CONVEX GEOMETRY *in the* CHARACTERISATION *of* QUANTUM RESOURCES

Bartosz Regula

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy in Mathematics

August 2018

Abstract

Various physical phenomena have found use as *resources* in quantum information processing tasks, and the study of their properties is necessary to provide optimal methods to harness their power. The thesis presents a series of results aiming to characterise the quantitative and operational aspects of general quantum resources, and in particular to establish methods applicable to a variety of resources and emphasise the similarities in their characterisation. Our approach relies on the underlying convex structure of quantum resource theories, employing techniques from convex analysis and optimisation to gain a better understanding of both the fundamental properties of quantum resources as well as our ability to manipulate them efficiently in information processing protocols such as resource distillation.

The first part of the thesis introduces a unified framework for resource quantification, establishing general properties of arbitrary convex quantum resource theories and providing insight into the common structure of many physically relevant resources. The second part of the thesis deals with the convex optimisation problems involved in the operational characterisation of two representative quantum resources, quantum entanglement and quantum coherence, where we in particular establish a detailed description of their distillation under several classes of operations, and introduce methods for the interconversion between the two resources. In the final part, we employ geometric methods to characterise the quantification of entanglement measures based on polynomial invariants and apply the results to investigate the monogamy properties of multipartite entanglement.

Contents

Acknowledgements 5

Publications 6

1 Introduction 7

PRELIMINARIES

2 Basics of quantum information 11
- 2.1 Quantum states and operators 11
- 2.2 Channels and measurements 13

3 Quantum resources 15
- 3.1 Quantum entanglement 15
- 3.2 Quantum coherence 20
- 3.3 Towards characterising general resources 22

4 Convex analysis and optimisation 24
- 4.1 Basic notions 24
- 4.2 Gauge functions 27
- 4.3 Convex optimisation and duality 32
- 4.4 Roof extensions 36

CONVEX GEOMETRY OF QUANTUM RESOURCES

5 Gauges for sets of quantum states 41
- 5.1 Selection of gauges 42
- 5.2 Base norms and robustness 48
- 5.3 Resource monotones 54
- 5.4 Relations with distance-based measures 61

6 Gauges for physical resources 66
- 6.1 k-coherence 66
- 6.2 Bipartite entanglement 70
- 6.3 k-partite entanglement 75
- 6.4 Magic states 77

OPERATIONAL CHARACTERISATION

7 One-shot distillation of quantum coherence 81
7.1 A family of coherence quantifiers 82
7.2 One-shot coherence distillation 87
7.3 Probabilistic distillation of coherence 95
7.4 Assisted distillation of coherence 104
8 One-shot entanglement manipulation beyond LOCC 112
8.1 A family of entanglement quantifiers 112
8.2 One-shot entanglement distillation 116
9 Conversion of nonclassicality into entanglement 128
9.1 Relations between nonclassicality and entanglement 129
9.2 Coherence conversion protocol 133
10 Polynomial measures of entanglement and monogamy 141
10.1 A geometric approach to entanglement quantification 141
10.2 Monogamy of entanglement 149
Bibliography 155
A Proofs and auxiliary results 178

Acknowledgements

First and foremost, I would like to thank my supervisor Gerardo Adesso for his tremendous support, for the opportunities he has given me, and for all the things that he has taught me during the time I spent in his research group. His advice and encouragement were indispensable in all of my achievements so far. His immediate responses to many late-night emails were also very much appreciated.

I am indebted to many collaborators and friends who have made the last three years an amazing experience. I would in particular like to thank Ludovico Lami for teaching me a great deal of things (and at the same time apologise for always bothering him with long discussions in the very late afternoon), Kun Fang and Xin Wang for the fruitful long-distance collaborations, and Ryuji Takagi for many fantastic discussions about a wide variety of topics. I am grateful to Alex Streltsov and Jamie Sikora for inviting me for research visits, during which I enjoyed many stimulating discussions. I also greatly appreciated Gilad Gour's hospitality during my stay at the University of Calgary. I am grateful to the many collaborators I had the pleasure of working with, including Thomas Bromley, Marco Cianciaruso, Sara Di Martino, Andrzej Dragan, Ivette Fuentes, Antony Lee, Soojoon Lee, Rosanna Nichols, Andreas Osterloh, Marco Piani, and Andreas Winter, in addition to the ones already mentioned above. Thanks also to all members, past and present, of the Quantum Correlations group in Nottingham: Ben, Carmine, Gökhan, Giorgio, Ioannis, Ludovico, Luis, Marco, Paul, Pietro, Rosanna, Ryuji, Tom, Tommaso, and Soojoon. There are too many people to thank for making the conferences and workshops I attended such memorable and fun experiences, but all of them are truly appreciated. I am also grateful to Karol Życzkowski and Ryszard Horodecki for their encouragement. And last, but certainly not least, I would like to thank Mario Berta and Tommaso Tufarelli for kindly agreeing to be the examiners for this thesis and providing many useful comments.

Publications

This thesis is based on the following publications:

[BR1] B. Regula and G. Adesso, "Entanglement Quantification Made Easy: Polynomial Measures Invariant under Convex Decomposition", Physical Review Letters **116**, 070504 (2016).

[BR2] B. Regula, A. Osterloh, and G. Adesso, "Strong monogamy inequalities for four qubits", Physical Review A **93**, 052338 (2016).

[BR3] B. Regula and G. Adesso, "Geometric approach to entanglement quantification with polynomial measures", Physical Review A **94**, 022324 (2016).

[BR4] B. Regula, "Convex geometry of quantum resource quantification", Journal of Physics A: Mathematical and Theoretical **51**, 045303 (2018).

[BR5] B. Regula, M. Piani, M. Cianciaruso, T. R. Bromley, A. Streltsov, and G. Adesso, "Converting multilevel nonclassicality into genuine multipartite entanglement", New Journal of Physics **20**, 033012 (2018).

[BR6] B. Regula, K. Fang, X. Wang, and G. Adesso, "One-Shot Coherence Distillation", Physical Review Letters **121**, 010401 (2018).

[BR7] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso, "Probabilistic distillation of quantum coherence", arXiv:1804.09500 (2018).

[BR8] B. Regula, L. Lami, and A. Streltsov, "Non-asymptotic assisted distillation of quantum coherence", arXiv:1807.04705 (2018).

During my PhD, I have additionally contributed to the following publications, which are not covered in this thesis:

[BR9] L. Lami, B. Regula, X. Wang, R. Nichols, A. Winter, and G. Adesso, "Gaussian quantum resource theories", arXiv:1801.05450 (2018).

[BR10] T. R. Bromley, M. Cianciaruso, S. Vourekas, B. Regula, and G. Adesso, "Accessible bounds for general quantum resources", Journal of Physics A: Mathematical and Theoretical **51**, 325303 (2018).

[BR11] N. Johnston, C.-K. Li, S. Plosker, Y.-T. Poon, and B. Regula, "Robustness of k-coherence and k-entanglement", arXiv:1806.00653 (2018).

1 Introduction

The overarching aim of quantum information theory is to understand how information can be transferred, processed, and encoded in physical systems. In recent years, many physical phenomena in quantum information have gone from being of purely theoretical interest to enjoying a variety of operational uses in quantum information protocols, enabling tasks impossible within the realm of classical physics, and prompting a technological revolution currently transforming areas such as communication, computation, cryptography, and sensing. The realisation that intrinsic physical properties of quantum systems can be regarded precisely as *resources* in information processing tasks sparked an investigation of quantum information in a resource-theoretic setting [1–6].

The prototypical example of a quantum resource is quantum entanglement, characterising how nonlocal correlations can significantly enhance the information processing capabilities of two agents restricted to spatially separated laboratories [7]. Following this paradigm, a variety of quantum resource theories have been investigated, including asymmetry [8,9], athermality [10], purity [11], nonlocality [12,13], coherence [8,14–16], nonclassicality [17–20], Einstein-Podolsky-Rosen steering [21], contextuality [22], magic state quantum computation [23], or non-Markovianity [24].

A general resource theory can be understood as the investigation of the properties of a restricted set of quantum states and our ability to manipulate them efficiently with a restricted set of quantum channels. In particular, two crucial aspects of such a characterisation are the *quantification* of the given resource, allowing us to discriminate which quantum states are the most useful in the given physical task, as well as the *conversion* of resourceful states, describing the limitations on transformations between quantum states under the restricted class of operations. Both of the fundamental aspects rely heavily on complicated optimisation problems, making them very difficult to answer in general terms, and frequently requiring resource-specific approaches which cannot be generalised to encompass multiple physically relevant resources.

The purpose of this thesis is to reveal the connections between seemingly unrelated resource theories, develop universal methods applicable to wide ranges of quantum resources, and characterise the quantitative as well as operational properties of a variety of resources. The common theme connecting all of our results is *convexity*, which underlies the structure of quantum states and quantum resources, allowing us to exploit it in an efficient and general description of their properties.

The thesis is organised as follows.

[1] Horodecki & Oppenheim, Int. J. Mod. Phys. B 2012

[2] Sperling & Vogel, Phys. Scr. 2015

[3] Brandão & Gour, PRL 2015

[4] del Rio et al., preprint 2015

[5] Coecke et al., Inf. Comput. 2016

[6] Gour, PRA 2017

[7] Horodecki et al., RMP 2009

[8] Gour & Spekkens, NJP 2008

[9] Lostaglio et al., PRX 2015

[10] Brandão et al., PRL 2013

[11] Horodecki et al., PRA 2003

[12] Gallego et al., PRL 2012

[13] de Vicente, JPA 2014

[14] Aberg, preprint 2006

[15] Baumgratz et al., PRL 2014

[16] Streltsov et al., RMP 2017

[17] Vogel & Sperling, PRA 2014

[18] Killoran et al., PRL 2016

[19] Theurer et al., PRL 2017

[20] Das et al., preprint 2017

[21] Gallego & Aolita, PRX 2015

[22] Grudka et al., PRL 2014

[23] Veitch et al., NJP 2014

[24] Wakakuwa, preprint 2017

In Chapters 2 and 3, we introduce the mathematical basics of quantum information and quantum resources. Chapter 3 in particular presents an operational characterisation of quantum entanglement, which can be seen as the basis for the definition of more general resource theories. We discuss the problems that one might encounter when trying to establish a fully general approach to quantum resources, and motivate the questions that the remainder of the thesis aims to answer. Chapter 4 then introduces the fundamentals of convex analysis and optimisation, which will form the foundation of the methods used in the subsequent chapters.

In Chapter 5, we introduce a unifying formalism for the quantification of quantum resources based on the gauge functions of convex sets, significantly simplifying the construction and characterisation of quantifiers of general quantum resources, and establishing a consolidated view of many resource measures. In particular, we show that many commonly used and well-known resource quantifiers — such as ones based on matrix norms, measures built through the convex roof, the so-called robustness measures, as well as various witness-based quantifiers — are all examples of such gauge functions, allowing us to relate them in a common geometric framework. This allows us to establish an extensive family of quantifiers for any given quantum resource, introduce easily verifiable criteria for a measure to satisfy desirable properties such as faithfulness and strong monotonicity under relevant free operations, and generalise known measures to new quantum resources very easily. Further, we show that many relations and bounds between the measures are in fact universal among quantum resources, and the proofs of such properties can be simplified in our framework. The applications of this formalism to a choice of physically relevant resources such as entanglement (including multipartite entanglement and entanglement of Schmidt number k), quantum coherence, and magic states are discussed in Chapter 6.

Chapter 7 establishes a novel framework for the one-shot distillation of quantum coherence, characterising the operational task in a realistic, non-asymptotic setting. We investigate coherence distillation under several different classes of relevant quantum operations, highlighting differences in their capabilities and establishing their fundamental limitations in state transformations. We then generalise this task to several different operational settings, including probabilistic distillation, where the protocol succeeds only with a certain probability, as well as assisted distillation, which is performed with the assistance of a third party. We provide detailed characterisation and comparison of the operational properties of quantum coherence, and in particular provide a complete description of distillation from pure input states for all classes of free operations.

In Chapter 8, we show that the approach of Chapter 7 generalises also to quantum entanglement, where we employ it to characterise the distillation of entanglement in the one-shot setting. We provide insight into entanglement manipulation under classes of operations which extend the fundamental class of

local operations and classical operations, such as separability-preserving operations and operations with a positive partial transpose. We establish in particular simplified properties of entanglement distillation for pure and maximally correlated states.

Chapter 9 develops an operational equivalence between the two resources of coherence and entanglement in multipartite settings, showing that the resources can always be interconverted with each other. We exemplify the results with an explicit protocol which converts quantum coherence into multipartite entanglement faithfully, and show that in additional to a qualitative equivalence, the resources are related quantitatively, in the sense that the amount of entanglement which can be generated is exactly equal to the coherence in the conversion protocol under a suitable choice of quantifiers.

Finally, Chapter 10 deals with the problem of quantifying a class of polynomial measures of entanglement. We present an exact solution to the convex roof, a notoriously difficult optimisation problem which makes the quantification of mixed-state entanglement difficult in general. By providing geometric insight into the structure of this problem, we solve it for a simplified class of states. The result finds an immediate application in the monogamy relations of multipartite entanglement, where we show that a natural generalisation of known three-qubit monogamy inequalities is violated in four-qubit systems.

Part I

Preliminaries

Basics of quantum information

2

This chapter reviews the basic concepts of quantum information and establishes the notation used throughout the thesis. The definitions follow standard literature [25–27].

[25] Nielsen & Chuang, 2011

[26] Wilde, 2017

[27] Watrous, 2018

2.1 QUANTUM STATES AND OPERATORS

The formalism of quantum information is built upon a complex Hilbert space $\mathcal{H}$ with the inner product $\langle \cdot, \cdot \rangle$, which we will take to be linear in the second argument. Vectors in $\mathcal{H}$ will be represented with the notation $|\psi\rangle$, while elements of the corresponding dual space of linear functionals $\mathcal{H}^\star$ are represented by $\langle \phi |$. By the Riesz representation theorem, we can uniquely identify each $|\phi\rangle \in \mathcal{H}$ with a vector $\langle \phi | \in \mathcal{H}^\star$ such that $\langle \phi | \psi \rangle = \langle |\phi\rangle , |\psi\rangle \rangle$, and we will hereafter follow the convention to use $\langle \phi | \psi \rangle$ to denote the inner product on $\mathcal{H}$. Throughout this thesis, we will concern ourselves with Hilbert spaces which are finite-dimensional, and this will hereafter be assumed without further comment.

The vector space of linear operators acting on $\mathcal{H}$ will be denoted as $\mathcal{L}(\mathcal{H})$, and the notation $A^\dagger$ will mean the adjoint of $A \in \mathcal{L}(\mathcal{H})$. An operator is called self-adjoint if $A = A^\dagger$, positive if it is self-adjoint and $\langle \psi | A | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$, and positive definite if it self-adjoint and $\langle \psi | A | \psi \rangle > 0$ for all $|\psi\rangle \in \mathcal{H}$. Notice that, when equipped with the Hilbert-Schmidt inner product $\langle A, B \rangle = \mathrm{Tr}(A^\dagger B)$, $\mathcal{L}(\mathcal{H})$ forms a Hilbert space of its own. We will often employ the notation $A \geq B$ to denote an operator inequality, equivalent to the fact that the operator $A - B$ is positive. We will use $\ker(A)$ to denote the kernel, and $\mathrm{supp}(A) \coloneqq \ker(A)^\perp$ to denote the support of an operator A. Notice that if A is self-adjoint, then $\mathrm{supp}(A)$ is the range of A. We will also write $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ for the space of linear maps from a Hilbert space $\mathcal{H}_1$ to another Hilbert space $\mathcal{H}_2$.

Since we are only concerned with finite-dimensional Hilbert spaces, it holds that $\mathcal{H} \cong \mathbb{C}^d$, and we will make no notational distinction between a vector $|\psi\rangle \in \mathcal{H}$ and its corresponding column vector $|\psi\rangle \in \mathbb{C}^d$ in a fixed basis, $\langle \phi | \in \mathcal{H}^\star$ and its corresponding row vector $\langle \phi | = |\phi\rangle^\dagger$ with † denoting conjugate transpose, or an operator $A \in \mathcal{L}(\mathcal{H})$ and its corresponding $d \times d$ matrix $A \in \mathbb{C}^{d \times d}$. Self-adjoint operators can then be identified with Hermitian matrices $\mathbb{H}$ (the dimension will be omitted when it is unambiguous), positive operators with positive semidefinite matrices $\mathbb{H}_+$, and positive definite operators with positive definite matrices $\mathbb{H}_{++}$. The norm induced on $\mathcal{H}$ by the inner product then corresponds

to the vector ℓ_2 norm, denoted as $\||\psi\rangle\|_{\ell_2} = \sqrt{\langle\psi|\psi\rangle}$, and the induced norm on $\mathcal{L}(\mathcal{H})$ corresponds to the Frobenius norm $\|A\|_2 = \sqrt{\operatorname{Tr}\langle A, A\rangle} = \sqrt{\operatorname{Tr} A^\dagger A}$.

Quantum states are often identified with a vector $|\psi\rangle \in \mathcal{H}$, normalised so that $\langle\psi|\psi\rangle = 1$. Strictly speaking, the physical state of a quantum system corresponds to a whole equivalence class of such vectors, in the sense that the vectors $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ with $\theta \in \mathbb{R}$ describe the same system and are physically indistinguishable. This motivates us to identify quantum states with the one-dimensional projections $|\psi\rangle\langle\psi|$, which are referred to as *pure states*. We will often use the notation $\psi \coloneqq |\psi\rangle\langle\psi|$ for pure-state projectors.

More generally, one might not have a perfect knowledge about the state of a quantum system — it can be the case that a physical system is described by a statistical ensemble of states $\{|\psi_i\rangle\}$, and the only information we have is that the system can be found in one of the states $|\psi_i\rangle$ with some corresponding probability p_i. Such a quantum system is said to be in a *mixed state*, described by the so-called density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Since each pure state $|\psi_i\rangle$ is normalised and a probability distribution satisfies $p_i \in \mathbb{R}_+$ and $\sum_i p_i = 1$, we then have that the most general representation of a quantum system corresponds to a matrix $\rho \in \mathbb{H}_+$ such that $\operatorname{Tr}(\rho) = 1$. We will use $\mathbb{D}$ or $\mathbb{D}(\mathcal{H})$ to denote the set of all such density matrices.

The lack of knowledge about the state of a quantum system is naturally quantified by the von Neumann entropy $S(\rho)$, defined as

$$S(\rho) \coloneqq -\operatorname{Tr}(\rho \log \rho) \tag{2.1}$$

where for any state ρ written in its spectral decomposition as $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ with $\lambda_i > 0$ we define $\log \rho \coloneqq \sum_i \log \lambda_i |\psi_i\rangle\langle\psi_i|$, and the logarithm is taken in base 2. The von Neumann entropy vanishes for any pure state, and reaches its maximal value of $\log d$ if and only if ρ is the maximally mixed state $\mathbb{1}/d$.

