subroutine $\mathcal{S}$ with $p(\lambda) = 1/|\lambda|$, $p(a|i,\lambda) = \delta_{ai}$, $p(b|j,\lambda) = \delta_{bj}$ and $\mathcal{E}_\lambda = \mathcal{N}_\lambda = id$. Let us now look at the corresponding classical (i.e. without access to entanglement) probability of guessing:

$$p^{DSD}_{\text{guess}}(\mathcal{G}^*) \tag{6.110}$$

$$= \max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} p^{DSD}_{\text{guess}}(\mathcal{G}^*, \mathbb{N}) \tag{6.111}$$

$$= \max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} \sum_{x,y} p^*(x,y)\ \text{tr}\left[N_{xy}\,\sigma^*_{xy}\right] \tag{6.112}$$

$$= \frac{1}{C}\max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} \sum_{x,y} \text{tr}\left[N_{xy}A_{xy}\right] \tag{6.113}$$

$$= \frac{1}{C}\max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} \sum_{x,y} \text{tr}\left[N_{xy}(C_x + B_y - F_{xy})\right] \tag{6.114}$$

$$= \frac{1}{C}\max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} \left(\sum_x \text{tr}\left[(N_x \otimes \mathbb{1})C_x\right] + \sum_y \text{tr}\left[(\mathbb{1}\otimes N_y)B_y\right] - \sum_{x,y}\text{tr}\left[N_{xy}F_{xy}\right]\right) \tag{6.115}$$

$$\leq \frac{1}{C}\max_{\mathbb{N}\in\mathcal{F}_{\text{BN}}} \left(\sum_x \text{tr}\left[N_x D\right] + \sum_y \text{tr}\left[N_y E\right]\right) \tag{6.116}$$

$$= \frac{1}{C}\left(\text{tr}\,D + \text{tr}\,E\right) \tag{6.117}$$

$$= \frac{1}{C}, \tag{6.118}$$

where the inequality follows since for all $\mathbb{N} \in \mathcal{F}_{\text{BN}}$ we have $\sum_{xy} \text{tr}\left[N_{xy}F_{xy}\right] \geq 0$. Combining bounds (6.109) and (6.118) leads to:

$$\max_{\mathcal{G}} \frac{p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M})}{p^{DSD}_{class}(\mathcal{G})} \geq \frac{p^{DSD}_{\text{guess}}(\mathcal{G}^*, \mathbb{M})}{p^{DSD}_{class}(\mathcal{G}^*)} \geq 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}). \tag{6.119}$$

In order to prove the upper bound notice that the first line of constraints in the primal formulation for RoBN (6.58) implies:

$$\forall\, a, b \qquad M'_{ab} = \widetilde{O}'_{ab} - \widetilde{N}'_{ab}, \tag{6.120}$$

where $\widetilde{O}'_{ab} = [1 + \mathbf{R}_{\text{BN}}(\mathbb{M})]\, O'_{ab}$ for all a, b and $\{O'_{ab}\} \in \mathcal{F}_{\text{BN}}$. This allows to write:

$$
\begin{aligned}
& p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}) && (6.121)\\
&= \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[M'_{ab}\, \sigma_{xy}\right]\ \delta_{xa}\delta_{yb} && (6.122)\\
&= \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[(\widetilde{O}'_{ab} - \widetilde{N}'_{ab})\sigma_{xy}\right]\ \delta_{xa}\delta_{yb} && (6.123)\\
&\leq \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[\widetilde{O}'_{ab}\sigma_{xy}\right]\ \delta_{xa}\delta_{yb} && (6.124)\\
&= \max_{\mathbb{M}' \prec_q \mathbb{M}} [1 + \mathbf{R}_{\text{BN}}(\mathbb{M}')] \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[O'_{ab}\sigma_{xy}\right]\ \delta_{xa}\delta_{yb} && (6.125)\\
&\leq \left(\max_{\mathbb{M}' \prec_q \mathbb{M}} [1 + \mathbf{R}_{\text{BN}}(\mathbb{M}')]\right)\left(\max_{\{O_{ab}\} \in \mathcal{F}_{\text{BN}}} \sum_{a,b,x,y} p(x,y)\ \text{tr}\left[O_{ab}\sigma_{xy}\right]\ \delta_{xa}\delta_{yb}\right) && (6.126)\\
&\leq [1 + \mathbf{R}_{\text{BN}}(\mathbb{M})]\, p^{DSD}_{\text{guess}}(\mathcal{G}), && (6.127)
\end{aligned}
$$

where the last inequality follows from the monotonicity of RoBN under quantum simulation. Combining bounds (6.119) and (6.127) yields:

$$
\max_{\mathcal{G}} \frac{p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M})}{p^{DSD}_{\text{guess}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}). \tag{6.128}
$$

6.5.4 Proof of Result 2

Before proving the result we recall the primal and dual formulation of the RoT quantifier. Let $\mathbb{\Lambda} = \{\Lambda_a\}$ be a teleportation instrument whose elements are defined as:

$$
\Lambda_a^{A\to B'}[\omega] := \text{tr}_{AA'}[(M_a^{AA'} \otimes \mathbb{1}^B)(\omega^A \otimes \rho^{A'B'})], \tag{6.129}
$$

for some measurement $M_a^{AA'}$ and a shared state $\rho^{A'B'}$. We denote the set of Choi-Jamiołkowski states corresponding to this of these subchannels with $\{J_a^{VB'}\}$, i.e. each $J_a^{VB'} := (\text{id}^V \otimes \Lambda_a^{A\to B'})[\phi_+^{VA}]$ with system V isomorphic to A. With these definitions RoT for a teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ can be written as:

$$
\begin{aligned}
\mathbf{R}_{\text{T}}(\mathbb{\Lambda}^{A\to B'}) = \min \quad & \text{tr}\,\widetilde{\sigma}^{B'}, && \iff \quad \max \quad \sum_a \text{tr}\left[A_a^{VB'} J_a^{VB'}\right] - 1,\\
s.t. \quad & J_a^{VB'} \leq F_a^{VB'} \quad \forall\, a, && s.t. \quad B^{VB'} - A_a^{VB'} = W_a^{VB'}\\
& \sum_a F_a^{VB'} = \frac{\mathbb{1}^V}{d} \otimes \widetilde{\sigma}^{B'}, && \{W_a^{VB'}\} \in \mathcal{F}^*_{\text{SEP}},\\
& \{F_a^{VB'}\} \in \mathcal{F}_{\text{T}} \quad \forall a, && B^{B'} = \mathbb{1}^{B'},\\
& \widetilde{\sigma}^{B'} \geq 0. && A_a^{VB'} \geq 0 \quad \forall a.
\end{aligned} \tag{6.130}
$$

Let us now proceed with the proof of Result 2.

Proof. As before, the proof consists of two steps. First we will show that $\mathbf{R}_{\text{T}}(\mathbb{\Lambda}^{A\to B'})$ lower bounds $\mathbf{R}_{\text{BN}}(\mathbb{M}^{AB})$ for a particular choice of local measurement $\mathbb{M}^{B'B}$. Then we

will show that for any choice of local measurements on Bob's side $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})$ is never larger than the teleportation quantifier $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})$.

Let $A_a^{VB'} \geq 0$, $W_a^{VB'} \in \mathcal{F}^*_{\mathrm{SEP}}$ and $B^{VB'}$ be optimal dual variables for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})$. Let $\{U_b^B\}$ for $b \in \{1,\ldots,d^2\}$ be a set of Pauli operators with respect to a basis $\{|i\rangle^B\}$. Consider the following measurement with $o_B = d^2$ outcomes:

$$M_b^{B'B} = (\mathrm{id}^{B'} \otimes \mathcal{U}_b^B)[\phi_+^{B'B}], \tag{6.131}$$

where $\mathcal{U}_b[\cdot] := U_b(\cdot)U_b^\dagger$. We are interested in the lower bound for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})$. Let us choose a set of dual variables in (6.85) inspired by the optimal dual variables for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{V\to B'})$:

$$A_{ab}^{AB} = (\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(A_a^{AB})^T], \quad B_b^{AB} = \frac{1}{d}(\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(B^{AB})^T], \quad C_a^{AB} = 0, \tag{6.132}$$

$$F_{ab}^{AB} = (\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(W_a^{AB})^T], \quad D^B = \frac{1}{d}\mathbb{1}^B, \qquad E^A = 0. \tag{6.133}$$

It can be verified by direct substitution that the above choice is feasible. In particular, the above choice for $\{F_{ab}^{AB}\}$ is feasible as $\mathcal{F}^*_{\mathrm{SEP}} \in \mathcal{F}^*_{\mathrm{BN}}$ and both sets are invariant under local unitaries. This leads to the following chain of inequalities:

$$1 + \max_{\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) \tag{6.134}$$

$$\geq \sum_{ab} \mathrm{tr}\left[A_{ab}^{AB} M_{ab}^{AB}\right] \tag{6.135}$$

$$= \sum_{ab} \mathrm{tr}\left[(\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(A_a^{AB})^T]\,\mathrm{tr}_{A'B'}\left[(M_a^{AA'} \otimes M_b^{B'B})(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B)\right]\right] \tag{6.136}$$

$$= \sum_{ab} \mathrm{tr}\left[(\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(A_a^{AB})^T]\,\mathrm{tr}_{A'B'}\left[(M_a^{AA'} \otimes (\mathrm{id}^{B'} \otimes \mathcal{U}_b^B)[\phi_+^{B'B}])(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B)\right]\right] \tag{6.137}$$

$$= \sum_{ab} \mathrm{tr}\left[\mathbb{1}^{A'B'} \otimes (\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B)[(A_a^{AB})^T]\left[(M_a^{AA'} \otimes (\mathrm{id}^{B'} \otimes \mathcal{U}_b^B)[\phi_+^{B'B}])(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B)\right]\right] \tag{6.138}$$

$$= \sum_{ab} \mathrm{tr}\left[\left(\mathbb{1}^{A'B'} \otimes (A_a^{AB})^T\right)\left(M_a^{AA'} \otimes \phi_+^{B'B}\right)\left(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B\right)\right] \tag{6.139}$$

$$= \frac{1}{d^2}\sum_{ab} \mathrm{tr}\left[A_a^{VB'}\,\mathrm{tr}_{AA'}\left[(\mathbb{1}^V \otimes M_a^{AA'} \otimes \mathbb{1}^{B'})(\phi_+^{VA} \otimes \rho^{A'B'})\right]\right] \tag{6.140}$$

$$= \sum_a \mathrm{tr}\left[A_a^{VB'} J_a^{VB'}\right] \tag{6.141}$$

$$= 1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{V\to B'}). \tag{6.142}$$

We now prove the upper bound. Notice that for any distributed measurement $\mathbb{M}^{AB}$ we can construct $\mathbb{M}^{VB} := \{M_{ab}^{VB}\}$ such that $M_{ab}^{VB} := d\ \mathrm{tr}_A[(\mathbb{1}^V \otimes M_{ab}^{AB})(\phi_+^{VA} \otimes \mathbb{1}^B)]$. This in turn can be written as:

$$M_{ab}^{VB} := d\ \mathrm{tr}_{AA'B'}\left[\left(\mathbb{1}^A \otimes M_a^{AA'} \otimes M_b^{B'B}\right)\left(\phi_+^{VA} \otimes \rho^{A'B'} \otimes \mathbb{1}^B\right)\right] \tag{6.143}$$

$$= d\ \mathrm{tr}_{B'}\left[\left(\mathbb{1}^V \otimes M_b^{B'B}\right)\left(J_a^{VB'} \otimes \mathbb{1}^B\right)\right]. \tag{6.144}$$

Note that we can always write $J_a^{VB'} \leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})]F_a^{VB'}$, where $\{F_a^{VB'}\}$ are Choi-Jamiołkowski operators of some classical teleportation instrument. This allows us to further rewrite (6.144) as:

$$M_{ab}^{VB} \leq d\,[1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})]\ \mathrm{tr}_A[(\mathbb{1}^V \otimes M_{ab}^{AB})(\phi_+^{VA} \otimes \mathbb{1}^B)] \tag{6.145}$$

$$= [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})]\, O_{ab}^{VB}. \tag{6.146}$$

Where $\{O_{ab}^{VB}\}$ is a free distributed measurement. Hence also $M_{ab}^{AB} \leq [1+\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})]\, O_{ab}^{AB}$ for some free distributed measurement $\{O_{ab}^{AB}\}$. This finally allows us to write:

$$\max_{\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) \leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})] \max_{\mathbb{M}^B} \sum_{ab} \mathrm{tr}\left[A_{ab}^{AB} O_{ab}^{AB}\right] \leq [1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})]. \tag{6.147}$$

This proves the lemma. □

6.5.5 Proof of Result 4

Let us recall that the conic program formulation of RoE is given by:

$$\begin{aligned} \mathbf{R}_{\mathrm{E}}(\rho^{A'B'}) = \min \quad & \mathrm{tr}\,\widetilde{\sigma}^{A'B'}, \\ s.t. \quad & \rho^{A'B'} \leq \widetilde{\sigma}^{A'B'} \\ & \widetilde{\sigma}^{A'B'} \in \mathcal{F}_{\mathrm{SEP}}. \end{aligned} \qquad \Longleftrightarrow \qquad \begin{aligned} \max \quad & \sum_a \mathrm{tr}\left[A^{A'B'}\rho^{A'B'}\right] - 1, \\ s.t. \quad & \mathbb{1}^{A'B'} - A^{A'B'} = W^{A'B'} \in \mathcal{F}^*_{\mathrm{SEP}}, \\ & A^{A'B'} \geq 0. \end{aligned} \tag{6.148}$$

The proof is based on three parts. First we use Result 2 to connect RoBN with RoT. Then we essentially parallel the steps taken in the proof of Result 2 to link RoT with RoE. It is worth mentioning that the link between RoT and RoE has already been obtained some time ago in (Cavalcanti, Skrzypczyk, and Supic, 2017). Here for convenience we state an independent proof.

Proof. Let us begin by noting that Result 2 implies:

$$\max_{\mathbb{M}^A,\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \max_{\mathbb{M}^A}\left[\max_{\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})\right] = \max_{\mathbb{M}^A} \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'}) \tag{6.149}$$

Let $A^{A'B'} \geq 0$, $W^{A'B'} \in \mathcal{F}^*_{\mathrm{SEP}}$ be optimal dual variables for $\mathbf{R}_{\mathrm{E}}(\rho^{A'B'})$. Let $\{U_a^{A'}\}$ for $a \in \{1, \ldots, d^2\}$ be a set of Pauli operators with respect to a basis $\{|i\rangle^{A'}\}$. Consider the following measurement with $o_A = d^2$ outcomes:

$$M_a^{AA'} = (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}]. \tag{6.150}$$

We are interested in the lower bound for $\mathbf{R}_{\mathrm{E}}(\rho^{A'B'})$, let us construct a set of (potentially sub-optimal) dual variables in the maximization (6.148) using the optimal set of dual variables for $\mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{V\to B'})$, i.e.:

$$A_a^{VB'} = ((\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'})[A^{VB'}],\ W_a^{VB'} = ((\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'})[(W^{VB'}],\ B^{VB'} = \frac{1}{d}\mathbb{1}^{VB'}. \tag{6.151}$$