2.1.1 *Composition of states*

Quantum states are often parts of a composite system, where one physical system consists of several subsystems indexed by $\{1, \ldots, n\}$, each described by its corresponding Hilbert space $\mathcal{H}_i$. The state space of such a composite system is then given by the tensor product $\mathcal{H}_{1\cdots n} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, and any state in this composite space corresponds to a density operator $\rho \in \mathcal{L}(\mathcal{H}_{1\cdots n}) = \mathcal{L}(\mathcal{H}_1) \otimes \cdots \otimes \mathcal{L}(\mathcal{H}_n)$. We will often use the shorthand notation $|ii\rangle \coloneqq |i\rangle \otimes |i\rangle$ where $\{|i\rangle\}$ denote some orthonormal bases of the corresponding subsystems.

The operation of *partial trace* Tr_j corresponds to discarding a given subsystem j. It is defined as the linear map satisfying

$$\operatorname{Tr}_j(B_1 \otimes \cdots \otimes B_n) = \operatorname{Tr}(B_j)\, B_1 \otimes \cdots B_{j-1} \otimes B_{j+1} \otimes \cdots \otimes B_n \tag{2.2}$$

where each $B_i \in \mathcal{L}(\mathcal{H}_i)$. More generally, we will denote by $\operatorname{Tr}_{j_1\cdots j_m}$ the partial trace over any collection of subsystems. Notice that $\operatorname{Tr} = \operatorname{Tr}_{1\cdots n}$.

For any state $\rho \in \mathbb{D}(\mathcal{H}_1)$, one can append an auxiliary Hilbert space $\mathcal{H}_2$ such that $\rho = \mathrm{Tr}_2\, \sigma$ for some (non-unique) extension $\sigma \in \mathbb{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. If the operator σ is pure, we call such an extension a *purification* of ρ and we will denote it as $\sigma = |\Psi_\rho\rangle\langle\Psi_\rho|$. Note that every state has a purification in a sufficiently large purifying Hilbert space — for an arbitrary density matrix $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, a purification can be obtained as

$$|\Psi_\rho\rangle = \sum_i \sqrt{\lambda_i}\, |\psi_i\rangle \otimes |i\rangle \tag{2.3}$$

where $\{|i\rangle\}$ are mutually orthogonal states on $\mathcal{H}_2$.

2.1.2 *Distance measures*

The trace norm (Schatten 1-norm) of an operator X is given by $\|X\|_1 = \mathrm{Tr}\sqrt{X^\dagger X}$. A common measure of distance between two states ρ, σ is given by the *trace distance* $\|\rho - \sigma\|_1$. This quantity vanishes if and only if $\rho = \sigma$, and reaches its maximal value 2 if and only if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)^\perp$. Another useful quantity is the *fidelity*, defined as $F(\rho,\sigma) \coloneqq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$ and satisfying $F(\rho,\sigma) = 1$ iff $\rho = \sigma$. Notably, if one of the states is a pure state $\psi = |\psi\rangle\langle\psi|$, then $F(\psi,\sigma) = \langle\psi|\sigma|\psi\rangle$, and in particular for two pure states we have $F(\psi,\phi) = |\langle\psi|\phi\rangle|^2$. An important property of the fidelity is that it can be computed as [28]

[28] Uhlmann, Rep. Math. Phys. 1976

$$F(\rho,\sigma) = \max_{|\Psi_\rho\rangle,|\Psi_\sigma\rangle} \left|\langle\Psi_\rho|\Psi_\sigma\rangle\right|^2 \tag{2.4}$$

where the maximisation is over all purifications of ρ and σ.

2.2 CHANNELS AND MEASUREMENTS

Consider the Hilbert spaces of two parties, $\mathcal{H}_A$ and $\mathcal{H}_{A'}$. Any linear map $\Lambda : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$ is referred to as a superoperator, and we will often write $\Lambda : A \to A'$ for simplicity. We will use id_A to denote the identity map on $\mathcal{L}(\mathcal{H}_A)$. For any two maps $\Lambda : A \to A'$ and $\Phi : B \to B'$, the linear map $\Lambda\otimes\Phi : A\otimes B \to A'\otimes B'$ is defined by $\Lambda \otimes \Phi(X \otimes Y) = \Lambda(X) \otimes \Phi(Y)$. A map Λ is called *Hermiticity-preserving* if $X \in \mathbb{H} \Rightarrow \Lambda(X) \in \mathbb{H}$, *positive* if $X \in \mathbb{H}_+ \Rightarrow \Lambda(X) \in \mathbb{H}_+$, *completely positive* if $\Lambda \otimes \mathrm{id}_{\hat{A}}$ is positive for $\hat{A} \cong A$, and *trace-preserving* if for any X we have $\mathrm{Tr}\,\Lambda(X) = \mathrm{Tr}\, X$. For any Λ, we define the adjoint map $\Lambda^\dagger : A' \to A$ as the unique map which satisfies

$$\langle \Lambda(X), Y\rangle = \left\langle X, \Lambda^\dagger(Y)\right\rangle \tag{2.5}$$

for all $X \in \mathcal{L}(H_A)$ and $Y \in \mathcal{L}(H_{A'})$.

Of particular interest are maps which correspond to physically allowed dynamics of a quantum system, that is, they always map a quantum state into a

valid quantum state. Such maps will be referred to as *quantum channels*, and they correspond exactly to the set of all completely positive and trace-preserving (CPTP) maps.

Superoperators admit several different representations which allow for an easier characterisation of their properties. One such representation is obtained through the *Choi-Jamiołkowski isomorphism*, which to each map $\Lambda : A \to A'$ associates an operator $J_\Lambda \in \mathcal{L}(A \otimes A')$ defined by

$$J_\Lambda \coloneqq \sum_{i,j} |i\rangle\langle j|_A \otimes \Lambda\big(|i\rangle\langle j|_A\big) \tag{2.6}$$

where $\{|i\rangle\}$ is an orthonormal basis for $\mathcal{H}_A$, and for simplicity of notation we do not differentiate between the space $\mathcal{H}_A$ and the space of an ancillary reference system isomorphic to $\mathcal{H}_A$. The action of the map Λ on an arbitrary operator X can be obtained as

$$\Lambda(X) = \mathrm{Tr}_A\left(J_\Lambda\left[X^T \otimes \mathbb{1}_{A'}\right]\right). \tag{2.7}$$

A map Λ is Hermiticity-preserving iff $J_\Lambda \in \mathbb{H}$, completely positive iff $J_\Lambda \in \mathbb{H}_+$, and trace-preserving iff $\mathrm{Tr}_{A'} J_\Lambda = \mathbb{1}_A$.

Another equivalent way to represent superoperators is the so-called *Kraus representation*, which allows us to write the action of any superoperator as

$$\Lambda(X) = \sum_j K_j X L_j^\dagger \tag{2.8}$$

where each $K_j, L_j \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_{A'})$. Note that such a representation is not unique. A map is trace-preserving iff $\sum_j L_j^\dagger K_j = \mathbb{1}_A$, and completely positive iff there exists a Kraus decomposition such that $\Lambda(X) = \sum_j K_j X K_j^\dagger$.

2.2.1 *Measurement*

A generalised quantum measurement is defined in terms of a positive operator-valued measure (POVM), which can be understood as a collection of operators $\{M_i\}$ with each each $M_i \in \mathbb{H}_+$ and such that $\sum_i M_i = \mathbb{1}$. The measurement $\{M_i\}$ applied to a quantum state ρ yields the ith outcome with probability $\langle M_i, \rho\rangle$. If the measurement is projective, i.e. $\{M_i\}$ are mutually orthogonal projection operators, then the state of the quantum system ρ after the measurement is given by

$$\frac{M_i \rho M_i}{\langle M_i, \rho\rangle} \tag{2.9}$$

provided that outcome i was measured. If $\{|i\rangle\}$ is an orthonormal basis of $\mathcal{H}$, the measurement formed by $\{|i\rangle\langle i|\}$ is referred to as the projective measurement in the basis $\{|i\rangle\}$.

3 Quantum resources

This chapter is meant to provide intuition about the mathematical framework for quantum resources, as well as motivate the main questions that we investigate in this thesis. We will begin with an operational characterisation of entanglement as an archetypal example of a resource theory, and later discuss about how some of the insights obtained in this way can be generalised to other resources.

3.1 QUANTUM ENTANGLEMENT

Entanglement was recognised as a fundamental resource in quantum information processing since the early days of the field. Indeed, shared entanglement between two parties enables them to perform protocols which are impossible within the realm of classical information processing, such as quantum teleportation [29] or superdense coding [30]. Entanglement has found use as a resource in a wide variety of tasks underlying quantum communication, computation, cryptography, and many other applications [7], and an efficient characterisation of entanglement manipulation remains at the forefront of quantum information research.

[29] Bennett et al., PRL 1993

[30] Bennett & Wiesner, PRL 1992

[7] Horodecki et al., RMP 2009

Intuitively, the entanglement of a pure state $|\psi_{AB}\rangle$ shared by two spatially separated parties is meant to capture nonlocal quantum correlations between the subsystems possessed by each party. If one treats entanglement as a resource which enables the parties to go beyond the restrictions of locality and classical correlations, this places a natural set of constraints on what kind of operations the separated parties are allowed to perform without using any entanglement — these are the local operations and classical communication (LOCC), establishing a fundamental paradigm in the theory of entanglement. In this setting, LOCC are *free operations*, and the two parties are allowed unlimited use of local quantum operations and classical communication without having to spend any of the resource (in the form of entanglement) to implement these operations.

The precise definition of entanglement is as follows. Given a pure state $|\psi_{AB}\rangle$ in the space $\mathcal{H}_A \otimes \mathcal{H}_B$, we call this state *separable* (or a *product state*) if there exist states $|\psi_A\rangle \in \mathcal{H}_A$, $|\psi_B\rangle \in \mathcal{H}_B$ such that

$$|\psi_{AB}\rangle = |\psi\rangle_A \otimes |\psi\rangle_B \,. \tag{3.1}$$

Conversely, if no such states exist, then the system is *entangled*.

Describing the entanglement of mixed states is much more involved. As a statistical mixture of pure states, a given mixed state can contain both quantum

as well as classical correlations, which have to be differentiated. Additionally, any mixed state ρ can be decomposed into a set of pure states $\{|\psi_i\rangle\}$ in infinitely many ways, with the different decompositions possessing, in general, different values of entanglement on average. To verify that a state ρ is separable, we then need to find at least one way of preparing the state using only local operations on its constituent subsystems, making the system fully characterised by the classical probabilities in this particular decomposition. In other words, we want to describe a mixed state as separable not only if it itself is a tensor product of states as in the pure-state case (3.1), but also if it is a convex combination of such product states.

Consider then a composite system ρ_{AB} in the space $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. If there exist states $\{\rho_A^{(k)}\} \in \mathcal{L}(\mathcal{H}_A)$, $\{\rho_B^{(k)}\} \in \mathcal{L}(\mathcal{H}_B)$ and coefficients $p_k \in \mathbb{R}_+$ with $\sum_k p_k = 1$ such that

$$\rho_{AB} = \sum_k p_k \, \rho_A^{(k)} \otimes \rho_B^{(k)} \,, \tag{3.2}$$

then the state ρ_{AB} is separable. Conversely, if no such states exist, the state is entangled. Separable states are the *free states* in the resource theory of entanglement.

As can be expected, such a definition makes verifying the separability of a state more difficult — in fact, it becomes computationally intractable with growing system dimensions [31]. One can then aim to establish useful separability criteria for a given state that would allow us to verify the lack (or the presence of) entanglement. A particularly simple and useful separability criterion for mixed states is the positive partial transpose (PPT) criterion [32,33], also known as the Peres-Horodecki criterion. The partial transpose operation, denoted T_A or T_B, is defined simply as the transposition applied to one part of a composite system; more precisely, for any operator $X \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ we define its partial transpose as $X^{T_B} \coloneqq (\mathrm{id}_A \otimes T_B)X$ where T_B is the transposition with respect to any orthonormal basis for $\mathcal{H}_B$. The PPT criterion then states that, if a bipartite state ρ_{AB} is separable, then its partial transpose with respect to either A or B is positive. This criterion is a sufficient condition for separability only for low-dimensional ($d_A d_B \leq 6$) systems, while in higher dimensions, there exist states which are entangled despite having a positive partial transpose [34].

[31] Gurvits, 2003

[32] Peres, PRL 1996

[33] Horodecki et al., PLA 1996

[34] Horodecki, PLA 1997

3.1.1 *Pure-state entanglement*

In order to characterise the entanglement of pure states, we will often use the following fundamental property: for every pure state $|\psi_{AB}\rangle$ of two systems A, B, there exist orthonormal bases $\{|a_i\rangle\} \in \mathcal{H}_A$ and $\{|b_i\rangle\} \in \mathcal{H}_B$ such that $|\psi\rangle$ can be written in its *Schmidt decomposition* as

$$|\psi_{AB}\rangle = \sum_{i=1} \lambda_i \, |a_i\rangle \otimes |b_i\rangle \tag{3.3}$$

where λ_i are non-negative *Schmidt coefficients* satisfying $\sum_i \lambda_i^2 = 1$. Notice that a bipartite system is in a pure product state iff it has precisely *one* non-zero Schmidt coefficient.

An important feature of pure-state entanglement is that it makes it impossible to assign a single pure state vector to the subsystems of an entangled state; instead, their description has to include the correlations they share with each other. This can be understood by noting that for a separable state $|\psi_{AB}\rangle$, the partial trace over either of its subsystems gives the pure state corresponding to the other subsystem:

$$\mathrm{Tr}_A\, \psi_{AB} = \psi_B. \tag{3.4}$$

However, in the case of entangled states, the state of the reduced subsystem is mixed, signifying a loss of information about it. We can then employ the von Neumann entropy (2.1) to quantify this mixedness. It follows from the Schmidt decomposition (3.3) that the partial trace of $|\psi_{AB}\rangle$ with respect to either A or B will have the squares of the Schmidt compositions as its non-zero eigenvalues. This means in particular that

$$S\left(\mathrm{Tr}_A\, \psi_{AB}\right) = S\left(\mathrm{Tr}_B\, \psi_{AB}\right) \tag{3.5}$$

for any bipartite state. We can conclude that the von Neumann entropy for any reduced subsystem of a bipartite state vanishes if and only if the state is separable. Indeed, this quantity is often referred to as the *entropy of entanglement*, and provides a particularly useful separability criterion and entanglement measure for all bipartite systems.

Since the von Neumann entropy is maximal only for maximally mixed states, the entropy of entanglement will be maximised by a state with maximally mixed reduced subsystems. The states with this property are the maximally entangled states (Bell states), defined as

$$|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle\,, \tag{3.6}$$

where $|ii\rangle$ can be any orthonormal product basis. Similarly as to how the *bit* was introduced to provide a way to quantify the information content of an information source [35], the *ebit* (or *singlet*), defined as the maximally entangled state of two qubits

[35] Shannon, Bell Syst. Tech. J. 1948

$$|\Psi_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{3.7}$$

can be understood as the basic unit of entanglement as a resource in quantum information processing — in the characterisation of information-theoretic tasks requiring the use of entanglement, the parties are usually assumed to possess entanglement in the form of shared ebits. Notice that a general maximally entangled state $|\Psi_d\rangle$ is equivalent to $\log d$ ebits.

We make note of the fact that, despite the name of maximally entangled states, the notion of what "maximal" entanglement is can also depend on the operational setting. As an example, in entanglement manipulations under local operations with no communication allowed, the so-called embezzling states constitute a resource significantly more powerful than ebits [36].

[36] van Dam & Hayden, PRA 2003

3.1.2 *Mixed-state entanglement and entanglement quantification*

Obviously, the fact that two quantum states are entangled does not mean that both of them are equally useful in any operational task which requires entanglement. Intuitively, if an entangled state ρ' can be obtained from another entangled state ρ by an application of a LOCC transformation, it should have quantitatively *less* entanglement, since the resource content cannot be increased by free operations. This motivates the definition of an *entanglement measure* as any non-negative function $f : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathbb{R}$ which vanishes for all separable states and satisfies $f(\Lambda(\rho)) \leq f(\rho)$ for any LOCC map Λ. If the measure f vanishes only on the set of separable states, it will be called faithful.

Based on the privileged status of the ebit in information processing, an operationally-motivated way to quantify the resource content of an arbitrary bipartite density matrix ρ can be obtained in two ways: the number of ebits one can obtain from ρ without spending any additional resource, or the number of ebits which are necessary to prepare the state ρ without spending any additional resource. These two ways correspond to two fundamental quantum information processing protocols: entanglement distillation [37] that is, the conversion of arbitrary quantum states into singlets with LOCC only, as well as entanglement dilution [38], that is, the conversion of singlets into arbitrary quantum states with LOCC only. Precisely, the *distillable entanglement* of a state ρ is defined as the best rate of distillation achievable with LOCC operations in the limit of infinitely many independent and identically distributed copies of ρ [39], that is,

[37] Bennett et al., PRL 1996

[38] Bennett et al., PRA 1996

[39] Rains, PRA 1999

$$\begin{aligned} E^{\infty}_{d,\mathrm{LOCC}}(\rho) &\coloneqq \sup\left\{ r \;\middle|\; \lim_{n\to\infty} \sup_{\Lambda\in\mathrm{LOCC}} F(\Lambda(\rho^{\otimes n}), \Psi_2^{\otimes rn}) = 1 \right\} \\ &= \sup\left\{ \lim_{n\to\infty} \frac{\log m}{n} \;\middle|\; \lim_{n\to\infty} \sup_{\Lambda\in\mathrm{LOCC}} F(\Lambda(\rho^{\otimes n}), \Psi_m) = 1 \right\}. \end{aligned} \tag{3.8}$$

Similarly, the *entanglement cost* is defined as the best rate of dilution of entanglement:

$$E^{\infty}_{c,\mathrm{LOCC}}(\rho) \coloneqq \inf\left\{ \lim_{n\to\infty} \frac{\log m}{n} \;\middle|\; \lim_{n\to\infty} \sup_{\Lambda\in\mathrm{LOCC}} F(\rho^{\otimes n}, \Lambda(\Psi_m)) = 1 \right\}. \tag{3.9}$$

Importantly, we know that for any pure state it holds that [37,38]

$$E^{\infty}_{d,\mathrm{LOCC}}(\psi) = E^{\infty}_{c,\mathrm{LOCC}}(\psi) = S(\mathrm{Tr}_B\, \psi), \tag{3.10}$$

which in particular implies that LOCC transformations between pure states are asymptotically *reversible*. Using the closed formula for the entanglement cost of a pure state, Bennett et al. [40] proposed to quantify the entanglement cost of a general mixed state ρ with the measure referred to as *entanglement of formation*,

[40] Bennett et al., PRA 1996

$$E_f(\rho) = \min\left\{ \sum_i p_i S(\mathrm{Tr}_B\, \psi_i) \;\middle|\; \rho = \sum_i p_i \psi_i \right\}, \tag{3.11}$$

where the optimisation is over all pure-state convex decompositions of ρ. The reasoning for this choice is as follows: since a mixed state ρ is defined as an ensemble of pure states, we should average over the entanglement of these states to obtain the corresponding entanglement of the ensemble. However, the decomposition of ρ into a set of pure states is not unique, and different decompositions have different entanglement in general. We therefore look for the *minimal* cost of creating entanglement, that is, the most economical decomposition of ρ, which leads to the definition in (3.11). This intuitive approach turned out to be not quite correct, and the entanglement cost was shown to be equal to the *regularised* entanglement of formation [41], that is,

[41] Hayden et al., JPA 2001

$$E^{\infty}_{c,\mathrm{LOCC}}(\rho) = \lim_{n\to\infty} \frac{1}{n} E_f(\rho^{\otimes n}). \tag{3.12}$$

The property that $E_f(\rho^{\otimes n}) \neq nE_f(\rho)$ is a non-trivial fact in itself, and relies on a series of fundamental results in quantum information theory [42–44].