It can be verified by direct substitution that the above choice is feasible. This leads to the following chain of inequalities:

$$1 + \max_{\mathbb{M}^A} \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'}) \tag{6.152}$$

$$\geq \sum_a \mathrm{tr}\left[A_a^{VB'} J_a^{VB'}\right] \tag{6.153}$$

$$= \sum_a \mathrm{tr}\left[((\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'})[A_a^{VB'}]\,\mathrm{tr}_{AA'}\left[(\mathbb{1}^V \otimes M_a^{AA'} \otimes \mathbb{1}^{B'})(\phi_+^{VA} \otimes \rho^{A'B'})\right]\right] \tag{6.154}$$

$$= \sum_a \mathrm{tr}\left[((\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'})[A_a^{VB'}]\,\mathrm{tr}_{AA'}\left[(\mathbb{1}^V \otimes (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}] \otimes \mathbb{1}^{B'})(\phi_+^{VA} \otimes \rho^{A'B'})\right]\right] \tag{6.155}$$

$$= \sum_a \mathrm{tr}\left[\mathbb{1}^{AA'} \otimes ((\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'})[A_a^{VB'}](\mathbb{1}^V \otimes (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}] \otimes \mathbb{1}^{B'})(\phi_+^{VA} \otimes \rho^{A'B'})\right] \tag{6.156}$$

$$= \frac{1}{d^2} \sum_a \mathrm{tr}\left[((\mathcal{U}_a^\dagger)^{A'} \otimes \mathrm{id}^{B'})A^{A'B'}(\mathcal{U}_a^{A'} \otimes \mathrm{id}^{B'})\rho^{A'B'}\right] \tag{6.157}$$

$$= \mathrm{tr}\left[A^{A'B'}\rho^{A'B'}\right] \tag{6.158}$$

$$= 1 + \mathbf{R}_{\mathrm{E}}(\rho^{A'B'}). \tag{6.159}$$

We now prove the upper bound. Notice that any teleportation instrument $\mathbb{\Lambda}^{A\to B'}$ expressed using Choi-Jamiołkowski operators $\{J_a^{VB'}\}$ satisfies:

$$J_a^{VB'} := \mathrm{tr}_{VA}\left[\left(M_a^{VA} \otimes \mathbb{1}^{B'}\right)\left(\phi_+^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B\right)\right] \tag{6.160}$$

$$\leq [1 + \mathbf{R}_{\mathrm{E}}(\rho^{A'B'})]\,\mathrm{tr}_{VA}\left[\left(M_a^{VA} \otimes \mathbb{1}^{B'}\right)\left(\phi_+^A \otimes \sigma^{A'B'} \otimes \mathbb{1}^B\right)\right] \tag{6.161}$$

$$= [1 + \mathbf{R}_{\mathrm{E}}(\rho^{A'B'})]O_a^{VB'}, \tag{6.162}$$

for some state $\sigma^{A'B'} \in \mathcal{F}_{\mathrm{SEP}}$ and corresponding (classical) teleportation operators $\{O_a^{VB'}\}$. In this way we can write:

$$\max_{\mathbb{M}^A}[1 + \mathbf{R}_{\mathrm{T}}(\mathbb{\Lambda}^{A\to B'})] = \max_{\mathbb{M}^A} \max_{\{A_a^{VB'}\}} \sum_a \mathrm{tr}\left[A_a^{VB'} J_a^{VB'}\right] \tag{6.163}$$

$$\leq [1 + \mathbf{R}_{\mathrm{E}}(\rho^{A'B'})] \sum_a \mathrm{tr}\left[A_a^{VB'} O_a^{VB'}\right] \tag{6.164}$$

$$\leq [1 + \mathbf{R}_{\mathrm{E}}(\rho^{A'B'})]. \tag{6.165}$$

This proves the lemma. □

6.5.6 Proof of Result 6

In this section, unless explicitly specified, all bipartite operators act on subsystems A and B. We begin by assuming that a distributed measurement $\mathbb{M}$ can be used to

simulate $\mathbb{M}^*$, that is $\mathbb{M} \succ_q \mathbb{M}^*$. We have:

$$p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}) = \max_{\mathbb{M} \succ_q \mathbb{M}'} \sum_{a,b} p(a,b)\ \text{tr}\left[M'_{ab}\,\sigma_{ab}\right] \tag{6.166}$$

$$\geq \max_{\mathbb{M}^* \succ_q \mathbb{M}'} \sum_{a,b} p(a,b)\ \text{tr}\left[M'_{ab}\,\sigma_{ab}\right] \tag{6.167}$$

$$= p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^*), \tag{6.168}$$

since the set $\{\mathbb{M}' | \mathbb{M}^* \succ_q \mathbb{M}'\}$ is a subset of $\{\mathbb{M}' | \mathbb{M} \succ_q \mathbb{M}'\}$. Now we are going to assume that $p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}) \geq p^{DSD}_{\text{guess}}(\mathcal{G}, \mathbb{M}^*)$ holds for all games $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ and show show that there always exist a subroutine $\mathcal{S}$ which allows to simulate $\mathbb{M}^*$ using $\mathbb{M}$. We thus have:

$$\forall\, \mathcal{G} \qquad \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b} p(a,b)\ \text{tr}\left[M'_{ab}\,\sigma_{ab}\right] - \max_{\mathbb{M}'' \prec_q \mathbb{M}^*} \sum_{a,b} p(a,b)\ \text{tr}\left[M''_{ab}\,\sigma_{ab}\right] \geq 0. \tag{6.169}$$

Let us now choose a particular subroutine in the second maximization, i.e.: $\mathcal{S}^* = \{p(\lambda) = \delta_{\lambda 0},\ p(a|i,\lambda) = \delta_{ai},\ p(b|j,\lambda) = \delta_{bj},\ U_\lambda = V_\lambda = \mathbb{1}\}$. In this way (6.169) implies:

$$\forall\, \mathcal{G} \qquad \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b} p(a,b)\ \text{tr}\left[(M'_{ab} - M^*_{ab})\,\sigma_{ab}\right] \geq 0. \tag{6.170}$$

Let us denote $\Delta_{ab} := M'_{ab} - M^*_{ab}$. Since both M'_{ab} and M^*_{ab} are measurements we have that $\sum_{a,b} \Delta_{ab} = 0$. This also means that only one of the two situations can hold: either (i) $\Delta_{ab} = 0$ for all a, b or (ii) there exists at least one Δ_{ab} with at least one negative eigenvalue.

We will now show by contradiction that (ii) cannot be true. Let us assume that (ii) holds and label the negative eigenvalue with $\lambda_{a^*b^*}$ and the associated eigenvector with $|\lambda_{a^*b^*}\rangle$. Then, since (6.170) holds for all games $\mathcal{G}$, it also holds for a particular game $\mathcal{G}^* = \{p(a,b) = \delta_{aa^*}\delta_{bb^*}, \sigma_{ab} = |\lambda_{a^*b^*}\rangle\langle\lambda_{a^*b^*}|\}$. Hence (6.170) implies:

$$\langle\lambda_{a^*b^*}|\,\Delta_{a^*b^*}\,|\lambda_{a^*b^*}\rangle = \lambda_{a^*b^*} < 0, \tag{6.171}$$

which is a contradiction. Hence we infer that (ii) cannot be true and the only possibility is that each operator Δ_{ab} is identically zero. This means that:

$$M^*_{ab} = M'_{ab} := \sum_{i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(U^\dagger_\lambda \otimes V^\dagger_\lambda)M_{ij}(U_\lambda \otimes V_\lambda), \tag{6.172}$$

i.e. $\mathbb{M}^*$ can be simulated using $\mathbb{M}$.

6.5.7 Proof of Result 7

The accessible min-information $I^{\text{acc}}_{min}(\mathcal{N})$ of a channel $\mathcal{N}$ is defined as (Wilde, 2013):

$$I^{\text{acc}}_{\text{min}}(\mathcal{N}) = \max_{\mathcal{E},\mathcal{D}} \left[H_{\text{min}}(X) - H_{\text{min}}(X|G)\right], \tag{6.173}$$

where the optimization is over all encodings $\mathcal{E} = \{p(x), \sigma_x\}$ and decodings $\mathcal{D} = \{D_g\}$ and the min-entropies are defined as:

$$H_{\min}(X) = -\log \max_x p(x), \tag{6.174}$$

$$H_{\min}(X|G) = -\log \left[\sum_g \max_x p(x, g) \right], \tag{6.175}$$

and $p(x, g)$ is the probability distribution induced by channel $\mathcal{N}$, i.e.:

$$p(x, g) = p(x)p(g|x) = p(x) \operatorname{tr} \left[\mathcal{N}[\sigma_x] D_g \right]. \tag{6.176}$$

Consider now encoding a bipartite random variable $X \times Y$ in an ensemble of bipartite quantum states, i.e.: $\mathcal{E} = \{p(x, y), \sigma_{xy}^{AB}\}$ and $\mathcal{D} = \{D_g^{A'B'}\}$ for $g = 1, \ldots, o_A \cdot o_B$. Moreover, consider the channel $\mathcal{N} = \mathcal{N}^{AB \to A'B'}$ to be a quantum-to-classical measurement channel, which can be written as:

$$\mathcal{N}^{AB \to A'B'}(\rho^{AB}) = \sum_{a,b} \operatorname{tr} \left[M_{ab}^{AB} \rho^{AB} \right] |a\rangle\langle a|_{A'} \otimes |b\rangle\langle b|_{B'}, \tag{6.177}$$

where $\mathbb{M} = \{M_{ab}^{AB}\}$ is a distributed measurement. We have:

$$I_{\min}^{\text{acc}}(\mathcal{N}^{AB \to A'B'}) \tag{6.178}$$

$$= \max_{\mathcal{E}, \mathcal{D}} \log \left[\sum_g \max_{x,y} p(x, y) \operatorname{tr} \left[\mathcal{N}^{AB \to A'B'}[\sigma_{xy}^{AB}] D_g^{A'B'} \right] \right] - \log \max_{a,b} p(a, b) \tag{6.179}$$

$$= \max_{\mathcal{E}, \mathcal{D}} \log \left[\sum_g \sum_{a,b} \max_{x,y} p(x, y) \operatorname{tr} \left[M_{ab}^{AB} \sigma_{xy}^{AB} \right] \operatorname{tr} \left[D_g^{A'B'} |a\rangle\langle a|_{A'} \otimes |b\rangle\langle b|_{B'} \right] \right] \tag{6.180}$$

$$- \log \max_{a,b} p(a, b) \tag{6.181}$$

$$= \log \left[\sum_{a,b} \max_{\mathcal{E}} \max_{x,y} p(x, y) \operatorname{tr} \left[M_{ab}^{AB} \sigma_{xy}^{AB} \right] \right] - \log \max_{a,b} p(a, b). \tag{6.182}$$

Notice now that we can always express the optimization over (x, y) as:

$$\max_{x,y} p(x, y) \operatorname{tr} \left[M_{ab}^{AB} \sigma_{xy}^{AB} \right] \tag{6.183}$$

$$= \max_{p(x|a)} \max_{p(y|b)} \sum_{x,y} p(x|a)p(y|b)p(x, y) \operatorname{tr} \left[M_{ab}^{AB} \sigma_{xy}^{AB} \right] \tag{6.184}$$

$$= \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a, \lambda)p(y|b, \lambda)p(x, y) \operatorname{tr} \left[M_{ab}^{AB} \sigma_{xy}^{AB} \right] \tag{6.185}$$

Notice further that if we carry out the optimisation of the above expression over $\mathcal{E}$ we can additionally write:

$$\max_{\mathcal{E}} \max_{x,y} p(x,y) \operatorname{tr}\left[M_{ab}^{AB} \sigma_{xy}^{AB}\right] \tag{6.186}$$

$$= \max_{\mathcal{E}} \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a,\lambda) p(y|b,\lambda) p(x,y) \operatorname{tr}\left[M_{ab}^{AB} \sigma_{xy}^{AB}\right] \tag{6.187}$$

$$= \max_{\mathcal{E}} \max_{\{\mathcal{E}_\lambda\},\{\mathcal{F}_\lambda\}} \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a,\lambda) p(y|b,\lambda) p(x,y) \times \tag{6.188}$$

$$\operatorname{tr}\left[M_{ab}^{AB} (\mathcal{E}_\lambda^A \otimes \mathcal{F}_\lambda^B)(\sigma_{xy}^{AB})\right] \tag{6.189}$$

$$= \max_{\mathcal{E}} \max_{\mathbb{N} \prec \mathbb{M}} \sum_x p(x,y) \operatorname{tr}\left[N_{ab}^{AB} \sigma_{xy}^{AB}\right] . \tag{6.190}$$

Hence we can further continue from (6.182) and write:

$$I_{\min}^{\text{acc}}(\mathcal{N}^{AB \to A'B'}) = \log \left[\sum_{a,b} \max_{\mathcal{E}} \max_{\mathbb{N} \prec \mathbb{M}} p(a,b) \operatorname{tr}\left[N_{ab}^{AB} \sigma_{ab}^{AB}\right] \right] - \log \max_{a,b} p(a,b) \tag{6.191}$$

$$= \max_{\mathcal{E}} \log \left[\max_{\mathbb{N} \prec \mathbb{M}} \sum_{a,b} p(a,b) \operatorname{tr}\left[M_{ab}^{AB} \sigma_{ab}^{AB}\right] \right] - \max_{a,b} p(a,b) \tag{6.192}$$

$$= \max_{\mathcal{E}} \log \left[p_{guess}^{DSD}(\mathcal{G}, \mathbb{M}^{AB}) \right] - \log \left[p_{guess}^{DSD}(\mathcal{G}) \right] \tag{6.193}$$

$$= \log \left[\max_{\mathcal{E}} \frac{p_{guess}^{DSD}(\mathcal{G}, \mathbb{M}^{AB})}{p_{guess}(\mathcal{G})} \right] \tag{6.194}$$

$$= \log \left[1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB})\right] . \tag{6.195}$$

Bibliography

Åberg, Johan (2014). "Catalytic Coherence". In: *Phys. Rev. Lett.* 113 (15). URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.150402.

— (2018). "Fully Quantum Fluctuation Theorems". In: *Phys. Rev. X* 8 (1). URL: https://link.aps.org/doi/10.1103/PhysRevX.8.011019.

Åberg, Johan (2013). "Truly work-like work extraction via a single-shot analysis". In: *Nat. Commun.* 4. URL: https://doi.org/10.1038/ncomms2712.

Acín, A. (2001). "Statistical Distinguishability between Unitary Operations". In: *Phys. Rev. Lett.* 87 (17). URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.177901.

Acin, Antonio, Nicolas Gisin, and Lluis Masanes (2006). "From Bell's Theorem to Secure Quantum Key Distribution". In: *Phys. Rev. Lett.* 97 (12). URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.120405.

Alhambra, Álvaro M. et al. (2016). "Fluctuating Work: From Quantum Thermodynamical Identities to a Second Law Equality". In: *Phys. Rev. X* 6 (4). URL: https://link.aps.org/doi/10.1103/PhysRevX.6.041017.

An, Shuoming et al. (2015). "Experimental test of the quantum Jarzynski equality with a trapped-ion system". In: *Nature Physics* 11.2. URL: https://doi.org/10.1038/nphys3197.

Aubrun, Guillaume and Ion Nechita (2007). "Catalytic Majorization and l_p Norms". In: *Communications in Mathematical Physics* 278.1. URL: http://dx.doi.org/10.1007/s00220-007-0382-4.

— (2009). "Stochastic domination for iterated convolutions and catalytic majorization". In: *Annales de lInstitut Henri Poincaré, Probabilités et Statistiques* 45.3. URL: http://dx.doi.org/10.1214/08-AIHP175.

Audenaert, Koenraad M R (2007). "A sharp continuity estimate for the von Neumann entropy". In: *Journal of Physics A: Mathematical and Theoretical* 40.28. URL: http://dx.doi.org/10.1088/1751-8113/40/28/S18.

Augusiak, R, M Demianowicz, and A Acín (2014). "Local hidden–variable models for entangled quantum states". In: *Journal of Physics A: Mathematical and Theoretical* 47.42. URL: https://doi.org/10.1088%2F1751-8113%2F47%2F42%2F424002.