[42] Shor, CMP 2004

[43] Hayden & Winter, CMP 2008

[44] Hastings, Nat. Phys. 2009

Both the entanglement cost and the entanglement of formation are faithful entanglement measures. This is not the case for distillable entanglement, since any state with positive partial transpose is, in fact, undistillable, i.e. $\rho^{T_B} \in \mathbb{H}_+ \Rightarrow E^{\infty}_{d,\mathrm{LOCC}}(\rho) = 0$ [45]. A complete characterisation of distillability is a long-standing open problem [46,47].

[45] Horodecki et al., PRL 1998

[46] Horodecki & Horodecki, QIC 2001

[47] Clarisse, preprint 2006

The approach taken by Bennett et al. in defining the entanglement of formation — the "worst-case" entanglement, i.e. the minimisation of a function over all pure-state decompositions — was later formalised by Uhlmann [48] as the *convex roof extension* of $S(\mathrm{Tr}_B\, \psi)$ (see Sec. 4.4 where we treat this in more detail). The convex roof has many useful properties and has become the standard way to quantify mixed-state entanglement in many operational settings. In particular, general pure-state entanglement monotones can be extended by the convex roof to entanglement monotones on the set of all mixed states [49]. However, the optimisation in Eq. (3.11) is in general very complicated, and a full quantification of the convex roof has only been established in limited special cases. Remarkably, the entanglement of formation has been obtained exactly for any state of two qubits [50], and constitutes a very rare instance of a fully general quantification of mixed-state entanglement.

[48] Uhlmann, OSID 1998

[49] Vidal, J. Mod. Opt. 2000

[50] Wootters, PRL 1998

Besides the operational and convex roof–based measures, a plethora of other entanglement quantifiers can be defined [7,51], based primarily on geometric

[7] Horodecki et al., RMP 2009

[51] Eltschka & Siewert, JPA 2014

considerations — e.g., one can define a faithful quantifier of entanglement simply by considering the distance to the set of separable states. That is, we can define

$$D_{\mathsf{SEP}}(\rho) \coloneqq \inf_{\sigma \in \mathsf{SEP}} D(\rho, \sigma) \tag{3.13}$$

where SEP denotes the set of separable states and D is some distance function such that $D(\Lambda(\rho), \Lambda(\sigma)) \leq D(\rho, \sigma)$ for arbitrary CPTP maps Λ [52,53], e.g. the trace distance. Another major class of entanglement measures can be obtained in a general formalism based on the so-called convex gauge functions, which we discuss in detail in Chapters 5 and 6.

[52] Vedral et al., PRL 1997

[53] Bengtsson & Życzkowski, 2007

3.2 QUANTUM COHERENCE

Quantum coherence is a fundamental property which truly distinguishes quantum systems from their classical counterparts: the ability to be in a superposition of classical states. To motivate the definition of coherent and incoherent states, notice that for any pure state written as a superposition of vectors in a given orthonormal basis $\{|i\rangle\}$ as $|\psi\rangle = \sum_i c_i |i\rangle$, performing a rank-one projective measurement in this basis then collapses the state $|\psi\rangle$ into one of the basis vectors $\{|i\rangle\}$ with the corresponding probability $|c_i|^2$. The uncertainty in the outcome of this measurement is a fundamentally quantum property, arising from the very nature of quantum mechanics, and not our lack of knowledge of the state of the system — indeed, the only way to avoid this uncertainty is to require $|\psi\rangle$ to be classical, in the sense that $|\psi\rangle\langle\psi| = |i\rangle\langle i|$.

As such, quantum coherence is fundamentally a basis-dependent concept: the set of free pure states of this resource theory — so-called incoherent (classical) pure states — is formed only by the projectors $|i\rangle\langle i|$ corresponding to the basis vectors $\{|i\rangle\}$, and the set of all incoherent density operators $\mathcal{I}$ consists of all convex combinations of such free pure states, that is, it corresponds to the set of density matrices diagonal in the given basis.

The resource-theoretic framework of quantum coherence in finite dimensions was established relatively recently [8,14,15,54] (see [16] for a comprehensive overview). A major conceptual difference between the framework of coherence and entanglement is that the definition of coherence as a resource lacks a clear *operational* interpretation. That is not to say that coherence has no applications in quantum information processing — in fact, quite the contrary — but rather that the physical setting does not make it clear what set of operations we should allow as the free operations, and what operational paradigm can mirror the natural setting of "spatially separated labs" of entanglement theory.

The lack of a unique, physically-motivated choice of free operations has not stopped the development of the resource theory of coherence, and in fact it has encouraged the characterisation of coherence under several different possible choices of free operations [16,55–58], necessitating a detailed comparison of the

[8] Gour & Spekkens, NJP 2008

[14] Aberg, preprint 2006

[15] Baumgratz et al., PRL 2014

[54] Levi & Mintert, NJP 2014

[16] Streltsov et al., RMP 2017

[55] Winter & Yang, PRL 2016

[56] Chitambar & Gour, PRL 2016

[57] Marvian & Spekkens, PRA 2016

[58] de Vicente & Streltsov, JPA 2017

properties and capabilities of each type of free operations. We outline the possible choices below.

The largest possible set of free operations are the *maximally incoherent operations (MIO)* [14], which are given by quantum channels $\mathcal{E}$ such that $\mathcal{E}(\rho) \in \mathcal{I}$ for every $\rho \in \mathcal{I}$. They are clearly the largest set of operations which cannot create coherence. A smaller class are the *dephasing-covariant incoherent operations (DIO)* [56,57], which are maps $\mathcal{E}$ which commute with the dephasing operation $\Delta(\cdot) \coloneqq \sum_i |i\rangle\langle i| \cdot |i\rangle\langle i|$, that is, $[\Delta, \mathcal{E}] = 0$. The *incoherent operations (IO)* [15] are those for which there exists a Kraus decomposition into incoherent Kraus operators, that is, $\{K_\ell\}$ such that $K_\ell \rho K_\ell^\dagger / \operatorname{Tr}(K_\ell \rho K_\ell^\dagger) \in \mathcal{I}$ for all ℓ and all $\rho \in \mathcal{I}$. These transformations can be interpreted as incoherent measurements which cannot create coherence even if postselection is applied to the individual measurement outcomes. The *strictly incoherent operations (SIO)* [55,59] are operations for which both $\{K_\ell\}$ and $\{K_\ell^\dagger\}$ are sets of incoherent operators. The following *strict* inclusions hold [57,60]:

[14] Aberg, preprint 2006

[56] Chitambar & Gour, PRL 2016

[57] Marvian & Spekkens, PRA 2016

[15] Baumgratz et al., PRL 2014

[55] Winter & Yang, PRL 2016

[59] Yadin et al., PRX 2016

[60] Chitambar & Gour, PRA 2016

$$\begin{aligned} \text{MIO} &\supsetneq \text{IO} \supsetneq \text{SIO} \\ \text{MIO} &\supsetneq \text{DIO} \supsetneq \text{SIO} \end{aligned} \tag{3.14}$$

and there is no inclusion between IO and DIO. While MIO is the largest possible choice of free operations, SIO can be regarded as the smallest known choice for which the resource theory does not become trivial in some regards, and which satisfies all properties desired from a valid choice of free operations [16].

[16] Streltsov et al., RMP 2017

The operational tasks such as coherence distillation and coherence dilution can then be defined analogously as in the resource theory of entanglement, where now the unit of the resource is represented by the maximally coherent state

$$|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle\,. \tag{3.15}$$

In particular, for any choice of operations $\mathcal{O} \in \{\text{MIO}, \text{DIO}, \text{IO}, \text{SIO}\}$, we define the distillable coherence and the coherence cost as

$$\begin{aligned} C_{d,\mathcal{O}}^{\infty}(\rho) &\coloneqq \sup\left\{ \lim_{n\to\infty} \frac{\log m}{n} \;\middle|\; \lim_{n\to\infty} \sup_{\Lambda\in\mathcal{O}} F(\Lambda(\rho^{\otimes n}), \Psi_m) = 1 \right\} \\ C_{c,\mathcal{O}}^{\infty}(\rho) &\coloneqq \inf\left\{ \lim_{n\to\infty} \frac{\log m}{n} \;\middle|\; \lim_{n\to\infty} \sup_{\Lambda\in\mathcal{O}} F(\rho^{\otimes n}, \Lambda(\Psi_m)) = 1 \right\}. \end{aligned} \tag{3.16}$$

Interestingly, it has been found that both of these quantities can be computed exactly for the set of incoherent operations [55], leading to

$$\begin{aligned} C_{d,\text{IO}}^{\infty}(\rho) &= S(\Delta(\rho)) - S(\rho) \\ C_{c,\text{IO}}^{\infty}(\rho) &= \min\left\{ \sum_i p_i S(\Delta(\psi_i)) \;\middle|\; \rho = \sum_i p_i \psi_i \right\}, \end{aligned} \tag{3.17}$$

where $C^{\infty}_{c,\mathrm{IO}}$ can be identified as the so-called coherence of formation. This shows in particular the much simpler structure of the resource theory of coherence compared to entanglement theory. We will consider the optimal rates under other choices of operations in Chapter 7.

The quantitative aspects of quantum coherence have been a large part of the focus of recent research into this resource theory [14–16,61,62]. One can establish some obvious parallels with the resource theory of entanglement, and many measures of coherence — both operational ones as well as ones based on geometric and algebraic considerations — can be defined by mirroring the corresponding entanglement measures. However, establishing that these measures have the desired properties, characterising the interrelations between them, and quantifying them for general states is still by no means trivial, as the results from entanglement theory cannot be applied directly in most cases. It then becomes a natural question to ask whether there is a better and more general way to approach this problem — that is, whether a universal quantitative approach to quantum resources, which could allow us to exploit similarities between general resource theories and eliminate the need to consider every single resource separately, can be established.

[14] Aberg, preprint 2006
[15] Baumgratz et al., PRL 2014
[16] Streltsov et al., RMP 2017
[61] Streltsov et al., PRL 2015
[62] Napoli et al., PRL 2016

3.3 TOWARDS CHARACTERISING GENERAL RESOURCES

We are now in position to define what we mean by a *resource theory* in the most general sense. We define a resource theory to consist of a set of free states $\mathcal{S}$, which do not possess the given resource, as well as a set of free operations $\mathcal{O}$, which cannot generate the given resource. The goal of the characterisation of general resources is then to understand two main tasks: resource quantification as well as resource conversion. The former deals with establishing methods for measuring the resource content of a given state, and the latter with describing the possible state transformations achievable under the class of free operations.

Notice that the operational description of a resource is heavily dependent on physical considerations, and the choice of the free operations reflects the physical restrictions placed on resource manipulation by the setting of the resource theory. In many cases, such as the aforementioned quantum coherence, a unique physically-motivated choices of free operations might not exist. In other cases, it might not be possible to define operational tasks such as "resource distillation", since arbitrary resource theories usually do not possess a standard unit of a resource such as a maximally entangled or maximally coherent states. This makes it difficult, if not impossible, to make precise statements about the operational capabilities of free operations of general quantum resource theories, and frequently requires resource-specific approaches.

However, it is often possible to characterise the set of free states very well. The leading question for our investigation is then: how much can we say about a resource theory knowing only the free states, but not the operations in con-

sideration? In particular, which results concerning the quantitative properties can be shown to hold regardless of the choice of free operations, or for a wide range of possible choices of operations? In Chapters 5 and 6, we aim to answer these question by presenting a mathematical framework for the quantification of resources, which shows that general statements about the properties of quantum resource theories can be made by only relying on an underlying convex structure of the relevant set of free states, establishing many relations which are universal among most resource theories, and straightforwardly generalising many results obtained in the resource theories of entanglement and coherence.

Coming back to the resource theories of coherence and entanglement, we investigate in Chapters 7 and 8 how changing the class of free operations in consideration affects the performance of distillation of the two resources, also generalising coherence and entanglement distillation to non-asymptotic settings. By construction, these approaches are resource-specific, but we focus on developing methods which encompass both of the resources and show interrelations between them. Another aspect of resource conversion — transformations which enable us to convert coherence into entanglement and vice versa — are explored in Chapter 9.

Convex analysis and optimisation

4

This chapter establishes some basic results in convex analysis, as well as introduces the formalism of gauge functions, which will be employed extensively throughout the thesis.

The definitions in this chapter follow standard literature, in particular Rockafellar [63] as well as Hiriart-Urruty and Lemaréchal [64,65].

[63] Rockafellar, 1970

[64] Hiriart-Urruty & Lemaréchal, 1993

[65] Hiriart-Urruty & Lemaréchal, 1993

4.1 BASIC NOTIONS

We will use $\mathbb{R}_+$ to denote non-negative reals and $\mathbb{R}_{++}$ to denote positive reals. Throughout this section, we will consider a general finite-dimensional real vector space $\mathcal{V}$ with an inner product $\langle \cdot, \cdot \rangle$, where for simplicity we will make no distinction between the space and its dual. We will then use $\| \cdot \|$ for the norm induced by the inner product.

Given a set $\mathcal{S}$, a linear combination of its elements is defined as $\sum_i t_i x_i$ for some $x_i \in \mathcal{S}$ and any choice of coefficients $t_i \in \mathbb{R}$; an affine combination is a linear combination with the additional constraint that $\sum_i t_i = 1$; and a convex combination of elements of $\mathcal{S}$ is an affine combination which satisfies $t_i \in \mathbb{R}_+ \ \forall i$. A set $\mathcal{S}$ is then called linear (respectively, affine / convex) if it contains any linear (respectively, affine / convex) combination of its elements. Notice in particular that any linear set is a subspace of $\mathcal{V}$, and any affine set can be thought of as a translated subspace (it need not contain the origin). Yet another equivalent definition is that a set is linear iff it is the kernel of a linear transformation $x \mapsto \Lambda(x)$, and similarly a set is affine iff it is the kernel of an affine transformation $x \mapsto \Lambda(x) + b$ for some $b \in \mathcal{V}$. Notice that the intersection of two linear / affine / convex sets is again linear / affine / convex.

The linear span (linear hull) of a set $\mathcal{S}$, denoted $\operatorname{span}(\mathcal{S})$, is the smallest linear set which contains $\mathcal{S}$; analogously, the affine hull $\operatorname{aff}(\mathcal{S})$ and the convex hull $\operatorname{conv}(\mathcal{S})$ are, respectively, the smallest affine set and the smallest convex set which contains $\mathcal{S}$. Alternatively, $\operatorname{span}(\mathcal{S})$ / $\operatorname{aff}(\mathcal{S})$ / $\operatorname{conv}(\mathcal{S})$ can be understood as the set of all linear / affine / convex combinations of the elements of $\mathcal{S}$; explicitly,

$$\operatorname{conv}(\mathcal{S}) = \left\{ \sum_i t_i x_i \;\middle|\; t_i \in \mathbb{R}_+,\ \sum_i t_i = 1,\ x_i \in \mathcal{S} \right\} \tag{4.1}$$

and similarly for the other cases.

A set $\mathcal{K}$ is called a *cone* if it contains any non-negative scalar multiple of its elements, i.e. $x \in \mathcal{K} \Rightarrow \lambda x \in \mathcal{K} \ \forall \lambda \in \mathbb{R}_+$. A convex cone $\mathcal{K}$ is then a cone which

contains any conic combination of elements in $\mathcal{K}$, that is, any linear combination which satisfies $t_i \in \mathbb{R}_+ \ \forall i$. An important object is the *dual cone* $\mathcal{S}^\star$ of an arbitrary set $\mathcal{S}$, defined as

$$\mathcal{S}^\star = \left\{ x \;\middle|\; \langle x, s \rangle \geq 0 \ \forall s \in \mathcal{S} \right\}. \tag{4.2}$$

This cone is always convex and closed, regardless of whether $\mathcal{S}$ is. Furthermore, $\mathcal{S}^{\star\star} \coloneqq (\mathcal{S}^\star)^\star$ is the so-called closed conic hull of $\mathcal{S}$, in the sense that it is the smallest closed convex cone which contains $\mathcal{S}$.

A hyperplane is an affine set defined as a translation of a linear subspace of dimension $\dim \mathcal{V} - 1$. Each hyperplane H can be written as

$$H = \left\{ x \;\middle|\; \langle x, y \rangle = a \right\} \tag{4.3}$$

for some $y \in \mathcal{V}$, $a \in \mathbb{R}$. Given two convex sets $\mathcal{S}_1, \mathcal{S}_2$, the hyperplane H is said to separate $\mathcal{S}_1$ and $\mathcal{S}_2$ if for all $x \in \mathcal{S}_1$ it holds that $\langle x, y \rangle \geq a$, while for all $x \in \mathcal{S}_2$ it holds that $\langle x, y \rangle \leq a$. *H strongly* separates the two sets if the inequalities are strict. An important property, often known as the strongly separating hyperplane theorem or the Hahn-Banach theorem, is that two disjoint closed and convex sets can always be strongly separated with a hyperplane provided that at least one of them is bounded.

A point $x \in \mathcal{S}$ is an *extreme point* of the convex set $\mathcal{S}$ if there is no way to express x as a convex combination of points $\{y_i\} \in \mathcal{S}$ except for taking $y_i = x \ \forall i$. We will use $\operatorname{ext}(\mathcal{S})$ to denote the set of all extreme points of $\mathcal{S}$. Importantly, for any convex set it holds that $\mathcal{S} = \operatorname{conv}(\operatorname{ext}(\mathcal{S}))$.

Define the ε-ball around a point x as $B_\varepsilon(x) \coloneqq \left\{ \widetilde{x} \in \mathcal{V} \;\middle|\; \|x - \widetilde{x}\| \leq \varepsilon \right\}$. The *relative interior* of a set $\mathcal{S}$ is then defined as the interior of $\mathcal{S}$ regarded as a subset of $\operatorname{aff}(\mathcal{S})$; precisely,

$$\operatorname{relint}(\mathcal{S}) = \left\{ x \in \operatorname{aff}(\mathcal{S}) \mid \exists\, \varepsilon \in \mathbb{R}_{++} \text{ s.t. } B_\varepsilon(x) \cap \operatorname{aff}(\mathcal{S}) \subseteq \mathcal{S} \right\}. \tag{4.4}$$

We will also use the notation $\operatorname{relint}_{\mathcal{X}}(\mathcal{S})$ to mean the relative interior of $\mathcal{S}$ regarded as a subset of $\mathcal{X}$.