Bae, Joonwoo, Dariusz Chruscinski, and Marco Piani (2019). "More Entanglement Implies Higher Performance in Channel Discrimination Tasks". In: *Phys. Rev. Lett.* 122 (14). URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.140404.

Bao, Xiao-Hui et al. (2012). "Quantum teleportation between remote atomic-ensemble quantum memories". In: *Proc. Natl. Acad. Sci.* 109.50. URL: https://www.pnas.org/content/109/50/20347.

Barrett, Jonathan (2002). "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality". In: *Phys. Rev. A* 65 (4). URL: https://link.aps.org/doi/10.1103/PhysRevA.65.042302.

Baumgratz, T., M. Cramer, and M.B. Plenio (2014). "Quantifying Coherence". In: *Physical Review Letters* 113.14. URL: http://dx.doi.org/10.1103/PhysRevLett.113.140401.

Bell, John S (1964). "On the einstein podolsky rosen paradox". In: *Physics Physique Fizika* 1.3.

Bennett, Charles H. and Stephen J. Wiesner (1992). "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states". In: *Phys. Rev. Lett.* 69 (20). URL: https://link.aps.org/doi/10.1103/PhysRevLett.69.2881.

Bennett, Charles H. et al. (1993). "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels". In: *Phys. Rev. Lett.* 70 (13). URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.1895.

Bennett, Charles H et al. (1996a). "Concentrating partial entanglement by local operations". In: *Physical Review A* 53.4.

Bennett, Charles H. et al. (1996b). "Mixed-state entanglement and quantum error correction". In: *Phys. Rev. A* 54 (5). URL: https://link.aps.org/doi/10.1103/PhysRevA.54.3824.

Bennett, Charles H. et al. (1999). "Unextendible Product Bases and Bound Entanglement". In: *Phys. Rev. Lett.* 82 (26). URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.5385.

Bennett, Charles H. et al. (2014). "The Quantum Reverse Shannon Theorem and Resource Tradeoffs for Simulating Quantum Channels". In: *IEEE Trans. Inf.* 60.5. URL: http://dx.doi.org/10.1109/TIT.2014.2309968.

Berta, Mario, Matthias Christandl, and Renato Renner (2011). "The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory". In: *Comm. Math. Phys.* 306.3. URL: http://dx.doi.org/10.1007/s00220-011-1309-7.

Boland, Philip J. (1989). "Majority Systems and the Condorcet Jury Theorem". In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 38.3. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.2307/2348873. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.2307/2348873.

Boschi, D. et al. (1998). "Experimental Realization of Teleporting an Unknown Pure Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels". In: *Phys. Rev. Lett.* 80 (6). URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.1121.

Bouwmeester, Dik et al. (1997). "Experimental quantum teleportation". In: *Nature* 390.6660. URL: https://doi.org/10.1038/37539.

Boyd, Stephen and Lieven Vandenberghe (2004). *Convex Optimization*.

Branciard, Cyril et al. (2013). "Measurement-Device-Independent Entanglement Witnesses for All Entangled Quantum States". In: *Phys. Rev. Lett.* 110 (6). URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.060405.

Brandão, Fernando G. S. L. and Gilad Gour (2015). "Reversible Framework for Quantum Resource Theories". In: *Phys. Rev. Lett.* 115 (7). URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.070503.

Brandão, Fernando et al. (2015). "The second laws of quantum thermodynamics". In: *PNAS* 112.11. URL: http://dx.doi.org/10.1073/pnas.1411728112.

Brassard, Gilles et al. (2006). "Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial". In: *Phys. Rev. Lett.* 96 (25). URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.250401.

Briegel, H.-J. et al. (1998). "Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication". In: *Phys. Rev. Lett.* 81 (26). URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.5932.

Buscemi, Francesco (2012). "All Entangled Quantum States Are Nonlocal". In: *Physical Review Letters* 108.20. URL: http://dx.doi.org/10.1103/PhysRevLett.108.200401.

Buscemi, Francesco, Eric Chitambar, and Wenbin Zhou (2020). "Complete Resource Theory of Quantum Incompatibility as Quantum Programmability". In: *Phys. Rev. Lett.* 124 (12). URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.120401.

Camati, Patrice A. et al. (2016). "Experimental Rectification of Entropy Production by Maxwell's Demon in a Quantum System". In: *Phys. Rev. Lett.* 117 (24). URL: https://link.aps.org/doi/10.1103/PhysRevLett.117.240502.

Campisi, Michele, Peter Hänggi, and Peter Talkner (2011). "Colloquium: Quantum fluctuation relations: Foundations and applications". In: *Reviews of Modern Physics* 83.3. URL: http://dx.doi.org/10.1103/RevModPhys.83.771.

Cavalcanti, D. and P. Skrzypczyk (2016). "Quantitative relations between measurement incompatibility, quantum steering, and nonlocality". In: *Phys. Rev. A* 93 (5). URL: https://link.aps.org/doi/10.1103/PhysRevA.93.052112.

Cavalcanti, Daniel, Paul Skrzypczyk, and Ivan Supic (2017). "All Entangled States can Demonstrate Nonclassical Teleportation". In: *Phys. Rev. Lett.* 119 (11). URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.110501.

Cavalcanti, Eric G., Michael J. W. Hall, and Howard M. Wiseman (2013). "Entanglement verification and steering when Alice and Bob cannot be trusted". In: *Phys. Rev. A* 87 (3). URL: https://link.aps.org/doi/10.1103/PhysRevA.87.032306.

Chida, Kensaku et al. (2017). "Power generator driven by Maxwell's demon". In: *Nat. Commun.* 8. Article. URL: https://doi.org/10.1038/ncomms15301.

Childs, Andrew M., John Preskill, and Joseph Renes (2000). "Quantum information and precision measurement". In: *Journal of Modern Optics* 47.2-3. eprint: https://www.tandfonline.com/doi/pdf/10.1080/09500340008244034. URL: https://www.tandfonline.com/doi/abs/10.1080/09500340008244034.

Chitambar, Eric and Gilad Gour (2019). "Quantum resource theories". In: *Rev. Mod. Phys.* 91 (2). URL: https://link.aps.org/doi/10.1103/RevModPhys.91.025001.

Chitambar, Eric et al. (2014). "Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)". In: *Communications in Mathematical Physics* 328.1. URL: http://dx.doi.org/10.1007/s00220-014-1953-9.

Ciganovi, N., N. J. Beaudry, and R. Renner (2014). "Smooth Max-Information as One-Shot Generalization for Mutual Information". In: *IEEE Transactions on Information Theory* 60.3.

Cleve, R. et al. (2004). "Consequences and limits of nonlocal strategies". In: *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*

Cohen, E G D and David Mauzerall (2004). "A note on the Jarzynski equality". In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.07. URL: https://doi.org/10.1088%2F1742-5468%2F2004%2F07%2Fp07006.

Cwiklinski, Piotr et al. (2015). "Limitations on the Evolution of Quantum Coherences: Towards Fully Quantum Second Laws of Thermodynamics". In: *Phys. Rev. Lett.* 115 (21). URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.210403.

Daftuar, Sumit and Matthew Klimesh (2001). "Mathematical structure of entanglement catalysis". In: *Phys. Rev. A* 64 (4). URL: https://link.aps.org/doi/10.1103/PhysRevA.64.042314.

Dam, Wim van (2013). "Implausible consequences of superstrong nonlocality". In: *Natural Computing* 12.1. URL: https://doi.org/10.1007/s11047-012-9353-6.

Dam, Wim van and Patrick Hayden (2003). "Universal entanglement transformations without communication". In: *Phys. Rev. A* 67.6. URL: http://dx.doi.org/10.1103/PhysRevA.67.060302.

Datta, Nilanjana (2009). "Max-relative entropy of entanglement, alias log robustness". In: *International Journal of Quantum Information* 07.02. eprint: https://

doi.org/10.1142/S0219749909005298. URL: https://doi.org/10.1142/S0219749909005298.

Debarba, Tiago et al. (2019). "Work estimation and work fluctuations in the presence of non-ideal measurements". In: *New Journal of Physics* 21.11. URL: http://dx.doi.org/10.1088/1367-2630/ab4d9d.

Deffner, Sebastian and Christopher Jarzynski (2013). "Information Processing and the Second Law of Thermodynamics: An Inclusive, Hamiltonian Approach". In: *Phys. Rev. X* 3 (4). URL: https://link.aps.org/doi/10.1103/PhysRevX.3.041003.

Designolle, Sébastien, Máté Farkas, and Jędrzej Kaniewski (2019a). "Incompatibility robustness of quantum measurements: a unified framework". In: *New Journal of Physics* 21.11. URL: https://doi.org/10.1088%2F1367-2630%2Fab5020.

— (2019b). "Incompatibility robustness of quantum measurements: a unified framework". In: *New Journal of Physics* 21.11. URL: https://doi.org/10.1088%2F1367-2630%2Fab5020.

Designolle, Sébastien et al. (2019). "Quantifying Measurement Incompatibility of Mutually Unbiased Bases". In: *Phys. Rev. Lett.* 122 (5). URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.050402.

Devetak, Igor, Aram W. Harrow, and Andreas J. Winter (2008). "A Resource Framework for Quantum Shannon Theory". In: *IEEE Transactions on Information Theory* 54.10. URL: http://dx.doi.org/10.1109/TIT.2008.928980.

Dinur, Irit, David Steurer, and Thomas Vidick (2013). *A parallel repetition theorem for entangled projection games*. arXiv: 1310.4113 [quant-ph].

Douarche, F et al. (2005). "An experimental test of the Jarzynski equality in a mechanical experiment". In: *Europhysics Letters (EPL)* 70.5. URL: https://doi.org/10.1209%2Fepl%2Fi2005-10024-4.

Du, Shuanping, Zhaofang Bai, and Yu Guo (2015). "Conditions for coherence transformations under incoherent operations". In: *Physical Review A* 91.5. URL: http://dx.doi.org/10.1103/PhysRevA.91.052120.

Duarte, Cristhiano, Raphael C. Drumond, and Marcelo Terra Cunha (2015). *Self-catalytic conversion of pure quantum states*. arXiv: 1504.06364 [quant-ph].

Ducuara, Andrés F., Patryk Lipka-Bartosik, and Paul Skrzypczyk (2020a). "Multiobject operational tasks for convex quantum resource theories of state-measurement pairs". In: *Phys. Rev. Research* 2 (3). URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.033374.

Ducuara, Andrés Felipe and Paul Skrzypczyk (2019). "Weight of informativeness, state exclusion games and excludible information". In: *arXiv preprint arXiv:1908.10347*.

Ducuara, Andrés F., Patryk Lipka-Bartosik, and Paul Skrzypczyk (2020b). "Multiobject operational tasks for convex quantum resource theories of state-measurement pairs". In: *Physical Review Research* 2.3. URL: http://dx.doi.org/10.1103/PhysRevResearch.2.033374.

Egloff, D et al. (2015). "A measure of majorization emerging from single-shot statistical mechanics". In: *New J. Phys.* 17.7. URL: https://doi.org/10.1088%2F1367-2630%2F17%2F7%2F073001.

Einstein, A., B. Podolsky, and N. Rosen (1935). "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10). URL: https://link.aps.org/doi/10.1103/PhysRev.47.777.

Ekert, Artur K. (1991). "Quantum cryptography based on Bell's theorem". In: *Phys. Rev. Lett.* 67 (6). URL: https://link.aps.org/doi/10.1103/PhysRevLett.67.661.

Elitzur, Avshalom C., Sandu Popescu, and Daniel Rohrlich (1992). "Quantum nonlocality for each pair in an ensemble". In: *Physics Letters A* 162.1. URL: http://www.sciencedirect.com/science/article/pii/037596019290952I.

Faist, P. (2016). *Quantum Coarse-Graining: An Information-theoretic Approach to Thermodynamics*.

Fannes, Mark (1973). "A continuity property of the entropy density for spin lattice systems". In: *Communications in Mathematical Physics* 31.4.

Furusawa, A. et al. (1998). "Unconditional Quantum Teleportation". In: *Science* 282.5389. URL: https://science.sciencemag.org/content/282/5389/706.

Goold, John et al. (2016). "The role of quantum information in thermodynamicsa topical review". In: *J. Phys. A* 49.14.

Gottesman, Daniel and Isaac L. Chuang (1999). "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations". In: *Nature* 402.6760. URL: https://doi.org/10.1038/46503.

Gour, Gilad et al. (2015). "The resource theory of informational nonequilibrium in thermodynamics". In: *Physics Reports* 583. URL: http://dx.doi.org/10.1016/j.physrep.2015.04.003.

Hardy, G.H. et al. (1952). *Inequalities*. URL: https://books.google.co.uk/books?id=t1RCSP8YKt8C.

Hardy, Godfrey H (1929). "Some simple inequalities satisfied by convex functions". In: *Messenger Math.* 58.

Hasegawa, Yasushi et al. (2019). "Experimental time-reversed adaptive Bell measurement towards all-photonic quantum repeaters". In: *Nat. Commun.* 10.1. URL: https://doi.org/10.1038/s41467-018-08099-5.

Hoang, Thai M. et al. (2018). "Experimental Test of the Differential Fluctuation Theorem and a Generalized Jarzynski Equality for Arbitrary Initial States". In: *Phys. Rev. Lett.* 120 (8). URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.080602.

Horodecki, Michał, Paweł Horodecki, and Ryszard Horodecki (1998). "Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?" In: *Phys. Rev. Lett.* 80 (24). URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.5239.

— (1999). "General teleportation channel, singlet fraction, and quasidistillation". In: *Phys. Rev. A* 60 (3). URL: https://link.aps.org/doi/10.1103/PhysRevA.60.1888.

Horodecki, Michał, Paweł Horodecki, and Jonathan Oppenheim (2003). "Reversible transformations from pure to mixed states and the unique measure of information". In: *Phys. Rev. A* 67 (6). URL: https://link.aps.org/doi/10.1103/PhysRevA.67.062104.

Horodecki, Michal and Jonathan Oppenheim (2013). "Fundamental limitations for quantum and nanoscale thermodynamics". In: *Nat. Commun.* 4. URL: https://doi.org/10.1038/ncomms3059.

Horodecki, Micha, Jonathan Oppenheim, and Carlo Sparaciari (2018). "Extremal distributions under approximate majorization". In: *J. Phys. A Math. Theor.* 51.30. URL: http://dx.doi.org/10.1088/1751-8121/aac87c.

Horodecki, Pawel et al. (2003). "Rank two bipartite bound entangled states do not exist". In: *Theoretical Computer Science* 292.3. Algorithms in Quantum Information Prcoessing. URL: http://www.sciencedirect.com/science/article/pii/S0304397501003760.

Howard, Mark and Earl Campbell (2017). "Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing". In: *Phys. Rev. Lett.* 118 (9). URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.090501.

Ito, Kosuke et al. (2019). "Generalized energy measurements and quantum work compatible with fluctuation theorems". In: *Phys. Rev. A* 99 (3). URL: https://link.aps.org/doi/10.1103/PhysRevA.99.032117.

Janzing, D. et al. (2000). "Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer's Principle and the Second Law". In: *International Journal of Theoretical Physics* 39.12. URL: https://doi.org/10.1023/A:1026422630734.

Jarzynski, C. (1997a). "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach". In: *Phys. Rev. E* 56 (5). URL: https://link.aps.org/doi/10.1103/PhysRevE.56.5018.

— (1997b). "Nonequilibrium Equality for Free Energy Differences". In: *Phys. Rev. Lett.* 78 (14). URL: https://link.aps.org/doi/10.1103/PhysRevLett.78.2690.