4.1.1 *Convex functions*

A function $f : \mathcal{V} \to \mathbb{R}$ is called convex if its epigraph

$$\operatorname{epi}(f) \coloneqq \left\{ (x, t) \;\middle|\; f(x) \leq t \right\} \in \mathcal{V} \oplus \mathbb{R} \tag{4.5}$$

is a convex set. Alternatively, f is convex if and only if

$$f\left(\sum_i t_i x_i \right) \leq \sum_i t_i f(x_i) \tag{4.6}$$

for any convex combination of elements $x_i \in \mathcal{V}$. A function is called concave if $-f$ is convex.

We will extend the above definition to functions which can take values in the extended real numbers $\mathbb{R} \cup \{\pm\infty\}$. The algebraic operations involving infinity are mostly obvious, such as $c + \infty = \infty$ or $c\infty = \operatorname{sgn}(c)\infty$ for any $c \in \mathbb{R}$, where $0\infty \coloneqq 0$ and sgn is the sign function. The operation $\infty - \infty$, which could lead to contradictions, is left undefined and will be avoided. We will hereafter work with functions which are *proper*, that is, $f(x) > -\infty$ for all x and $f(x) < \infty$ for at least one x, as improper functions can lead to undesired behaviour — in particular, an improper convex function cannot take any finite values, because it would violate convexity. The *effective domain* of a proper function f is then defined as

$$\operatorname{dom}(f) \coloneqq \left\{ x \in \mathcal{V} \;\middle|\; f(x) < \infty \right\}. \tag{4.7}$$

To see the applicability of this convention, define the function $\widetilde{f} : \operatorname{dom}(f) \to \mathbb{R}$ as $\widetilde{f}(x) = f(x)\ \forall x \in \operatorname{dom}(f)$, and notice that if $\widetilde{f}$ is convex on $\operatorname{dom}(f)$, then $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ is convex on the whole space $\mathcal{V}$. This allows us to derive results about convex functions without having to explicitly consider their domains.

Define the *indicator function* associated with a set $\mathcal{S}$ as

$$\delta_{\mathcal{S}}(x) \coloneqq \begin{cases} 0 & x \in \mathcal{S} \\ \infty & x \notin \mathcal{S}. \end{cases} \tag{4.8}$$

The indicator function is convex as long as $\mathcal{S}$ is (notice, however, that this would not be true if $\delta_{\mathcal{S}}$ was allowed to take a finite value outside of $\mathcal{S}$). Another basic convention we will take is $\inf \varnothing = \infty$ and analogously $\sup \varnothing = -\infty$, which in particular allows us to use the indicator function to express the minimisation of any function as

$$\inf_{x \in \mathcal{S}} f(x) = \inf \left(f(x) + \delta_{\mathcal{S}}(x) \right). \tag{4.9}$$

A function f is lower semicontinuous (l.s.c.) if

$$\liminf_{x \to x_0} f(x) \geq f(x_0)\ \forall\, x. \tag{4.10}$$

Convex and l.s.c. functions are often called closed, since for proper convex functions lower semicontinuity is equivalent to their epigraph being closed.

The convex conjugate (also known as the Fenchel conjugate or the Legendre-Fenchel transform) of a function $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ is defined as

$$f^*(y) \coloneqq \sup_x \langle y, x \rangle - f(x). \tag{4.11}$$

This function is always l.s.c. and convex — the convex biconjugate f^{**} is in fact the largest l.s.c. and convex function smaller than f, often referred to as the convex closure of f. An example of a convex conjugate is

$$\delta_{\mathcal{S}}^*(y) = \sup_{x \in \mathcal{S}} \langle y, x \rangle, \tag{4.12}$$

called the support function of $\mathcal{S}$.

Finally, we will employ the operation of *infimal convolution*, defined for two proper functions f, g as

$$(f \mathbin{\dot{\vee}} g)(x) \coloneqq \inf_{x=y+z} f(y) + g(z). \tag{4.13}$$

Many important functions can be written as an infimal convolution — notice, for example, that

$$(f \mathbin{\dot{\vee}} \delta_{\mathcal{S}})(x) = \inf_{y \in \mathcal{S}} f(x-y) \tag{4.14}$$

which is an expression that one commonly encounters e.g. when minimising the distance to a particular set $\mathcal{S}$. This operation satisfies several useful properties. As long as f and g are both convex and l.s.c., so is $f \mathbin{\dot{\vee}} g$. Furthermore, the following fact will be useful.

Lemma 4.1. *If f and g are proper, convex, l.s.c., and such that* $\operatorname{dom} f \cap \operatorname{dom} g \neq \varnothing$, *then*

$$(f \mathbin{\dot{\vee}} g)^\star = f^\star + g^\star. \tag{4.15}$$

If, furthermore, $\operatorname{relint}(\operatorname{dom} f) \cap \operatorname{relint}(\operatorname{dom} g) \neq \varnothing$, *then*

$$(f + g)^\star = f^\star \mathbin{\dot{\vee}} g^\star. \tag{4.16}$$

The proof is rather technical and we refer e.g. to [63, Thm. 16.4] for details. The infimal convolution has an intuitive geometric interpretation: as long as the functions satisfy the assumptions of Lemma 4.1, then

[63] Rockafellar, 1970

$$\operatorname{epi}(f \mathbin{\dot{\vee}} g) = \operatorname{epi} f + \operatorname{epi} g \tag{4.17}$$

where we have defined the Minkowski sum of arbitrary sets $\mathcal{A}, \mathcal{B}$ as

$$\mathcal{A} + \mathcal{B} \coloneqq \left\{ a + b \;\middle|\; a \in \mathcal{A},\ b \in \mathcal{B} \right\}. \tag{4.18}$$

All of the above considerations extend to the infimal convolution of any finite number of proper functions,

$$(f_1 \mathbin{\dot{\vee}} \cdots \mathbin{\dot{\vee}} f_n)(x) = \inf_{x=x_1+\ldots+x_n} f_1(x_1) + \ldots + f_n(x_n). \tag{4.19}$$

4.2 GAUGE FUNCTIONS

The definitions in this section follow [63, §14-15] in particular; see also [64,66,67], where much of this information is reviewed and expanded on.

[64] Hiriart-Urruty & Lemaréchal, 1993

[66] Bach, preprint 2013

[67] Friedlander et al., SIAM J. Optim. 2014

Given a set $\mathcal{S}$, we define its *polar set* $\mathcal{S}^\circ$ as

$$\mathcal{S}^\circ = \left\{ x \;\middle|\; \langle x, s \rangle \leq 1 \ \forall s \in \mathcal{S} \right\} \tag{4.20}$$

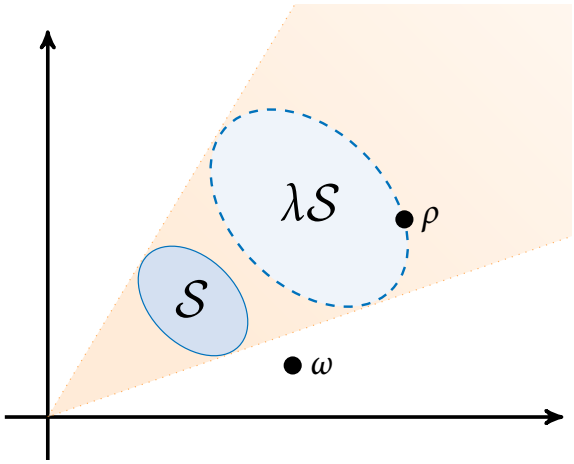


Figure 4.1: Gauge function of a set. The gauge function $\gamma_{\mathcal{S}}$ has a finite value at the point ρ, while ω lies outside of the domain of this gauge.

The polar set is always convex and closed, and always contains the origin. The bipolar set, given by

$$\mathcal{S}^{\circ\circ} \coloneqq (\mathcal{S}^{\circ})^{\circ} = \overline{\operatorname{conv}(\mathcal{S} \cup \{0\})} \tag{4.21}$$

where $\overline{\cdot}$ denotes closure, is the smallest closed convex set containing $\mathcal{S}$ and the origin. Note that the polar set of a cone $\mathcal{K}$ is given by $\mathcal{K}^{\circ} = -\mathcal{K}^{\star}$, where we recall that $\mathcal{K}^{\star}$ denotes the dual cone.

The *gauge function* $\gamma_{\mathcal{S}}$ for a set $\mathcal{S}$ (also known as the Minkowski functional) is defined as

$$\gamma_{\mathcal{S}}(x) = \inf\left\{\lambda \in \mathbb{R}_+ \;\middle|\; x \in \lambda\mathcal{S}\right\} = \inf_{s\in\mathcal{S}}\left\{\lambda \in \mathbb{R}_+ \;\middle|\; x = \lambda s\right\} \tag{4.22}$$

where we note that the effective domain of $\gamma_{\mathcal{S}}$ is the cone generated by $\mathcal{S}$, that is $\left\{\mu s \;\middle|\; \mu \in \mathbb{R}_+,\, s \in \mathcal{S}\right\} =: \mathbb{R}_+\mathcal{S}$. Notice that, if $\mathcal{S}$ is a cone, then $\gamma_{\mathcal{S}} = \delta_{\mathcal{S}}$.

One of the most common examples of gauge functions are norms, which are the gauge functions of their unit balls. Norms are defined as finite functions satisfying the following axioms: absolute homogeneity (of degree 1), subadditivity (triangle inequality), and positivity everywhere except the origin. This corresponds to gauge functions of sets $\mathcal{S}$ which are convex, compact, centrally symmetric ($\mathcal{S} = -\mathcal{S}$), and such that $0 \in \operatorname{int}(\mathcal{S})$. In general, we have the following relations between sets and their corresponding gauges [63,68]:

[63] Rockafellar, 1970

[68] Freund, Math. Program. 1987

$\mathcal{S}$ is convex and	$\gamma_{\mathcal{S}}$ is
—	convex positively homogeneous, i.e. $\gamma_{\mathcal{S}}(\alpha x) = \alpha\gamma_{\mathcal{S}}(x) \forall \alpha \in \mathbb{R}_{++}$
closed	lower semicontinuous
compact	positive everywhere except 0, i.e. $\gamma_{\mathcal{S}}(x) > 0 \ \forall x \neq 0$
centrally symmetric	symmetric (even), i.e. $\gamma_{\mathcal{S}}(x) = \gamma_{\mathcal{S}}(-x)$
$0 \in \operatorname{int}(\mathcal{S})$	finite everywhere, i.e. $\gamma_{\mathcal{S}}(x) < \infty \ \forall x$

where we note that evenness together with positive homogeneity implies absolute homogeneity, and convexity with positive homogeneity implies subadditivity.

An important property of any gauge corresponding to a closed convex set $\mathcal{S}$ is that if $0 \in \mathcal{S}$, then the unit ball of $\gamma_{\mathcal{S}}$ is given exactly by $\mathcal{S}$, and each such unit ball *uniquely* determines its corresponding gauge — more generally, all closed convex gauge functions are in a one-to-one correspondence with closed convex sets containing the origin. In full generality, the unit ball of a convex gauge is given by

$$\left\{ x \;\middle|\; \gamma_{\mathcal{S}}(x) \leq 1 \right\} = \mathcal{S}^{\circ\circ} \tag{4.23}$$

which in particular means that $\gamma_{\mathcal{S}} = \gamma_{\mathcal{S}^{\circ\circ}}$.

4.2.1 *Polarity*

The relation between a gauge function of a convex set $\mathcal{S}$ and its polar set $\mathcal{S}^\circ$ is established through convex conjugation. In particular, notice that

$$\begin{aligned} \gamma_{\mathcal{S}}^{\star}(y) &= \sup \langle x, y \rangle - \gamma_{\mathcal{S}}(x) \\ &= \delta_{\mathcal{S}^\circ}(y), \end{aligned} \tag{4.24}$$

which can be seen by noting that if $y \notin \mathcal{S}^\circ$ then we can always choose $x \in \mathcal{S}$ such that $\langle x, y \rangle > \gamma_{\mathcal{S}}(x)$ and then use the homegeneity of $\gamma_{\mathcal{S}}$, and on the other hand if $y \in \mathcal{S}^\circ$ then $\langle x, y \rangle \leq \gamma_{\mathcal{S}}(x) \ \forall x$ holds by definition of the polar set.

It is well-known that any norm $\|\cdot\|$ and its dual $\|\cdot\|^\circ$ satisfy the Cauchy-Schwarz inequality, $\langle x, y \rangle \leq \|x\|\|y\|^\circ$. In an analogous fashion, we define the *polar function* of a convex gauge $\gamma_{\mathcal{S}}$ as the function which satisfies the Cauchy-Schwarz inequality most closely:

$$\gamma_{\mathcal{S}}^{\circ}(y) \coloneqq \inf \left\{ \mu \in \mathbb{R}_+ \;\middle|\; \langle x, y \rangle \leq \mu \gamma_{\mathcal{S}}(x) \ \forall x \right\}. \tag{4.25}$$

Using the positive homogeneity of $\gamma_{\mathcal{S}}$, this can be written as

$$\begin{aligned} \gamma_{\mathcal{S}}^{\circ}(y) &= \sup \left\{ \langle x, y \rangle \;\middle|\; \gamma_{\mathcal{S}}(x) \leq 1 \right\} \\ &= \sup_{x \in \mathcal{S}^{\circ\circ}} \langle x, y \rangle \\ &= \left[\sup_{x \in \mathcal{S}} \langle x, y \rangle \right]_+ \end{aligned} \tag{4.26}$$

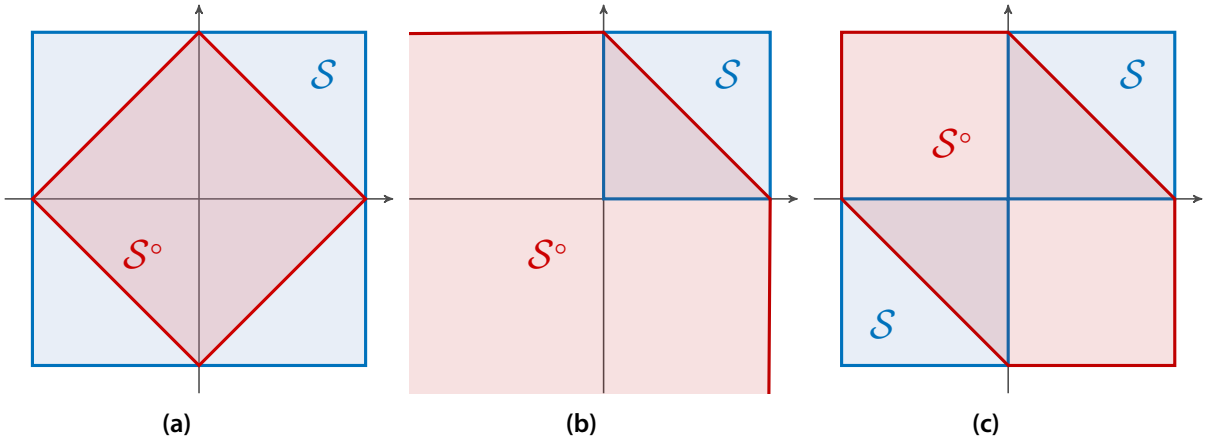


Figure 4.2: Examples of polars of different sets in $\mathbb{R}^2$. The set $\mathcal{S}$ in (a) can be identified with the unit ball of the ℓ_∞ norm, whose polar set $\mathcal{S}^\circ$ is the unit ball of the ℓ_1 norm. Restricting the set $\mathcal{S}$ to the non-negative orthant in (b) leads to an unbounded polar $\mathcal{S}^\circ$. The boundedness is then restored in (c) by considering a set consisting of the restricted unit ball mirrored through the origin. Notice that all cases leave the part of the polar in the non-negative orthant unaffected.

where $[\lambda]_+ \coloneqq \max\{0, \lambda\}$. Notice in particular that $\gamma^\circ_{\mathcal{S}} = \delta^\star_{\mathcal{S}^{\circ\circ}}$. But since we know that $\gamma^\star_{\mathcal{S}^\circ} = \delta_{\mathcal{S}^{\circ\circ}}$ and the closedness of $\mathcal{S}^\circ$ implies that $\gamma_{\mathcal{S}^\circ} = \gamma^{\star\star}_{\mathcal{S}^\circ} = \delta^\star_{\mathcal{S}^{\circ\circ}}$, we then have $\gamma^\circ_{\mathcal{S}} = \gamma_{\mathcal{S}^\circ}$. This provides a justification for the name of "polar function" by expressing it as a gauge of the polar set $\mathcal{S}^\circ$. When $\gamma_{\mathcal{S}}$ defines a valid norm, $\gamma^\circ_{\mathcal{S}}$ is its dual norm.

Dually, by considering the polar function of the set $\mathcal{S}^\circ$ we can also write

$$\gamma_{\mathcal{S}}(x) = \gamma_{\mathcal{S}^{\circ\circ}}(x) = \sup_{y \in \mathcal{S}^\circ} \langle x, y \rangle \,, \tag{4.27}$$

where we note that $0 \in \mathcal{S}^\circ$ for any $\mathcal{S}$. We will refer to this as the dual formulation of the gauge function $\gamma_{\mathcal{S}}$.

The polarity operation induces a one-to-one symmetric correspondence between closed convex gauge functions, which means that any gauge of a compact convex set is uniquely determined by its polar. Note that the polarity operation is inclusion-reversing, i.e. $\mathcal{S}_1 \subseteq \mathcal{S}_2 \Rightarrow \mathcal{S}_1^\circ \supseteq \mathcal{S}_2^\circ$; this means that $\gamma_{\mathcal{S}_1}(x) \geq \gamma_{\mathcal{S}_2}(x)\ \forall x \in \mathrm{dom}(\gamma_{\mathcal{S}_1})$ and $\gamma^\circ_{\mathcal{S}_1}(y) \leq \gamma^\circ_{\mathcal{S}_2}(y)\ \forall y \in \mathrm{dom}(\gamma^\circ_{\mathcal{S}_2})$ for such sets.

We remark the useful fact that a closed and convex set $\mathcal{S}$ is bounded if and only if $0 \in \mathrm{int}(\mathcal{S}^\circ)$. For any compact and convex $\mathcal{S}$, $\gamma^\circ_{\mathcal{S}}$ then has full domain.