Jonathan, Daniel and Martin B. Plenio (1999). "Entanglement-Assisted Local Manipulation of Pure Quantum States". In: *Phys. Rev. Lett.* 83.17. URL: http://dx.doi.org/10.1103/PhysRevLett.83.3566.

Kimble, H. J. (2008). "The quantum internet". In: *Nature* 453. URL: https://doi.org/10.1038/nature07127.

Kitaev, A Yu (1997). "Quantum computations: algorithms and error correction". In: *Russian Mathematical Surveys* 52.6. URL: https://doi.org/10.1070%2Frm1997v052n06abeh002155.

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2020). "Extreme points and majorization: Economic applications". In: *Available at SSRN*.

Koski, J. V. et al. (2014). "Experimental Observation of the Role of Mutual Information in the Nonequilibrium Dynamics of a Maxwell Demon". In: *Phys. Rev. Lett.* 113 (3). URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.030601.

Koski, J. V. et al. (2015). "On-Chip Maxwell's Demon as an Information-Powered Refrigerator". In: *Phys. Rev. Lett.* 115 (26). URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.260602.

Kurchan, Jorge (2000). "A quantum fluctuation theorem". In: *arXiv preprint cond-mat/0007360*.

Landi, Gabriel T. and Mauro Paternostro (2020). *Irreversible entropy production, from quantum to classical*. arXiv: 2009.07668 [quant-ph].

Leuenberger, Michael N., Michael E. Flatté, and D. D. Awschalom (2005). "Teleportation of Electronic Many-Qubit States Encoded in the Electron Spin of Quantum Dots via Single Photons". In: *Phys. Rev. Lett.* 94 (10). URL: https://link.aps.org/doi/10.1103/PhysRevLett.94.107401.

Leung, Debbie, Ben Toner, and John Watrous (2008). "Coherent state exchange in multi-prover quantum interactive proof systems". In: arXiv: 0804.4118 [quant-ph].

Leung, Debbie and Bingjie Wang (2014). "Characteristics of universal embezzling families". In: *Phys. Rev. A* 90.4.

Lewenstein, Maciej and Anna Sanpera (1998). "Separability and Entanglement of Composite Quantum Systems". In: *Phys. Rev. Lett.* 80 (11). URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.2261.

Li, Ke (2014). "Second-order asymptotics for quantum hypothesis testing". In: *The Annals of Statistics* 42.1. URL: http://dx.doi.org/10.1214/13-AOS1185.

Li, Zheng-Da et al. (2020). "Measurement-Device-Independent Entanglement Witness of Tripartite Entangled States and Its Applications". In: *Phys. Rev. Lett.* 124 (16). URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.160503.

Linden, N. and S. Popescu (1999). "Bound entanglement and teleportation". In: *Phys. Rev. A* 59 (1). URL: https://link.aps.org/doi/10.1103/PhysRevA.59.137.

Lipka-Bartosik, Patryk, Paweł Mazurek, and Michał Horodecki (2021). "Second law of thermodynamics for batteries with vacuum state". In: *Quantum* 5. URL: https://doi.org/10.22331/q-2021-03-10-408.

Lipka-Bartosik, Patryk and Paul Skrzypczyk (2020). "Operational advantages provided by nonclassical teleportation". In: *Phys. Rev. Research* 2 (2). URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.023029.

— (2021a). "All States are Universal Catalysts in Quantum Thermodynamics". In: *Phys. Rev. X* 11 (1). URL: https://link.aps.org/doi/10.1103/PhysRevX.11.011061.

— (2021b). *Catalytic quantum teleportation*. arXiv: 2102.11846 [quant-ph].

Lipka-Bartosik, Patryk et al. (2021). "Operational Significance of the Quantum Resource Theory of Buscemi Nonlocality". In: *PRX Quantum* 2 (2). URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.020301.

Liu, Yunchao and Xiao Yuan (2019). "Operational Resource Theory of Quantum Channels". In: *arXiv preprint arXiv:1904.02680*.

Liu, Zi-Wen, Xueyuan Hu, and Seth Lloyd (2017). "Resource Destroying Maps". In: *Physical Review Letters* 118.6. URL: http://dx.doi.org/10.1103/PhysRevLett.118.060502.

Liu, Zi-Wen and Andreas Winter (2019). *Resource theories of quantum channels and the universal role of resource erasure*. arXiv: 1904.04201 [quant-ph].

Lostaglio, Matteo (2019). "An introductory review of the resource theory approach to thermodynamics". In: *Rep. Prog. Phys.* 82.11. URL: https://doi.org/10.1088%2F1361-6633%2Fab46e5.

Lostaglio, Matteo, David Jennings, and Terry Rudolph (2015). "Description of quantum coherence in thermodynamic processes requires constraints beyond free energy". In: *Nat. Commun.* 6.1.

Marshall, Albert W., Ingram Olkin, and Barry C. Arnold (2011). *Inequalities: Theory of Majorization and its Applications*. Second. Vol. 143.

Marvian, Iman and Robert W Spekkens (2013). "The theory of manipulations of pure state asymmetry: I. Basic tools, equivalence classes and single copy transformations". In: *New Journal of Physics* 15.3. URL: http://dx.doi.org/10.1088/1367-2630/15/3/033001.

Masanes, Lluis and Jonathan Oppenheim (2017). "A general derivation and quantification of the third law of thermodynamics". In: *Nat. Commun.* 8.

Müller, Markus P. (2018). "Correlating Thermal Machines and the Second Law at the Nanoscale". In: *Phys. Rev. X* 8 (4). URL: https://link.aps.org/doi/10.1103/PhysRevX.8.041051.

Müller-Lennert, Martin et al. (2013). "On quantum Rényi entropies: A new generalization and some properties". In: *Journal of Mathematical Physics* 54.12. URL: http://dx.doi.org/10.1063/1.4838856.

Napoli, Carmine et al. (2016). "Robustness of Coherence: An Operational and Observable Measure of Quantum Coherence". In: *Phys. Rev. Lett.* 116 (15). URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.150502.

Ng, N H Y et al. (2015). "Limits to catalysis in quantum thermodynamics". In: *New J. Phys.* 17.8. URL: http://dx.doi.org/10.1088/1367-2630/17/8/085004.

Nielsen, M. A. (1999). "Conditions for a Class of Entanglement Transformations". In: *Phys. Rev. Lett.* 83 (2). URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.436.

Olmschenk, S et al. (2009). "Quantum teleportation between distant matter qubits". In: *Science* 323.5913.

Oszmaniec, Michał and Tanmoy Biswas (2019). "Operational relevance of resource theories of quantum measurements". In: *Quantum* 3. URL: https://doi.org/10.22331/q-2019-04-26-133.

Oszmaniec, Michał et al. (2017). "Simulating Positive-Operator-Valued Measures with Projective Measurements". In: *Phys. Rev. Lett.* 119 (19). URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.190501.

P., S. Peterson J. et al. (2016). "Experimental demonstration of information to energy conversion in a quantum system at the Landauer limit". In: *Proc. Royal Soc. A* 472.2188. URL: https://doi.org/10.1098/rspa.2015.0813.

Palmieri, Benoit and David Ronis (2007). "Jarzynski equality: Connections to thermodynamics and the second law". In: *Phys. Rev. E* 75 (1). URL: https://link.aps.org/doi/10.1103/PhysRevE.75.011133.

Pan, Jian-Wei et al. (1998). "Experimental Entanglement Swapping: Entangling Photons That Never Interacted". In: *Phys. Rev. Lett.* 80 (18). URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.3891.

Piani, Marco and John Watrous (2015). "Necessary and Sufficient Quantum Information Characterization of Einstein-Podolsky-Rosen Steering". In: *Phys. Rev. Lett.* 114 (6). URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.060404.

Pirandola, S. et al. (2015). "Advances in quantum teleportation". In: *Nature Photonics* 9. URL: https://doi.org/10.1038/nphoton.2015.154.