4.2.2 *Atomic gauges*

Following the terminology of [69] and subsequent works, given a compact set $\mathcal{C}$ we define the *atomic gauge* as the gauge function of the convex hull $\mathrm{conv}(\mathcal{C})$:

[69] Chandrasekaran et al., Found Comput Math 2012

$$\Gamma_{\mathcal{C}}(x) \coloneqq \gamma_{\mathrm{conv}(\mathcal{C})}(x). \tag{4.28}$$

The terminology comes from understanding $\mathcal{C}$ as a set of "atoms", often enjoying some simple structure which we want $\Gamma_{\mathcal{C}}$ to characterise. In particular, notice that $\Gamma_{\mathcal{C}}$ can be expressed as an optimisation over conic combinations of such atoms:

$$\begin{aligned}\Gamma_{\mathcal{C}}(x) &= \inf\left\{\lambda \;\middle|\; x = \lambda \sum_i c_i a_i,\ a_i \in \mathcal{C},\ c_i \in \mathbb{R}_+,\ \sum_i c_i = 1\right\} \\ &= \inf\left\{\sum_i c_i \;\middle|\; x = \sum_i c_i a_i,\ a_i \in \mathcal{C},\ c_i \in \mathbb{R}_+\right\}.\end{aligned} \tag{4.29}$$

Recall that all norms are gauge functions, but in addition many common norms can be conveniently expressed as an optimisation over simple sets of atoms — as two representative examples in $\mathbb{R}^{d\times d}$, consider the trace norm (Schatten 1-norm), which is given for $\mathcal{C} = \left\{xy^T \;\middle|\; x, y \in \mathbb{R}^d,\ \|x\|_{\ell_2} = \|y\|_{\ell_2} = 1\right\}$, and the operator norm (Schatten ∞-norm), which corresponds to $\mathcal{C}$ being the set of all orthogonal matrices. We will look into the structure of similar norms in the succeeding chapters, where a suitable choice of atoms will enable us to develop a general geometric description of quantum resources.

By Carathéodory's theorem, it suffices to consider combinations of at most $\dim(\mathcal{V})$ elements in (4.29). Note that in general $\Gamma_{\mathcal{C}}$ need not be symmetric or finite everywhere, and so it only defines a valid norm when the linear span of $\mathcal{C}$ is the whole space and $\mathcal{C}$ is centrally symmetric. The domain of $\Gamma_{\mathcal{C}}$ is the conic hull of $\mathcal{C}$, given by $\mathbb{R}_+ \operatorname{conv}(\mathcal{C}) = \mathcal{C}^{**}$. Notice that $\operatorname{conv}(\mathcal{C})^\circ = \mathcal{C}^\circ$ and $\operatorname{conv}(\mathcal{C})^\star = \mathcal{C}^\star$.

An important property is how atomic gauges of intersections and unions can be characterised using the gauges of their constituent parts. Let us consider arbitrary compact set $\mathcal{C}$ and $\mathcal{D}$. It is then easy to see that

$$\gamma_{\mathcal{C}^{\circ\circ}\cap\mathcal{D}^{\circ\circ}} = \max\{\Gamma_{\mathcal{C}}, \Gamma_{\mathcal{D}}\}, \tag{4.30}$$

and, more generally, $(\operatorname{conv}\mathcal{C} \cap \operatorname{conv}\mathcal{D})^\circ = \operatorname{conv}(\mathcal{C}^\circ \cup \mathcal{D}^\circ)$. To consider $\Gamma_{\mathcal{C}\cup\mathcal{D}}$, we first observe the general composition rule $(\mathcal{C}\cup\mathcal{D})^\circ = \mathcal{C}^\circ \cap \mathcal{D}^\circ$, using which we get

$$\Gamma^\star_{\mathcal{C}\cup\mathcal{D}} = \delta_{(\mathcal{C}\cup\mathcal{D})^\circ} = \delta_{\mathcal{C}^\circ\cap\mathcal{D}^\circ} = \delta_{\mathcal{C}^\circ} + \delta_{\mathcal{D}^\circ} = \Gamma^\star_{\mathcal{C}} + \Gamma^\star_{\mathcal{D}}. \tag{4.31}$$

Notice that the conditions $\operatorname{relint}(\operatorname{dom}\delta_{\mathcal{C}^\circ}) \cap \operatorname{relint}(\operatorname{dom}\delta_{\mathcal{D}^\circ}) \neq \varnothing$ in Lemma 4.1 reduce in this case to $\operatorname{relint}\mathcal{C}^\circ \cap \operatorname{relint}\mathcal{D}^\circ \neq \varnothing$. This is satisfied for any two compact $\mathcal{C}$ and $\mathcal{D}$, since from their compactness it follows that 0 is in the interior of their polars, and thus

$$\Gamma_{\mathcal{C}\cup\mathcal{D}} = \Gamma_{\mathcal{C}} \mathbin{\dot\vee} \Gamma_{\mathcal{D}}. \tag{4.32}$$

This extends also to the union of more than two sets, in the sense that

$$\Gamma_{\mathcal{C}_1\cup\cdots\cup\mathcal{C}_n} = \Gamma_{\mathcal{C}_1} \mathbin{\dot\vee} \cdots \mathbin{\dot\vee} \Gamma_{\mathcal{C}_n}. \tag{4.33}$$

Note that the conditions in Lemma 4.1 can be satisfied even if one of the sets is unbounded — as a simple example, the above composition relation holds for n

sets $\{C_i\}$, $n-1$ of which are compact, with at most one $\mathcal{C}_n$ being unbounded, closed, and such that $\mathcal{C}_n^\circ \neq \{0\}$ (i.e. $\mathcal{C}_n$ is not the whole space), since this is sufficient to guarantee the existence of a point in the relative interior of $\mathcal{C}_n^\circ$ which is in the neighbourhood of the origin (and thus in the interior of the other $\mathcal{C}_i^\circ$).

We will hereafter mostly work with atomic gauges, and for simplicity we will often write Γ_C even if $\mathcal{C}$ is already convex and the notation is superfluous.

4.2.3 *Complex vector spaces*

In a complex vector space, we define the polar with respect to the real inner product $\operatorname{Re}\langle x, y\rangle$, and similarly for other definitions. A generalisation of the concept of a centrally symmetric set is then a *balanced set*, that is $\mathcal{S}$ s.t. $s \in \mathcal{S} \Rightarrow \xi s \in \mathcal{S}\ \forall \xi \in \mathbb{C} : |\xi| = 1$. In particular, for a balanced set we have $\Gamma_{\mathcal{S}}(x) = \Gamma_{\mathcal{S}}(\xi x)$ for all ξ as above, and we can then take a simplified definition of a polar function:

$$\Gamma_{\mathcal{S}}^\circ(x) = \left[\sup_{a\in\mathcal{S}} \operatorname{Re}\langle x, a\rangle\right]_+ = \sup_{a\in\mathcal{S}} |\langle x, a\rangle| . \tag{4.34}$$

Any convex, balanced, compact set which spans the whole space defines a valid norm. We will henceforth only encounter balanced sets in complex vector spaces, so we do not expand on other properties and definitions — see e.g. [70] for more general cases.

[70] Bourbaki, 2003

4.3 CONVEX OPTIMISATION AND DUALITY

There are many equivalent ways to express general convex optimisation problems (see e.g. [63,71]). In the following, we will derive a particularly useful form which allows us to easily represent many problems which we will encounter later, while retaining full generality.

[63] Rockafellar, 1970

[71] Boyd & Vandenberghe, 2004

In the most general sense, convex optimisation is simply the minimisation of a convex function over a convex set. That is, any problem which we can express as

$$\widetilde{c} \coloneqq \inf_{x\in\mathcal{C}} f(x) \tag{4.35}$$

for some convex set $\mathcal{C}$ and a convex objective function $f : \mathcal{V} \to \mathbb{R}$ is a convex optimisation problem. The value $\widetilde{c}$ is the *optimal value*, and the set of all points which satisfy the constraint of the given problem — in this case, all $x \in \mathcal{C}$ — is referred to as the *feasible set*. We will denote by $\widetilde{x}$ any feasible x which achieves the infimum, that is, $\widetilde{c} = f(\widetilde{x})$. Note that this optimal solution does not need to exist, and even if it does, it does not have to be unique. In the case where $\widetilde{c} = -\infty$, the problem is called unbounded, and if $\mathcal{C} = \varnothing$, we have $\widetilde{c} = \infty$ and the problem is referred to as infeasible.

The first observation we can make is that one can, without loss of generality, take f to be linear. This is because the program defined as

$$\inf_{x,t} \left\{ t \in \mathbb{R} \;\middle|\; f(x) \leq t,\ x \in \mathcal{C} \right\} \tag{4.36}$$

can be straightforwardly seen to be equivalent to (4.35), in the sense that every feasible solution to (4.35) gives a feasible solution $(x, f(x))$ to problem (4.36) and vice versa, and similarly every optimal value of one problem gives a feasible optimal value of the other, hence they must be equal. Noting that (4.36) is a valid convex optimisation problem, since the feasible set is obtained from the epigraph of the convex function f over the convex set $\mathcal{C}$, we can equivalently redefine convex optimisation to be the minimisation of a linear objective function. This means that we can write any convex optimisation problem as

$$\inf_{x \in \mathcal{C}} \langle a, x \rangle \tag{4.37}$$

for some fixed a and a convex set $\mathcal{C}$ by suitably enlarging the optimisation space from the original vector space $\mathcal{V}$ to $\mathcal{V} \oplus \mathbb{R}$.

A particularly important class of convex optimisation problems are the so-called *conic programs*, where the feasible set takes the form $\mathcal{K} \cap \mathcal{F}$ with $\mathcal{K}$ being a convex cone and $\mathcal{F}$ an affine set. Crucially, it is important to note that we can consider only sets of this form without any loss of generality: indeed, any general convex optimisation problem of the form (4.37) with $a, x \in \mathcal{V}$ can be expressed as the conic program

$$\inf_{x' \in \mathcal{K} \cap \mathcal{F}} \langle a', x' \rangle \tag{4.38}$$

in the enlarged space $\mathcal{V} \oplus \mathbb{R}$, where $\mathcal{K} \coloneqq \left\{ (c, t) \;\middle|\; c \in t\mathcal{C},\ t \in \mathbb{R}_+ \right\}$, $\mathcal{F}$ is defined as $\left\{ (r, s) \;\middle|\; r \in \mathcal{V},\ s = 1 \right\}$, and $a' \coloneqq a \oplus 0$.

Let us therefore consider a general convex optimisation problem defined as

$$\widetilde{p} \coloneqq \inf \left\{ \langle a, x \rangle \;\middle|\; \Lambda(x) = b,\ x \in \mathcal{K} \right\} \tag{P}$$

where $\mathcal{K} \subseteq \mathcal{V}$ is a convex cone, $\Lambda : \mathcal{V} \to \mathcal{V}'$ is an arbitrary linear map, and $a \in \mathcal{V}$, $b \in \mathcal{V}'$ are fixed. To simplify some of the considerations, we will assume that $\mathcal{K}$ is closed.

We will hereafter refer to the program (P) as the *primal* problem, and use $\mathcal{F}_P \coloneqq \left\{ x \;\middle|\; \Lambda(x) = b,\ x \in \mathcal{K} \right\}$ to denote its feasible set. Notice that $\mathcal{F}_P$ is the intersection of a closed convex cone and an affine (and thus closed) set, and so it is also closed.

4.3.1 *Lagrange duality*

Given the primal in (P), we define the *Lagrangian* for this problem as

$$\mathcal{L}(x; \mu, \eta) \coloneqq \langle a, x \rangle - \langle \mu, \Lambda(x) - b \rangle - \langle \eta, x \rangle \tag{4.39}$$

where $\mu \in \mathcal{V}'$ and $\eta \in \mathcal{V}$ are referred to as Lagrange multipliers. It follows immediately from this definition that, for any feasible $x \in \mathcal{F}_P$, any $\mu \in \mathcal{V}'$, and any $\eta \in \mathcal{K}^\star$, we necessarily have

$$\langle a, x \rangle \geq \mathcal{L}(x; \mu, \eta). \tag{4.40}$$

Further, for any $x \notin \mathcal{F}_P$ the sets $\mathcal{F}_P$ and $\{x\}$ can be strongly separated by the closedness of $\mathcal{F}_P$, and we thus have

$$\sup_{\eta \in \mathcal{K}^\star} \mathcal{L}(x; \mu, \eta) = \begin{cases} \langle a, x \rangle & \text{if } x \in \mathcal{F}_P \\ \infty & \text{otherwise} \end{cases} \tag{4.41}$$

which in particular tells us that

$$\widetilde{p} = \inf_x \sup_{\eta \in \mathcal{K}^\star} \mathcal{L}(x; \mu, \eta), \tag{4.42}$$

where μ is unconstrained. The *Lagrange dual* of the problem (P) is then defined by interchanging the minimisation and maximisation in the above:

$$\widetilde{d} \coloneqq \sup_{\eta \in \mathcal{K}^\star} \inf_x \mathcal{L}(x; \mu, \eta). \tag{4.43}$$

To provide a stronger motivation for this definition, one can intuitively think of the Lagrangian as a way to lower bound to the optimal value of the problem (P). Indeed, defining the so-called dual function as

$$g(\mu, \eta) \coloneqq \inf_x \mathcal{L}(x; \mu, \eta) \tag{4.44}$$

we have precisely the desired property: for any $\mu \in \mathcal{V}'$ and $\eta \in \mathcal{K}^\star$, it necessarily holds that $g(\mu, \eta) \leq \widetilde{p}$. The Lagrange dual is then defined simply as the best such lower bound.

In the case of the conic program (P), the dual function $g(\mu, \eta)$ can be computed exactly, simplifying the characterisation of the dual problem. Defining $\Lambda^\dagger : \mathcal{V}' \to \mathcal{V}$ as the adjoint map, that is, the unique linear map such that $\langle w, \Lambda(v) \rangle = \left\langle \Lambda^\dagger(w), v \right\rangle \ \forall v \in \mathcal{V},\ w \in \mathcal{V}'$, we can see that

$$\begin{aligned} g(\mu, \eta) &= \inf_x \left\langle x, a - \Lambda^\dagger(\mu) - \eta \right\rangle + \langle \mu, b \rangle \\ &= \begin{cases} \langle \mu, b \rangle & \text{if } a - \Lambda^\dagger(\mu) - \eta = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned} \tag{4.45}$$

since $\left\langle x, a - \Lambda^\dagger(\mu) - \eta \right\rangle$ is a linear function of x. Noting that the conditions $a - \Lambda^\dagger(\mu) - \eta = 0$ and $\eta \in \mathcal{K}^\star$ are together equivalent to $a - \Lambda^\dagger(\mu) \in \mathcal{K}^\star$, we then have the Lagrange dual as

$$\widetilde{d} = \sup \left\{ \langle \mu, b \rangle \;\middle|\; a - \Lambda^\dagger(\mu) \in \mathcal{K}^* \right\}. \tag{D}$$

The set $\mathcal{F}_D \coloneqq \left\{ \mu \;\middle|\; a - \Lambda^\dagger(\mu) \in \mathcal{K}^* \right\}$ is the dual feasible set. We remark the easily verifiable fact that the Lagrange dual of (D) is

$$\inf \left\{ \langle a, x \rangle \;\middle|\; \Lambda(x) = b,\ x \in \mathcal{K}^{**} \right\} \tag{4.46}$$

and is therefore equal to the primal.

A particularly useful class of optimisation problems are *semidefinite programs* (SDP), which correspond to the case when $\mathcal{V} = \mathbb{H}$ (the space of Hermitian matrices), $\mathcal{K} = \mathbb{H}_+$ (the cone of positive semidefinite matrices), and the inner product is given by $\langle A, B \rangle = \mathrm{Tr}(AB)$. In this particular case, the cone is in fact self-dual, i.e. $\mathcal{K}^* = \mathcal{K}$. Semidefinite programs are very commonly encountered in quantum information [27] and beyond [72], and their popularity is amplified by the fact that very efficient solvers exist which can solve SDPs to arbitrary precision in polynomial time (polynomial in the dimension of the problem and in $\log(\frac{1}{\varepsilon})$, where ε is the allowed additive error) [72].

[27] Watrous, 2018

[72] Vandenberghe & Boyd, SIAM Rev. 1996

4.3.2 *Strong duality*

The fact that we always have $\widetilde{d} \leq \widetilde{p}$ is referred to as weak duality. We have seen that it holds for all convex optimisation problems, and in fact it extends also to non-convex primal problems. However, a very desirable property is the so-called *strong duality*, that is, the equality $\widetilde{p} = \widetilde{d}$. One of the most appealing aspects of convex optimisation problems is precisely the fact that the conditions under which strong duality occurs are easier to characterise. Among them, one of the most useful is Slater's condition, formalised as follows.

Theorem 4.2 (Slater's condition). *If both the primal and dual feasible sets are non-empty and at least one of them has a non-empty relative interior, then strong duality holds.*

See e.g. [64,71] for a proof.

[64] Hiriart-Urruty & Lemaréchal, 1993

[71] Boyd & Vandenberghe, 2004

Strong duality is a particular case of a so-called minimax theorem, allowing us to interchange the maximisation and minimisation in an optimisation problem without affecting its optimal value. We will in particular use a general minimax formulation from [73], although in simplified form, which we state as follows.

[73] Sion, Pac. J. Math. 1958

Theorem 4.3 (Sion's minimax theorem). *Let $\mathcal{C}$ and $\mathcal{D}$ be convex sets, at least one of which is compact. Then*

$$\inf_{x \in \mathcal{C}} \sup_{y \in \mathcal{D}} \langle x, y \rangle = \sup_{y \in \mathcal{D}} \inf_{x \in \mathcal{C}} \langle x, y \rangle. \tag{4.47}$$

4.4 ROOF EXTENSIONS

The intuition behind the convex roof approach is due to Uhlmann [48,74]. He describes the basic idea of the procedure as a way to extend entanglement measures "as linearly as possible" to the set of mixed states. It is formalised as follows.

[48] Uhlmann, OSID 1998

[74] Uhlmann, Entropy 2010

Let $\mathcal{S}$ be a compact convex set, and let $\mathrm{ext}(\mathcal{S})$ be the set of its extreme points, such that $\mathcal{S} = \mathrm{conv}(\mathrm{ext}(\mathcal{S}))$. Given a function f defined on $\mathrm{ext}(\mathcal{S})$, we now would like to consider a function $\widetilde{f}$ defined on the whole set $\mathcal{S}$ such that the two functions are equal on $\mathrm{ext}(\mathcal{S})$ — we will call such functions extensions of f. Extensions of this kind were briefly investigated in the mathematical literature before Uhlmann's work [75,76] and have also received attention in other areas after finding use in quantum information [77–80]. A particularly useful class of them is defined as follows:

[75] Lima, Proc. LMS 1972

[76] Peters & Wakker, Econ. Lett. 1986

[77] Tawarmalani & Sahinidis, Math. Program. 2002

[78] Bucicovschi & Lebl, J. Convex Anal. 2013

[79] Yan, J. Convex Anal. 2014

[80] Shcherbatyi & Andres, 2016

Definition 4.4. A function $\widetilde{f} : \mathcal{S} \to \mathbb{R}_+$ is a **roof extension** of $f : \mathrm{ext}(\mathcal{S}) \to \mathbb{R}_+$ if for every point $x \in \mathcal{S}$ there exists at least one extremal convex decomposition $\{c_i, \pi_i\}$ of the form $x = \sum_i c_i \pi_i$ with $c_i \in \mathbb{R}_+$, $\sum_i c_i = 1$, $\pi_i \in \mathrm{ext}(\mathcal{S})$, such that $\widetilde{f}(x) = \sum_i c_i f(\pi_i)$.

The justification for the terminology of *roofs* can be understood by noting that if we know a set of extremal points $\{\pi_1, \ldots, \pi_k\}$ which constitutes an optimal decomposition for some $x \in \mathcal{S}$, the roof extension $\widetilde{f}$ will be affine on the convex hull of these points — that is, the same set of points will be an optimal decomposition for any other convex combination thereof. This means that the graph of $\widetilde{f}$ consists of flat (affine) pieces covering the set $\mathcal{S}$, not unlike a roof covering a floor (see Fig. 4.3).

This definition, however, might seem arbitrary — since there are no restrictions on the choice of a roof, it does not tell us much about the properties of the function f. Here is where the convexity of $\widetilde{f}$ comes useful.

Theorem 4.5 (Uhlmann [48]). *If a convex roof extension $f^\cup$ of $f : \mathrm{ext}(\mathcal{S}) \to \mathbb{R}_+$ onto $\mathcal{S}$ exists, then it is unique and satisfies the following properties:*

[48] Uhlmann, OSID 1998

(i) $f^\cup$ is the pointwise largest convex extension of f from $\mathrm{ext}(\mathcal{S})$ to $\mathcal{S}$

(ii) $f^\cup$ is the pointwise smallest roof extension of f from $\mathrm{ext}(\mathcal{S})$ to $\mathcal{S}$

(iii) $f^\cup$ is given by

$$f^\cup(x) = \inf\left\{ \sum_i c_i f(\pi_i) \;\middle|\; \sum_i c_i \pi_i = x,\ \sum_i c_i = 1,\ c_i \in \mathbb{R}_+ \right\} \tag{4.48}$$

with the infimum taken over all extremal convex decompositions $\{c_i, \pi_i\}$ with each $\pi_i \in \mathrm{ext}(\mathcal{S})$.

The result tells us that finding the largest convex function on $\mathcal{S}$ which coincides with f on $\mathrm{ext}(\mathcal{S})$ is equivalent to finding the unique convex roof extension of f. For a lower semicontinuous function f, an optimal decomposition realising the infimum in eq. (4.48) always exists.

Since we will be dealing with functions $f : \mathrm{ext}(\mathcal{S}) \to \mathbb{R}_+ \cup \{\infty\}$, we extend the definition by taking $f^{\cup}(x) = \infty$ when $x \notin \mathrm{conv}(\mathrm{dom}\, f)$. For a function g already defined on the whole set $\mathcal{S}$, we can define its convex roof extension as the convex roof extension of the restriction of g to $\mathrm{ext}(\mathcal{S})$ — if g is convex, it is then clear from the above Theorem that $g^{\cup}(x) \geq g(x)\ \forall x$.

Note also the related concept of the *concave roof* $f^{\cap}$: it is the pointwise smallest concave extension (or pointwise largest roof extension) of f, and can be calculated by replacing the minimisation with a maximisation in Eq. (4.48):

$$f^{\cap}(x) = \sup_{\{c_i,\pi_i\}} \left\{ \sum_i c_i\, f(\pi_i) \;\middle|\; \sum_i c_i\, \pi_i = x,\ \sum_i c_i = 1,\ c_i \in \mathbb{R}_+ \right\}. \tag{4.49}$$

One can notice a similarity between the concept of the convex roof and the *convex envelope* (also known as the *convex hull* or the *largest convex underestimator*). For a function g which is already defined on the whole convex set $\mathcal{S}$, its convex envelope is the largest convex function smaller than g. In practice, finding the convex envelope means that the optimisation in eq. (4.48) is performed over all decompositions of x into $s_i \in \mathcal{S}$, instead of $\pi_i \in \mathrm{ext}(\mathcal{S})$ as in the case of the convex roof extension. For this reason, in quantum information literature the convex envelope has sometimes been referred to as the *mixed convex roof*, identifying $\mathrm{ext}(\mathcal{S})$ with the set of pure quantum states. Note that the convex envelope of a concave function defined on $\mathcal{S}$ coincides with its convex roof extension, since the infimum of any concave function over a bounded convex set is equal to the infimum over its extremal points [63, 32.2].

[63] Rockafellar, 1970

A simple relation between the convex roof and the formalism of atomic gauges can be obtained as follows.

Proposition 4.6. *Given a compact set $\mathcal{C} \subseteq \mathrm{ext}(\mathcal{S})$, for any $x \in \mathcal{S}$ we have*

$$\Gamma_{\mathcal{C}}(x) = \gamma^{\cup}_{\mathcal{C}}(x). \tag{4.50}$$

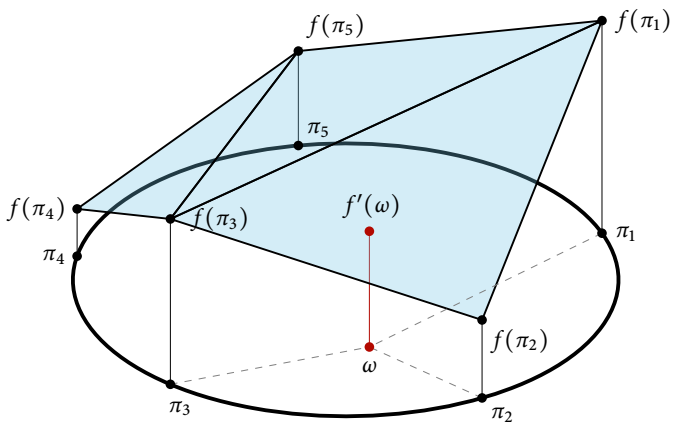


Figure 4.3: Constructing a roof extension. The roof $\widetilde{f}$ is constructed for a choice of points $\pi_i \in \operatorname{ext}(\mathcal{S})$ and a function $f : \operatorname{ext}(\mathcal{S}) \to \mathbb{R}$, where $\operatorname{ext}(\mathcal{S})$ is the circle. The value of $\widetilde{f}$ at any $\omega \in \mathcal{S}$ is given by the height of the "roof" above it.

Proof. The atomic gauge $\Gamma_{\mathcal{C}}$ can be written as

$$\begin{aligned}
\Gamma_{\mathcal{C}}(x) &= \inf\left\{ \sum_i c_i \;\middle|\; x = \sum_i c_i x_i,\ c_i \in \mathbb{R}_+,\ x_i \in \mathcal{C} \right\} \\
&\overset{\text{(i)}}{=} \inf\left\{ \sum_i c_i\, \gamma_{\mathcal{C}}(x_i) \;\middle|\; x = \sum_i c_i x_i,\ c_i \in \mathbb{R}_+,\ x_i \in \mathcal{C} \right\} \\
&\overset{\text{(ii)}}{=} \inf\left\{ \sum_i c_i\, \gamma_{\mathcal{C}}(x_i) \;\middle|\; x = \sum_i c_i x_i,\ c_i \in \mathbb{R}_+,\ \sum_i c_i = 1,\ x_i \in \mathcal{C}^{**} \right\} \\
&\overset{\text{(iii)}}{=} \inf\left\{ \sum_i c_i\, \gamma_{\mathcal{C}}(x_i) \;\middle|\; x = \sum_i c_i x_i,\ c_i \in \mathbb{R}_+,\ \sum_i c_i = 1,\ x_i \in \operatorname{ext}(\mathcal{S}) \right\} \\
&= \gamma_{\mathcal{C}}^{\cup}(x).
\end{aligned} \tag{4.51}$$

The line (i) follows since, without loss of generality, we can consider only x_i such that $\gamma_{\mathcal{C}}(x_i) = 1$. To see this, notice that for any $c_i' x_i'$ with $\gamma_{\mathcal{C}}(x_i') \in (0,1)$ we can define $x_i \coloneqq x_i'/\gamma_{\mathcal{C}}(x_i')$ and $c_i \coloneqq c_i' \gamma_{\mathcal{C}}(x_i')$ such that $c_i' \gamma_{\mathcal{C}}(x_i') = c_i \gamma_{\mathcal{C}}(x_i) = c_i$. The step (ii) follows similarly, and (iii) is a straightforward consequence of the fact that $\gamma_{\mathcal{C}}(x_i) = \infty$ for any $x_i \notin \mathcal{C}^{**}$, so the optimisation is effectively constrained to be over $x_i \in \mathcal{C}^{**}$, in the sense that the infimum (if finite) will necessarily be given for x in a decomposition into $x_i \in \mathcal{C}^{**}$. ∎

We remark that, following a very similar reasoning as in the proof above, one can write any gauge function of this form as

$$\Gamma_{\mathcal{C}}(x) = \inf\left\{ \sum_i \gamma_{\mathcal{C}}(x_i) \;\middle|\; x = \sum_i x_i,\ x_i \in \operatorname{ext}(\mathcal{S}) \right\}. \tag{4.52}$$

The applications of this result and further relations of this kind will be explored in subsequent chapters.

Part II

Convex geometry of quantum resources

Gauges for sets of quantum states

5

Throughout the development of the resource theory of entanglement, various measures were established [7,81], many of which have been adapted to other resource theories recently [3,6,15,16,19,82,83]. However, defining and characterising the measures of a given quantum resource is usually a cumbersome task — the investigation of such functions typically has to be approached in a resource-dependent way, and desired properties of the quantifiers have to be explicitly verified for every resource in consideration. Although some connections between the various quantities are known, there are very few known results which provide a common framework relating them and their features together.

In this chapter, we introduce a unifying formalism based on the convex geometry of the underlying sets of quantum states, significantly simplifying the construction and characterisation of quantifiers of general quantum resources. We employ the concept of gauge functions to establish a consolidated view of many resource measures. In particular, we show that many commonly used and well-known quantifiers — such as ones based on matrix norms, measures built through the convex roof, the so-called robustness measures, as well as various witness-based quantifiers — are all examples of such gauge functions, allowing us to relate them in a common geometric framework. This lets us establish an extensive family of quantifiers for any given quantum resource, introduce easily verifiable criteria for a measure to satisfy desirable properties such as faithfulness and strong monotonicity under relevant free operations, and generalise known measures to new quantum resources very easily. Further, we show that many relations and bounds between the measures, some of which known in the resource theories of entanglement and coherence, are in fact universal among quantum resources, and the proofs of such properties can be simplified in the present framework.

To begin with, recall that the essential element of the characterisation of a general resource theory are the free states, that is, states not possessing a given resource. We therefore begin by defining the set of resource-free normalised pure state vectors $\mathcal{V} \subseteq \mathbb{C}^d$, which we will assume to be non-empty. Another intuitive assumption is that the set $\mathcal{V}$ should be compact, which ensures the continuity of the given resource theory [3]. Since the global phase factor $e^{i\theta}|\psi\rangle$ of a quantum state $|\psi\rangle$ is physically irrelevant, we further assume that $|\psi\rangle \in \mathcal{V}$ implies $e^{i\theta}|\psi\rangle \in \mathcal{V}\ \forall\theta \in \mathbb{R}_+$, which means the set is balanced and so $0 \in \mathrm{conv}(\mathcal{V})$.

[7] Horodecki et al., RMP 2009
[81] Plenio & Virmani, QIC 2007
[3] Brandão & Gour, PRL 2015
[6] Gour, PRA 2017
[15] Baumgratz et al., PRL 2014
[16] Streltsov et al., RMP 2017
[19] Theurer et al., PRL 2017
[82] Piani et al., PRA 2016
[83] Howard & Campbell, PRL 2017

We then get the atomic gauge

$$\Gamma_{\mathcal{V}}(|\psi\rangle) = \inf\left\{ \sum_i c_i \;\middle|\; |\psi\rangle = \sum_i c_i |v_i\rangle\,,\ |v_i\rangle \in \mathcal{V},\ c_i \in \mathbb{R}_+ \right\} \tag{5.1}$$
$$= \sup_{|x\rangle \in \mathcal{V}^\circ} |\langle \psi | x \rangle|$$
$$\Gamma^\circ_{\mathcal{V}}(|\psi\rangle) = \sup_{|y\rangle \in \mathcal{V}} |\langle \psi | y \rangle|\,. \tag{5.2}$$

Under the assumptions above, $\Gamma_{\mathcal{V}}$ is a valid norm for $\mathbb{C}^d$ whenever $0 \in \operatorname{int}(\operatorname{conv} \mathcal{V})$, i.e. $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$. This property is very desirable for the set $\mathcal{V}$, as we will see later, but we do not assume it to hold in all cases. Note that, by the compactness of $\mathcal{V}$, the infima and suprema in (5.1)-(5.2) will be achieved as long as they are finite.

Given the set $\mathcal{V}$ defined at the level of vectors, we now define our set of interest: the set of resource-free density matrices, given by the convex hull of the set

$$\mathcal{S}_+ \coloneqq \left\{ |v\rangle\langle v| \;\middle|\; |v\rangle \in \mathcal{V} \right\} \subseteq \mathbb{D}. \tag{5.3}$$

Since all density matrices which do not lie in $\operatorname{conv}(\mathcal{S}_+)$ are resourceful states in the given resource theory, we would now like to use the set $\mathcal{S}_+$ to introduce a gauge function which could effectively quantify this resource. The atomic gauge $\Gamma_{\mathcal{S}_+}$ itself is not very useful: since the domain of this function is $\mathcal{S}_+^{\star\star} = \mathbb{R}_+ \operatorname{conv}(\mathcal{S}_+)$ (the convex cone generated by $\mathcal{S}_+$) and all density operators lie in the hyperplane defined by $\langle \mathbb{1}, \rho \rangle = 1$, there are actually no density operators in the domain of $\Gamma_{\mathcal{S}_+}$ which do not lie in the convex hull of $\mathcal{S}_+$. That is, for any density matrix ρ we have

$$\begin{aligned} \Gamma_{\mathcal{S}_+}(\rho) &\leq 1 \quad \text{if } \rho \in \operatorname{conv}(\mathcal{S}_+), \\ \Gamma_{\mathcal{S}_+}(\rho) &= \infty \quad \text{if } \rho \notin \operatorname{conv}(\mathcal{S}_+). \end{aligned} \tag{5.4}$$

We will now look at ways of circumventing this problem by making a different choice of the gauge function to use.

This chapter is based on results published in [BR4]. [BR4] Regula, JPA 2018

5.1 SELECTION OF GAUGES

Throughout this chapter, we will be working in several different vector spaces: the underlying Hilbert space, identified with $\mathbb{C}^d$; the real vector space of self-adjoint linear operators, idenitifed with the space of Hermitian matrices $\mathbb{H}$ and endowed with the real inner product $\langle A, B \rangle = \operatorname{Tr}(AB)$; as well as the complex vector space of linear operators acting on $\mathbb{C}^d$, identified with the space of complex square matrices $\mathbb{C}^{d\times d}$ and endowed with the (degenerate) real inner product $\operatorname{Re}\langle A, B \rangle = \operatorname{Re}\operatorname{Tr}(A^\dagger B)$. Unless otherwise specified, the default setting will be the real vector space $\mathbb{H}$.

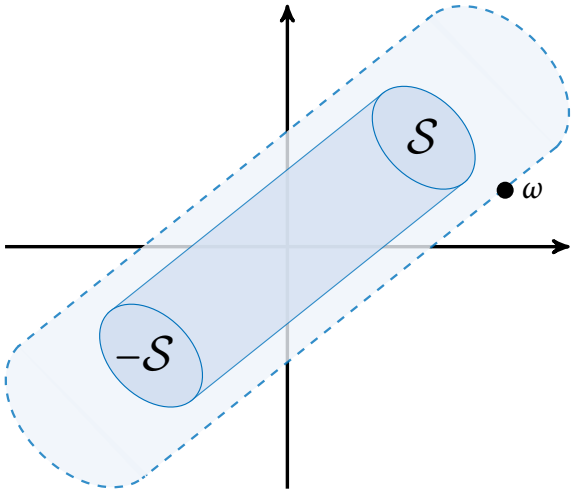


Figure 5.1: Extending the domain of a gauge. The point ω, which lies outside of the domain of the gauge $\gamma_{\mathcal{S}}$ (cf. Fig. 4.1), is included in the domain of $\Gamma_{\mathcal{S}\cup(-\mathcal{S})}$.

We will use $\mathbb{H}_1$ to denote the set $\left\{ X \in \mathbb{H} \;\middle|\; \langle \mathbb{1}, X\rangle = 1 \right\}$ of unit trace Hermitian matrices, $\mathbb{H}_+$ to denote the cone of positive semidefinite matrices, and $\mathbb{D} = \mathbb{H}_1 \cap \mathbb{H}_+$ to denote the set of valid density matrices. We will use $|x\rangle$ to denote vectors in the Hilbert space which are not necessarily normalised, and reserve the standard Greek letters ψ, ϕ etc. to represent normalised pure states. Similarly, Greek letters ρ, σ, ω will denote normalised density operators.

Our aim now is to define a choice of non-trivial atomic gauges which provide a natural way to quantify the given resource.

An intuition for the possible choices of suitable quantifiers can be obtained by looking at the simple example where $\mathcal{V}$ is taken to be the set of *all* normalised pure state vectors, and thus $\Gamma_{\mathcal{V}}(\cdot) = \|\cdot\|_{\ell_2}$. In this case, we have $\mathrm{conv}(\mathcal{S}_+) = \mathbb{D}$, and the gauge $\Gamma_{\mathbb{D}}$ can be thought of as the trace norm $\|\cdot\|_1$, but defined only on the limited domain $\mathbb{R}_+\mathbb{D} = \mathbb{H}_+$. It is straightforward to see that the convex roof extension of $\Gamma_{\mathcal{V}}$ corresponds to the same function. A natural way to extend the domain of this gauge beyond $\mathbb{H}_+$ is to symmetrise the set under consideration, and consider the gauge $\Gamma_{\mathbb{D}\cup(-\mathbb{D})}$ instead — since the origin is now contained in the convex hull of the set, the function takes finite values for all Hermitian matrices (see Fig. 5.1). In fact, $\Gamma_{\mathbb{D}\cup(-\mathbb{D})}$ is simply equal to the trace norm itself [84], with the domain now extended to $\mathbb{H}$. One can go a step further and, instead of limiting ourselves to the real vector space of Hermitian matrices, define the set $\mathcal{S} = \left\{ |a\rangle\langle b| \;\middle|\; |a\rangle, |b\rangle \in \mathcal{V} \right\}$. The gauge $\Gamma_{\mathcal{S}}$ is then precisely the trace norm $\|\cdot\|_1$ in its most general formulation, which has full domain in $\mathbb{C}^{d\times d}$.

[84] Hartkämper & Neumann, 1974

In these simple examples, we have seen that different ways of extending the domain of the gauge $\Gamma_{\mathcal{S}_+}$ actually correspond to the same function (on their

effective domains), and so there is little ambiguity in the choice of the appropriate extension. This does not, however, hold in more general cases, and so it becomes useful to be able to characterise all the different possible gauges that we can obtain from the initial set $\mathcal{V}$.