Pironio, Stefano et al. (2009). "Device-independent quantum key distribution secure against collective attacks". In: *New Journal of Physics* 11.4. URL: http://dx.doi.org/10.1088/1367-2630/11/4/045021.

Popescu, Sandu (1994). "Bell's inequalities versus teleportation: What is nonlocality?" In: *Phys. Rev. Lett.* 72 (6). URL: https://link.aps.org/doi/10.1103/PhysRevLett.72.797.

Rajagopal, A. K. and R. W. Rendell (2002). "Robust and fragile entanglement of three qubits: Relation to permutation symmetry". In: *Phys. Rev. A* 65 (3). URL: https://link.aps.org/doi/10.1103/PhysRevA.65.032328.

Regula, Bartosz (2017). "Convex geometry of quantum resource quantification". In: *Journal of Physics A: Mathematical and Theoretical* 51.4. URL: http://dx.doi.org/10.1088/1751-8121/aa9100.

Renner, Renato (2005). *Security of Quantum Key Distribution*. arXiv: quant-ph/0512258 [quant-ph].

Richens, Jonathan G. and Lluis Masanes (2016). "Work extraction from quantum systems with bounded fluctuations in work". In: *Nat. Commun.* 7. URL: https://doi.org/10.1038/ncomms13511.

Rosset, Denis, David Schmid, and Francesco Buscemi (2019). *Characterizing nonclassicality of arbitrary distributed devices*. eprint: arXiv:1911.12462.

Rényi, Alfréd (1961). "On Measures of Entropy and Information". In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. URL: https://projecteuclid.org/euclid.bsmsp/1200512181.

Schmid, David, Denis Rosset, and Francesco Buscemi (2020). "The type-independent resource theory of local operations and shared randomness". In: *Quantum* 4. URL: https://doi.org/10.22331/q-2020-04-30-262.

Schmid, David et al. (2020). *Why standard entanglement theory is inappropriate for the study of Bell scenarios*. arXiv: 2004.09194 [quant-ph].

Schrödinger, E. (1936). "Probability relations between separated systems". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 32.3.

Sherson, Jacob F. et al. (2006). "Quantum teleportation between light and matter". In: *Nature* 443.7111. URL: https://doi.org/10.1038/nature05136.

Shor, Peter W. (1997). "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM Journal on Computing* 26.5. URL: http://dx.doi.org/10.1137/S0097539795293172.

Simon, Christoph and Julia Kempe (2002). "Robustness of multiparty entanglement". In: *Phys. Rev. A* 65 (5). URL: https://link.aps.org/doi/10.1103/PhysRevA.65.052327.

Skrzypczyk, Paul, Nicolas Brunner, and Sandu Popescu (2009). "Emergence of Quantum Correlations from Nonlocality Swapping". In: *Phys. Rev. Lett.* 102 (11). URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.110402.

Skrzypczyk, Paul and Noah Linden (2019). "Robustness of Measurement, Discrimination Games, and Accessible Information". In: *Phys. Rev. Lett.* 122 (14). URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.140403.

Skrzypczyk, Paul, Anthony J. Short, and Sandu Popescu (2014). "Work extraction and thermodynamics for individual quantum systems". In: *Nat. Commun.* 5. URL: https://doi.org/10.1038/ncomms5185.

Steiner, Michael (2003). "Generalized robustness of entanglement". In: *Physical Review A* 67.5. URL: http://dx.doi.org/10.1103/PhysRevA.67.054305.

Streltsov, Alexander et al. (2018). "Maximal coherence and the resource theory of purity". In: *New Journal of Physics* 20.5. URL: https://doi.org/10.1088%2F1367-2630%2Faac484.

Supic, Ivan, Paul Skrzypczyk, and Daniel Cavalcanti (2017). "Measurement-device-independent entanglement and randomness estimation in quantum networks". In: *Phys. Rev. A* 95 (4). URL: https://link.aps.org/doi/10.1103/PhysRevA.95.042340.

— (2019). "Methods to estimate entanglement in teleportation experiments". In: *Phys. Rev. A* 99 (3). URL: https://link.aps.org/doi/10.1103/PhysRevA.99.032334.

Supic, Ivan et al. (2020). "Self-testing and certification using trusted quantum inputs". In: *New Journal of Physics* 22.7. URL: https://doi.org/10.1088%2F1367-2630%2Fab90d1.

Takagi, Ryuji and Quntao Zhuang (2018). "Convex resource theory of non-Gaussianity". In: *Phys. Rev. A* 97 (6). URL: https://link.aps.org/doi/10.1103/PhysRevA.97.062337.

Takagi, Ryuji et al. (2019). "Operational Advantage of Quantum Resources in Subchannel Discrimination". In: *Phys. Rev. Lett.* 122 (14). URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.140402.

Talkner, Peter, Eric Lutz, and Peter Hänggi (2007). "Fluctuation theorems: Work is not an observable". In: *Physical Review E* 75.5. URL: http://dx.doi.org/10.1103/PhysRevE.75.050102.

Tasaki, Hal (2000). "Jarzynski relations for quantum systems and some applications". In: *arXiv preprint cond-mat/0009244*.

Tomamichel, Marco and Masahito Hayashi (2013). "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks". In: *IEEE Transactions on Information Theory* 59.11. URL: http://dx.doi.org/10.1109/TIT.2013.2276628.

Turgut, S (2007). "Catalytic transformations for bipartite pure states". In: *Journal of Physics A: Mathematical and Theoretical* 40.40. URL: https://doi.org/10.1088%2F1751-8113%2F40%2F40%2F012.

Umegaki, Hisaharu (1962). "Conditional expectation in an operator algebra, IV (entropy and information)". In: *Kodai Mathematical Seminar Reports*. Vol. 14. 2. Department of Mathematics, Tokyo Institute of Technology.

Uola, Roope, Tristan Kraft, and Alastair A. Abbott (2020). "Quantification of quantum dynamics with input-output games". In: *Phys. Rev. A* 101 (5). URL: https://link.aps.org/doi/10.1103/PhysRevA.101.052306.

Vaidman, Lev (1994). "Teleportation of quantum states". In: *Phys. Rev. A* 49 (2). URL: https://link.aps.org/doi/10.1103/PhysRevA.49.1473.

Vidal, Guifré and Rolf Tarrach (1999). "Robustness of entanglement". In: *Physical Review A* 59.1.

Vinjanampathy, Sai and Janet Anders (2016). "Quantum thermodynamics". In: *Contemp. Phys.* 57.4. URL: http://dx.doi.org/10.1080/00107514.2016.1201896.

Watrous, John (2018). *The Theory of Quantum Information*.

Werner, Reinhard F. (1989). "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model". In: *Phys. Rev. A* 40 (8). URL: https://link.aps.org/doi/10.1103/PhysRevA.40.4277.

Wilde, Mark M. (2013). *Quantum Information Theory*.

Wilde, Mark M., Andreas Winter, and Dong Yang (2014). "Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy". In: *Communications in Mathematical Physics* 331.2. URL: http://dx.doi.org/10.1007/s00220-014-2122-x.

Winter, Andreas and Dong Yang (2016). "Operational Resource Theory of Coherence". In: *Physical Review Letters* 116.12. URL: http://dx.doi.org/10.1103/PhysRevLett.116.120404.

Zanin, G. L. et al. (2019). "Experimental Quantum Thermodynamics with Linear Optics". In: *Brazilian Journal of Physics* 49.5. URL: https://doi.org/10.1007/s13538-019-00700-6.

Zhang, Xingjian and Qi Zhao (2020). "Simultaneous certification of entangled states and measurements in bounded dimensional semiquantum games". In: *Phys. Rev. Research* 2 (3). URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.033400.

Zukowski, M. et al. (1993). ""Event-ready-detectors" Bell experiment via entanglement swapping". In: *Phys. Rev. Lett.* 71 (26). URL: https://link.aps.org/doi/10.1103/PhysRevLett.71.4287.