Returning to the case of general quantum resources, we can take inspiration from the above procedure and consider the atomic gauge with respect to the symmetrised set $\mathcal{S}_+ \cup (-\mathcal{S}_+)$. This extends the domain of the gauge function to $\operatorname{span}(\mathcal{S}_+)$, that is, the vector space generated by $\mathcal{S}_+$. We then get the gauge

$$\begin{aligned}\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(\rho) &= \inf_{S_i\in\mathcal{S}_+\cup(-\mathcal{S}_+)}\left\{\sum_i c_i \;\middle|\; \rho = \sum_i c_i\, S_i,\ c_i\in\mathbb{R}_+\right\}\\ &= \inf_{|v_i\rangle\in\mathcal{V}}\left\{\sum_i |c_i| \;\middle|\; \rho = \sum_i c_i\, |v_i\rangle\langle v_i|,\ c_i\in\mathbb{R}\right\}\\ \Gamma^\circ_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(\rho) &= \sup\left\{|\langle v|\rho|v\rangle| \;\middle|\; |v\rangle\in\mathcal{V}\right\}.\end{aligned} \tag{5.5}$$

Analogously to Prop. 4.6, the above can be equivalently expressed as

$$\begin{aligned}\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(\rho) &= \inf_{X_i\in\mathbb{H}}\left\{\sum_i \Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(X_i) \;\middle|\; \rho = \sum_i X_i,\ \operatorname{rank}(X_i)=1\right\}\\ &= \inf_{|x_i\rangle\in\mathbb{C}^d}\left\{\sum_i |c_i|\,\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(|x_i\rangle\langle x_i|) \;\middle|\; \rho = \sum_i c_i\,|x_i\rangle\langle x_i|,\ c_i\in\mathbb{R}\right\}.\end{aligned} \tag{5.6}$$

$\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}$ defines a valid norm for Hermitian matrices as long as $\operatorname{span}(\mathcal{S}_+) = \mathbb{H}$.

Another straightforward way to define a quantifier is to consider instead the set

$$\mathcal{S} \coloneqq \left\{|a\rangle\langle b| \;\middle|\; |a\rangle\in\mathcal{V},\, |b\rangle\in\mathcal{V}\right\}, \tag{5.7}$$

which we note to be balanced in the complex vector space $\mathbb{C}^{d\times d}$ with the Hilbert-Schmidt inner product $\operatorname{Re}\langle\cdot,\cdot\rangle$. The corresponding atomic gauge is:

$$\begin{aligned}\Gamma_{\mathcal{S}}(\rho) &= \inf_{|v_i\rangle,|w_i\rangle\in\mathcal{V}}\left\{\sum_i c_i \;\middle|\; \rho = \sum_i c_i\,|v_i\rangle\langle w_i|,\ c_i\in\mathbb{R}_+\right\}\\ \Gamma^\circ_{\mathcal{S}}(\rho) &= \sup\left\{|\langle v|\rho|w\rangle| \;\middle|\; |v\rangle, |w\rangle\in\mathcal{V}\right\}.\end{aligned} \tag{5.8}$$

This gauge admits an alternative representation, as we show below. The result is a slight generalisation of a lemma in [85,86], where norms for real matrices are considered.

[85] Richard et al., 2014

[86] Richard et al., preprint 2014

Lemma 5.1. $\Gamma_{\mathcal{S}}(\rho) = \inf_{|x_i\rangle,|y_i\rangle\in\mathbb{C}^d}\left\{\sum_i \Gamma_{\mathcal{V}}(|x_i\rangle)\,\Gamma_{\mathcal{V}}(|y_i\rangle) \;\middle|\; \rho = \sum_i |x_i\rangle\langle y_i|\right\}.$

Proof. The function on the right-hand side can be noticed to be the gauge function of the set

$$\begin{aligned}\mathcal{S}' &\coloneqq \left\{|a\rangle\langle b| \;\middle|\; \Gamma_{\mathcal{V}}(|a\rangle)\le 1,\, \Gamma_{\mathcal{V}}(|b\rangle)\le 1\right\}\\ &= \left\{|a\rangle\langle b| \;\middle|\; |a\rangle, |b\rangle\in\mathcal{V}^{\circ\circ}\right\}.\end{aligned} \tag{5.9}$$

The straightforward inclusion $\mathcal{S} \subseteq \mathcal{S}'$ gives $\mathcal{S}^{\circ\circ} \subseteq \mathcal{S}'^{\circ\circ} = \mathcal{S}'$, as $\mathcal{S}'$ is already a closed convex set containing the origin. On the other hand, noticing that for any $|a\rangle\langle b| \in \mathcal{S}'$ we can write $|a\rangle = \sum_i c_i |a_i\rangle$ and $|b\rangle = \sum_j d_j |b_j\rangle$, where $c_i, d_j \in \mathbb{R}_+$, $|a_i\rangle, |b_j\rangle \in \mathcal{V}$, and $\sum_i c_i, \sum_j d_j \leq 1$, we get $|a\rangle\langle b| = \sum_{i,j} c_i d_j |a_i\rangle\langle b_j|$ with $\sum_{i,j} c_i d_j \leq 1$, which means that $|a\rangle\langle b| \in \mathcal{S}^{\circ\circ}$ and so $\mathcal{S}' \subseteq \mathcal{S}^{\circ\circ}$. The sets $\mathcal{S}'$ and $\mathcal{S}^{\circ\circ}$ are therefore equal, and so $\Gamma_{\mathcal{S}} = \Gamma_{\mathcal{S}'}$. ∎

Gauges of the form encountered on the right-hand side of Lemma 5.1 are referred to as nuclear gauges [66,87]. $\Gamma_{\mathcal{S}}$ defines a valid norm for $\mathbb{C}^{d\times d}$, called a generalised nuclear norm, as long as $\Gamma_{\mathcal{V}}$ is a norm for $\mathbb{C}^d$, i.e. $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$.

[66] Bach, preprint 2013

[87] Jameson, 1987

Considering the values of the gauge function $\Gamma_{\mathcal{V}}$ for pure states is appealing, as this function is typically significantly simpler to compute than gauges defined at the level of matrices. We can then extend $\Gamma_{\mathcal{V}}$ to general quantum states through the convex roof. To ensure homogeneity at the level of projectors $|\psi\rangle\langle\psi|$, we will take the extension of $\Gamma_{\mathcal{V}}(|\psi\rangle)^2$, giving

$$\begin{aligned}\Gamma^{\cup}_{\mathcal{S}_+}(\rho) &\coloneqq \left(\Gamma^2_{\mathcal{V}}\right)^{\cup}(\rho) \\ &= \inf\left\{ \sum_i p_i \Gamma_{\mathcal{V}}(|\psi_i\rangle)^2 \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \sum_i p_i = 1,\ p_i \in \mathbb{R}_+ \right\}\end{aligned} \tag{5.10}$$

where the optimisation is over normalised pure states. The convex roof extension can therefore be regarded as a natural extension of $\Gamma_{\mathcal{S}_+}$ out of its limited domain, since it will be finite for any density matrix as long as $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$. We can characterise it as a gauge function as follows.

Proposition 5.2. *For any $\rho \in \mathbb{D}$, the convex roof extension $\Gamma^{\cup}_{\mathcal{S}_+}(\rho)$ is equal to the atomic gauge $\Gamma_{\mathcal{S}^{\cup}_+}(\rho)$, where*

$$\mathcal{S}^{\cup}_+ \coloneqq \left\{ |a\rangle\langle a| \;\middle|\; |a\rangle \in \operatorname{conv}(\mathcal{V}) \right\}. \tag{5.11}$$

Proof. By Prop. 4.6, the gauge function of $\mathcal{S}^{\cup}_+$ can be written as

$$\Gamma_{\mathcal{S}^{\cup}_+}(\rho) = \inf\left\{ \sum_i p_i \Gamma_{\mathcal{V}}(|x_i\rangle)^2 \;\middle|\; \rho = \sum_i p_i |x_i\rangle\langle x_i|, \sum_i p_i = 1,\ p_i \in \mathbb{R}_+ \right\}, \tag{5.12}$$

where we now optimise over general rank-one terms. The result follows simply by noting that we can renormalise any such decomposition $\rho = \sum_i p_i |x_i\rangle\langle x_i|$ as $\rho = \sum_i p'_i |\psi_i\rangle\langle\psi_i|$ where $p'_i \coloneqq p_i \||x_i\rangle\|^2_{\ell_2}$ and $|\psi_i\rangle \coloneqq |x_i\rangle / \||x_i\rangle\|_{\ell_2}$, giving

$$\sum_i p'_i = \sum_i p_i \operatorname{Tr}|x_i\rangle\langle x_i| = \operatorname{Tr}\left(\sum_i p_i |x_i\rangle\langle x_i|\right) = \operatorname{Tr}\rho = 1. \tag{5.13}$$

Using the homogeneity of $\Gamma_{\mathcal{V}}$, this ensures that $\Gamma_{\mathcal{S}^{\cup}_+}(\rho) = \Gamma^{\cup}_{\mathcal{S}_+}(\rho)$. ∎

Before we proceed with the characterisation of the introduced atomic gauges, let us make explicit the effective domains on which the functions are finite.

Gauge	Effective domain	$\mathbb{D}$ in the effective domain?
$\Gamma_{\mathcal{V}}$	$\operatorname{span}(\mathcal{V})$	(not applicable)
$\Gamma^{\circ}_{\mathcal{V}}$	$\mathbb{C}^d$	(not applicable)
$\Gamma_{\mathcal{S}_+}$	$\mathcal{S}_+^{**}$	no, $\operatorname{dom}(\Gamma_{\mathcal{S}_+}) \cap \mathbb{D} = \operatorname{conv}(\mathcal{S}_+)$
$\Gamma_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}$	$\operatorname{span}(\mathcal{S}_+)$	if $\operatorname{relint}_{\mathbb{D}}(\operatorname{conv}\mathcal{S}_+) \neq \varnothing$
$\Gamma_{\mathcal{S}}$	$\operatorname{span}(\mathcal{S})$	if $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$
$\Gamma^{\cup}_{\mathcal{S}_+}$	$\mathcal{S}_+^{\cup **}$	if $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$
$\Gamma^{\circ}_{\mathcal{S}_+}, \Gamma^{\circ}_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}, \Gamma^{\cup \circ}_{\mathcal{S}_+}$	$\mathbb{H}$	yes
$\Gamma^{\circ}_{\mathcal{S}}$	$\mathbb{C}^{d \times d}$	yes

In the above, $\operatorname{relint}_{\mathbb{D}}(\operatorname{conv}\mathcal{S}_+)$ is the interior of $\operatorname{conv}(\mathcal{S}_+)$ relative to $\mathbb{D}$. The set $\mathcal{S}_+^{\cup **}$ can be understood as the set of positive semidefinite matrices supported on the subspace spanned by $\mathcal{V}$, i.e.

$$\mathcal{S}_+^{\cup **} = \left\{ X \in \mathbb{H}_+ \;\middle|\; \operatorname{supp}(X) \subseteq \operatorname{span}(\mathcal{V}) \right\}. \tag{5.14}$$

We also remark that in general we have

$$\begin{aligned} \Gamma_{\mathcal{S}_+}(\rho) &\geq \Gamma_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) \\ \Gamma_{\mathcal{S}_+}(\rho) &\geq \Gamma^{\cup}_{\mathcal{S}_+}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) \end{aligned} \tag{5.15}$$

and reverse inequalities for the polar gauges, which follows from the set inclusion of their corresponding unit balls.

Theorem 5.3. *For pure states $|\psi\rangle\langle\psi|$ s.t. $|\psi\rangle \in \operatorname{span}(\mathcal{V})$, we have*

$$\begin{aligned} \Gamma^{\cup}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) &= \Gamma_{\mathcal{V}}(|\psi\rangle)^2 \\ \Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|) &= \Gamma_{\mathcal{V}}(|\psi\rangle)^2. \end{aligned} \tag{5.16}$$

For arbitrary pure states we also have

$$\begin{aligned} \Gamma^{\circ}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) &= \Gamma^{\circ}_{\mathcal{V}}(|\psi\rangle)^2 \\ \Gamma^{\circ}_{\mathcal{S}}(|\psi\rangle\langle\psi|) &= \Gamma^{\circ}_{\mathcal{V}}(|\psi\rangle)^2. \end{aligned} \tag{5.17}$$

Proof. The case of $\Gamma^{\cup}_{\mathcal{S}_+}$ is obvious from the definition.

For $\Gamma_{\mathcal{S}}$, notice that by the compactness of $\mathcal{V}$ there exist $\lambda^\star \in \mathbb{R}_+$ and $|v^\star\rangle \in \operatorname{conv}(\mathcal{V})$ such that

$$\lambda^\star = \min\left\{ \lambda \in \mathbb{R}_+ \;\middle|\; |\psi\rangle \in \lambda \operatorname{conv}(\mathcal{V}) \right\} = \Gamma_{\mathcal{V}}(|\psi\rangle) \tag{5.18}$$

with $|\psi\rangle = \lambda^\star |v^\star\rangle$. This gives $|\psi\rangle\langle\psi| = \lambda^{\star 2} |v^\star\rangle\langle v^\star|$ with $|v^\star\rangle\langle v^\star| \in \operatorname{conv}(\mathcal{S})$, and so

$$\Gamma_{\mathcal{V}}(|\psi\rangle)^2 = \lambda^{\star 2} \geq \inf\left\{ \eta \in \mathbb{R}_+ \;\middle|\; |\psi\rangle\langle\psi| \in \eta \operatorname{conv}(\mathcal{S}) \right\} = \Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|). \tag{5.19}$$

Now, from the dual formulation of the atomic gauge we get

$$\begin{aligned} \Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|) &= \sup\left\{ |\langle |\psi\rangle\langle\psi| , X\rangle| \;\middle|\; X \in \mathcal{S}^\circ \right\} \\ &\geq \sup\left\{ |\langle |\psi\rangle\langle\psi| , |x\rangle\langle x|\rangle| \;\middle|\; |x\rangle \in \mathcal{V}^\circ \right\} \\ &= \sup\left\{ |\langle\psi|x\rangle|^2 \;\middle|\; |x\rangle \in \mathcal{V}^\circ \right\} \\ &= \Gamma_{\mathcal{V}}(|\psi\rangle)^2 \end{aligned} \tag{5.20}$$

since $|x\rangle \in \mathcal{V}^\circ \Rightarrow |x\rangle\langle x| \in \mathcal{S}^\circ$, and so the claim follows.

The case of $\Gamma^\circ_{\mathcal{S}_+}$ follows immediately from the definition. For $\Gamma^\circ_{\mathcal{S}}$, we have

$$\begin{aligned} \Gamma^\circ_{\mathcal{S}}(|\psi\rangle\langle\psi|) &= \sup\left\{ |\langle x|\psi\rangle \langle\psi|y\rangle| \;\middle|\; |x\rangle , |y\rangle \in \mathcal{V} \right\} \\ &= \left(\sup_{|x\rangle\in\mathcal{V}} |\langle x|\psi\rangle| \right) \left(\sup_{|y\rangle\in\mathcal{V}} |\langle y|\psi\rangle| \right) \\ &= \Gamma^\circ_{\mathcal{V}}(|\psi\rangle)\Gamma^\circ_{\mathcal{V}}(|\psi\rangle) \end{aligned} \tag{5.21}$$

since the optimisation is over two compact sets of non-negative real numbers. ∎

Remark 5.4. We can also note a stronger property that

$$\begin{aligned} \Gamma_{\mathcal{S}}(|\psi\rangle\langle\phi|) &= \Gamma_{\mathcal{V}}(|\psi\rangle)\, \Gamma_{\mathcal{V}}(|\phi\rangle) \\ \Gamma^\circ_{\mathcal{S}}(|\psi\rangle\langle\phi|) &= \Gamma^\circ_{\mathcal{V}}(|\psi\rangle)\, \Gamma^\circ_{\mathcal{V}}(|\phi\rangle). \end{aligned} \tag{5.22}$$

Proposition 5.5. *For positive semidefinite matrices P, we have*

$$\Gamma^\circ_{\mathcal{S}_+}(P) = \Gamma^\circ_{\mathcal{S}}(P) = \Gamma^\circ_{\mathcal{S}_+^\cup}(P) = \Gamma^\circ_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}(P) \tag{5.23}$$

In the proof of the Proposition, we will make use of the following lemma.

Lemma 5.6. *For any $P \in \mathbb{H}_+$, it holds that*

$$\sup\left\{ |\langle a|P|b\rangle| \;\middle|\; |a\rangle , |b\rangle \in \mathcal{V} \right\} = \sup\left\{ \langle a|P|a\rangle \;\middle|\; |a\rangle \in \mathcal{V} \right\}. \tag{5.24}$$

Proof. First, notice that the left-hand side clearly cannot be smaller than the right-hand side. By the compactness of $\mathcal{V}$, there exists a $|a^\star\rangle$ which realises

the maximum on the right-hand side. Writing P in its spectral decomposition $P = \sum_i \chi_i |p_i\rangle\langle p_i|$, we then have for any $|a\rangle$, $|b\rangle$ that

$$\begin{aligned} |\langle a|P|b\rangle|^2 &= \left|\sum_i \chi_i \langle a|p_i\rangle \langle p_i|b\rangle\right|^2 \\ &\leq \left(\sum_i \chi_i |\langle a|p_i\rangle|^2\right)\left(\sum_i \chi_i |\langle b|p_i\rangle|^2\right) \\ &\leq \left(\sum_i \chi_i |\langle a^\star|p_i\rangle|^2\right)\left(\sum_i \chi_i |\langle a^\star|p_i\rangle|^2\right) \\ &= \langle a^\star|P|a^\star\rangle^2 \end{aligned} \tag{5.25}$$

using the Cauchy-Schwarz inequality, which proves the Lemma. ∎

Proof of Proposition 5.5. By the above Lemma, one immediately obtains $\Gamma^\circ_{\mathcal{S}_+}(P) = \Gamma^\circ_{\mathcal{S}}(P)$ from the definitions. It is also straightforward to see that $\Gamma^\circ_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(P)$ follows in a similar way. For $\Gamma^\circ_{\mathcal{S}_+^\cup}$, we have that

$$\begin{aligned} \Gamma^\circ_{\mathcal{S}_+^\cup}(P) &= \sup\left\{\langle P, Z\rangle \;\middle|\; Z \in \operatorname{conv}(\mathcal{S}_+^\cup)\right\} \\ &= \sup\left\{\left\langle P, \sum_i p_i |v_i\rangle\langle v_i|\right\rangle \;\middle|\; |v_i\rangle \in \operatorname{conv}(\mathcal{V}),\ p_i \in \mathbb{R}_+,\ \sum_i p_i = 1\right\} \\ &= \sup\left\{\sum_i p_i \langle v_i|P|v_i\rangle \;\middle|\; |v_i\rangle \in \operatorname{conv}(\mathcal{V}),\ p_i \in \mathbb{R}_+,\ \sum_i p_i = 1\right\} \\ &= \sup\left\{\sum_i p_i \langle v_i|P|v_i\rangle \;\middle|\; |v_i\rangle \in \mathcal{V},\ p_i \in \mathbb{R}_+,\ \sum_i p_i = 1\right\} \\ &= \sup\left\{\langle P, Z\rangle \;\middle|\; Z \in \mathcal{S}_+\right\} \\ &= \Gamma^\circ_{\mathcal{S}_+}(P) \end{aligned} \tag{5.26}$$

where the fourth equality follows since each $\langle v_i|P|v_i\rangle$ is a convex quadratic form in $|v_i\rangle$ by the positive semidefiniteness of P [71], and the maximum of a convex function over a bounded convex set is reached at an extremal point of the set [63, 32.2]. ∎

[71] Boyd & Vandenberghe, 2004

[63] Rockafellar, 1970

5.2 BASE NORMS AND ROBUSTNESS

[84] Hartkämper & Neumann, 1974

If the set $\mathcal{S}_+$ spans the whole space of Hermitian matrices, the gauge $\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}$ defines a valid norm for $\mathbb{H}$ called the *base norm* [84]. Base norms have found a variety of uses in state and channel discrimination [88–91]. More generally, one can consider other gauges of the type $\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}$ with $\mathcal{X}$ being another set satisfying some desired properties. A connection between such gauges and measures commonly used in entanglement theory — robustness and negativity —

[88] Matthews et al., CMP 2009

[89] Reeb et al., JMP 2011

[90] Jenčová, JMP 2014

[91] Lami et al., CMP 2018

was noticed by Vidal and Werner [92] as well as Rudolph [93] and later briefly expanded on by Plenio and Virmani [81]. We extend and further generalise the relation between gauges of this type and resource quantifiers.

[92] Vidal & Werner, PRA 2002
[93] Rudolph, QIP 2005
[81] Plenio & Virmani, QIC 2007

Assuming that $\mathcal{X}$ is a compact set of Hermitian operators, we know that in general we have

$$\begin{aligned}\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}(\rho) &= (\Gamma_{\mathcal{S}_+} \mathbin{\dot\vee} \Gamma_{-\mathcal{X}})(\rho) \\ &= \inf_{\rho=A-B} \Gamma_{\mathcal{S}_+}(A) + \Gamma_{-\mathcal{X}}(-B) \\ &= \inf \Gamma_{\mathcal{S}_+}(\rho+B) + \Gamma_{\mathcal{X}}(B)\end{aligned} \tag{5.27}$$

where the optimisation can be taken to be over arbitrary Hermitian B, but effectively we are constrained to $B \in \mathcal{X}^{**}$ since $\Gamma_{\mathcal{X}}(B) = \infty$ for any $B \notin \mathcal{X}^{**}$. This problem admits a simplified form when the set $\mathcal{X}$ consists of trace-one matrices only:

$$\begin{aligned}\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}(\rho) &= \inf_{B\in\mathcal{X}^{**}} \Gamma_{\mathcal{S}_+}(\rho+B) + \operatorname{Tr} B \\ &= \inf\left\{ \operatorname{Tr}(\rho+2B) \;\middle|\; B\in\mathcal{X}^{**},\ \rho+B\in\mathcal{S}_+^{**}\right\}\end{aligned} \tag{5.28}$$

where we have used that $\Gamma_{\mathcal{X}}(B) = \operatorname{Tr} B$ for any $B \in \mathcal{X}^{**}$, and similarly for $\mathcal{S}_+^{**}$. By defining a general quantity which we will refer to as the *robustness* with respect to $\mathcal{X}$ as

$$R^{\mathcal{X}}_{\mathcal{S}_+}(\rho) \coloneqq \inf_{Z\in\operatorname{conv}\mathcal{X}} \left\{ \lambda\in\mathbb{R}_+ \;\middle|\; \rho+\lambda Z\in\mathcal{S}_+^{**}\right\}, \tag{5.29}$$

we can write $\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}(\rho) = 1 + 2R^{\mathcal{X}}_{\mathcal{S}_+}(\rho)$ as long as $\mathcal{X}\subseteq\mathbb{H}_1$. We remark that both the robustness and the gauge $\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}$ can, of course, be defined for more general sets $\mathcal{X}\notin\mathbb{H}_1$, and we will later encounter also such cases.

For the remainder of this section, we will assume that $\mathcal{X}\subseteq\mathbb{D}$, which gives $R^{\mathcal{X}}_{\mathcal{S}_+}$ an intuitive physical interpretation as the robustness of ρ to noise in $\mathcal{X}$ [94], that is, the smallest amount of mixing with a state in the set $\operatorname{conv}(\mathcal{X})$ necessary in order for the resulting mixture to be in $\operatorname{conv}(\mathcal{S}_+)$ (see Fig. 5.2). Specifically, we have

[94] Vidal & Tarrach, PRA 1999

$$R^{\mathcal{X}}_{\mathcal{S}_+}(\rho) = \inf_{\omega\in\operatorname{conv}\mathcal{X}} \left\{ \lambda\in\mathbb{R}_+ \;\middle|\; \frac{\rho+\lambda\omega}{1+\lambda}\in\operatorname{conv}(\mathcal{S}_+)\right\}. \tag{5.30}$$

The robustness itself is also a gauge function, albeit corresponding to an unbounded set: a straightforward manipulation gives

$$\begin{aligned}R^{\mathcal{X}}_{\mathcal{S}_+}(\rho) &= \Gamma_{\mathcal{S}_+^{**}\cup(-\mathcal{X})}(\rho) \\ &= \Gamma_{\mathcal{S}_+\cup(-\mathcal{X}^{**})}(\rho) - 1.\end{aligned} \tag{5.31}$$

Alternatively, by introducing the generalised inequality $A\succeq_{\mathcal{X}} B \iff A-B\in\mathcal{X}^{**}$, we can also write the above as

$$\begin{aligned}R^{\mathcal{X}}_{\mathcal{S}_+}(\rho) &= \inf_{\omega\in\operatorname{conv}\mathcal{X}} \left\{ \lambda\in\mathbb{R}_+ \;\middle|\; \rho\succeq_{\mathcal{S}_+} -\lambda\,\omega\right\} \\ &= \inf_{\sigma\in\operatorname{conv}\mathcal{S}_+} \left\{ \lambda\in\mathbb{R}_+ \;\middle|\; \rho\preceq_{\mathcal{X}} (1+\lambda)\,\sigma\right\}.\end{aligned} \tag{5.32}$$

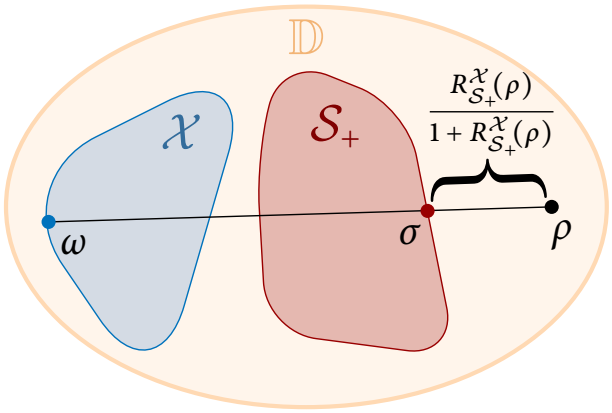


Figure 5.2: Robustness of a resource. The quantity $R^{\mathcal{X}}_{\mathcal{S}_+}(\rho)$ corresponds to the least weight of a state in $\mathrm{conv}(\mathcal{X}) \subseteq \mathbb{D}$ in a convex combination with ρ, such that the convex combination lies in $\mathrm{conv}(\mathcal{S}_+)$.

We note the similarity of this expression to geometric concepts such as the Thompson metric [95,96] and Hilbert's metric [97,98], the latter having been applied to the study of distinguishability norms in quantum information [89].

Two choices of the set $\mathcal{X}$ will be particularly useful: following the naming conventions of the theory of quantum entanglement, we introduce the *(standard) robustness* $R^{\mathcal{S}_+}_{\mathcal{S}_+}$ and the *generalised robustness* $R^{\mathbb{D}}_{\mathcal{S}_+}$ [94]. The latter is, up to logarithm, equivalent to the so-called max-relative entropy [99]. We recall that we generally the subscript in $\succeq_{\mathbb{D}}$ and simply use $\succeq$ to refer to the inequality with respect to the positive semidefinite cone (Löwner partial order).

By the above consideration, the standard robustness is an affine function of the base gauge $\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}$, and the generalised robustness of $\Gamma_{\mathcal{S}_+\cup(-\mathbb{D})}$. We recall that the effective domain of $\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}$ is $\mathrm{span}(\mathcal{S}_+)$, and so again we have $\mathbb{D} \subseteq \mathrm{dom}(R^{\mathcal{S}_+}_{\mathcal{S}_+})$ only if $\mathrm{relint}_{\mathbb{D}}(\mathrm{conv}\,\mathcal{S}_+) \neq \varnothing$. The domain of the generalised robustness $R^{\mathbb{D}}_{\mathcal{S}_+}$ consists of all states supported on the subspace spanned by $\mathcal{V}$, i.e. $\mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+}) = \mathcal{S}_+^{\cup\star\star}$, and so $\mathrm{span}(\mathcal{V}) = \mathbb{C}^d$ is necessary and sufficient for $R^{\mathbb{D}}_{\mathcal{S}_+}$ to be finite for all states.

We now consider the dual formulations of the gauges. Noting that $(\mathcal{S}_+ \cup -\mathbb{D})^\circ = \mathcal{S}_+^\circ \cap (-\mathbb{D})^\circ$ and similarly for $\mathcal{S}_+ \cup (-\mathcal{S}_+)$, we have

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) = \frac{1}{2}\sup\left\{\langle \rho, W'\rangle \;\middle|\; W' \in \mathcal{S}_+^\circ \cap (-\mathbb{D})^\circ\right\} - \frac{1}{2} \tag{5.33}$$

$$R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) = \frac{1}{2}\sup\left\{\langle \rho, W'\rangle \;\middle|\; W' \in \mathcal{S}_+^\circ \cap (-\mathcal{S}_+)^\circ\right\} - \frac{1}{2}. \tag{5.34}$$

With the change of variables $W' = \mathbb{1} - 2W$, we obtain a common representation

[95] Thompson, Proc. Amer. Math. Soc. 1963
[96] Cobzaș & Rus, 2014
[97] Hilbert, Math. Ann. 1895
[98] Bushell, Arch. Rat. Mech. Anal. 1973
[89] Reeb et al., JMP 2011
[94] Vidal & Tarrach, PRA 1999
[99] Datta, IEEE TIT 2009

of the Lagrange duals of the robustnesses [100]: [100] Brandão, PRA 2005

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) = \sup\left\{ -\langle \rho, W\rangle \;\middle|\; W \in \mathcal{S}_+^\star \cap \mathbb{D}^\circ \right\} \tag{5.35}$$

$$R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) = \sup\left\{ -\langle \rho, W\rangle \;\middle|\; W \in \mathcal{S}_+^\star \cap \mathcal{S}_+^\circ \right\}. \tag{5.36}$$

It follows that strong Lagrange duality always holds for the robustness quantifiers by virtue of their being gauge functions. This will be considered in more generality in Sec. 5.3.2.

5.2.1 *Relations between robustness and other gauges*

We proceed by characterising various relations between the robustness measures and the previously introduced atomic gauges, and in particular showing that the generalised robustness $R^{\mathbb{D}}_{\mathcal{S}_+}$ always reduces to the vector atomic gauge $\Gamma_{\mathcal{V}}$ on pure states.

Theorem 5.7. *For any state $\rho \in \mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$, we have*

$$\Gamma_{\mathcal{S}}(\rho) - 1 \geq R^{\mathbb{D}}_{\mathcal{S}_+}(\rho). \tag{5.37}$$

Proof. First, note that the generalised robustness can be expressed as

$$\begin{aligned}
2R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) + 1 &= \Gamma_{\mathcal{S}_+ \cup (-\mathbb{D})}(\rho) \\
&= \sup\left\{ \langle \rho, W\rangle \;\middle|\; W \in \mathcal{S}_+^\circ,\ \langle W, \sigma\rangle \geq -1\ \forall \sigma \in \mathbb{D} \right\} \\
&= 2\sup\left\{ \langle \rho, W'\rangle \;\middle|\; W' \in \mathcal{S}_+^\circ,\ W' \in \mathbb{D}^* \right\} - 1 \qquad (5.38a) \\
&= 2\sup\left\{ \langle \rho, W'\rangle \;\middle|\; W' \in \mathcal{S}_+^\circ,\ W' \in \mathbb{H}_+ \right\} - 1 \qquad (5.38b)
\end{aligned}$$

where in (5.38a) we make the change of variables $W' = \frac{1}{2}(W + \mathbb{1})$, and (5.38b) follows from the self-duality of the positive semidefinite cone $\mathbb{H}_+$ and the fact that $\mathbb{R}_+\mathbb{D} = \mathbb{H}_+$. We now have

$$\begin{aligned}
\Gamma_{\mathcal{S}}(\rho) &= \sup\left\{ \mathrm{Re}\,\langle \rho, Z\rangle \;\middle|\; Z \in \mathcal{S}^\circ \right\} \\
&\geq \sup\left\{ \mathrm{Re}\,\langle \rho, Z\rangle \;\middle|\; Z \in \mathbb{H}_+, Z \in \mathcal{S}^\circ \right\} \\
&= \sup\left\{ \langle \rho, Z\rangle \;\middle|\; Z \in \mathbb{H}_+,\ |\langle a|Z|b\rangle| \leq 1\ \forall\, |a\rangle, |b\rangle \in \mathcal{V} \right\} \\
&= \sup\left\{ \langle \rho, Z\rangle \;\middle|\; Z \in \mathbb{H}_+,\ \langle a|Z|a\rangle \leq 1\ \forall\, |a\rangle \in \mathcal{V} \right\} \\
&= \sup\left\{ \langle \rho, Z\rangle \;\middle|\; Z \in \mathbb{H}_+, Z \in \mathcal{S}_+^\circ \right\} \\
&= R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) + 1
\end{aligned}$$

where the third equality follows from Lemma 5.6. ∎

Theorem 5.8. *For any pure state $|\psi\rangle\langle\psi| \in \mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$, the generalised robustness is equivalent to the vector atomic gauge for the set $\mathcal{V}$:*

$$R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = \Gamma_{\mathcal{V}}(|\psi\rangle)^2 - 1 \tag{5.39}$$

Proof. The fact that we have $R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) \leq \Gamma_{\mathcal{V}}(|\psi\rangle)^2 - 1$ follows from Thm. 5.7, since by Thm. 5.3 we have $\Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|) = \Gamma_{\mathcal{V}}(|\psi\rangle)^2$ for any $|\psi\rangle \in \mathrm{span}(\mathcal{V})$.

To show that $R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) \geq \Gamma_V(|\psi\rangle\langle\psi|)^2 - 1$, from the dual characterisation of the robustness we have

$$\begin{aligned}
2R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) + 1 &= 2\sup\left\{\langle|\psi\rangle\langle\psi|, W'\rangle \;\middle|\; W' \in \mathcal{S}_+^\circ,\ W' \in \mathbb{H}_+\right\} - 1 \\
&\geq 2\sup\left\{\langle|\psi\rangle\langle\psi|, |w\rangle\langle w|\rangle \;\middle|\; |w\rangle\langle w| \in \mathcal{S}_+^\circ\right\} - 1 \\
&= 2\sup\left\{|\langle\psi|w\rangle|^2 \;\middle|\; \langle|w\rangle\langle w|, |v\rangle\langle v|\rangle \leq 1\ \forall\, |v\rangle \in \mathcal{V}\right\} - 1 \\
&= 2\sup\left\{|\langle\psi|w\rangle|^2 \;\middle|\; |w\rangle \in \mathcal{V}^\circ\right\} - 1 \\
&= 2\Gamma_{\mathcal{V}}(|\psi\rangle)^2 - 1.
\end{aligned} \tag{5.40}$$

■

Let us briefly discuss the implications of the above Theorem. It shows that the quantification of the robustness $R^{\mathbb{D}}_{\mathcal{S}_+}$ for pure states, which can be a very non-trivial problem to solve for specific resource theories explicitly [101–104], always reduces to an optimisation over the underlying Hilbert space only. This generalises results previously obtained in the resource theory of entanglement [101] and coherence [82], providing novel computable formulas and solving conjectures raised in the literature regarding the quantification of robustness measures of different resources [102,104,105], as we will discuss in detail in Chapter 6.

[101] Steiner, PRA 2003

[102] Clarisse, JPA 2006

[103] Cavalcanti et al., PRA 2005

[104] Ringbauer et al., preprint 2017

[105] Johnston & Kribs, Houst. J. Math. 2015

Remark 5.9. Notice that the proofs of Theorems 5.7 and 5.8 do not rely on the fact that $\mathcal{S}_+$ consists of normalised density matrices. We can obtain a more general result in the following form.

Corollary 5.10. *Let $\mathcal{Q} \coloneqq \mathrm{conv}\left\{|v\rangle\langle v| \;\middle|\; |v\rangle \in \mathcal{V}\right\}$ where $\mathcal{V} \subseteq \mathbb{C}^d$ is a balanced and compact set. For any $|x\rangle \in \mathbb{C}^d$, it then holds that*

$$\sup\left\{\langle x|W|x\rangle \;\middle|\; W \in \mathbb{H}_+ \cap \mathcal{Q}^\circ\right\} = \sup\left\{|\langle x|w\rangle|^2 \;\middle|\; |w\rangle \in \mathcal{V}^\circ\right\}. \tag{5.41}$$

We can furthermore relate the standard robustness $R^{\mathcal{S}_+}_{\mathcal{S}_+}$ with the other considered gauge functions.

Theorem 5.11. *For any state $\rho \in \mathrm{span}(\mathcal{S}_+)$ we have*

$$R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) - 1 \tag{5.42}$$

Proof. First, recall that by Theorems 5.3 and 5.8 we have

$$R^{\mathcal{S}_+}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) \geq R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = \Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|) - 1 \tag{5.43}$$

for any $|\psi\rangle\langle\psi| \in \mathrm{dom}(R^{\mathcal{S}_+}_{\mathcal{S}_+})$. Together with the fact that $2R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho)+1 = \Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(\rho)$, it means that we have $\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(|\psi\rangle\langle\psi|) \geq 2\Gamma_{\mathcal{S}}(|\psi\rangle\langle\psi|) - 1$. We then use the representation of $\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}$ in Eq. (5.6) to write

$$\begin{aligned}
\Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}(\rho) &= \inf\left\{ \sum_i |c_i|\, \Gamma_{\mathcal{S}_+\cup(-\mathcal{S}_+)}\left(|a_i\rangle\langle a_i|\right) \;\middle|\; \rho = \sum_i c_i\, |a_i\rangle\langle a_i|\,,\; c_i \in \mathbb{R} \right\} \\
&\geq 2\inf\left\{ \sum_i |c_i|\, \Gamma_{\mathcal{S}}\left(|a_i\rangle\langle a_i|\right) \;\middle|\; \rho = \sum_i c_i\, |a_i\rangle\langle a_i|\,,\; c_i \in \mathbb{R} \right\} - 1 \\
&\geq 2\inf\left\{ \sum_i |c_i|\, \Gamma_{\mathcal{S}}\left(|a_i\rangle\langle b_i|\right) \;\middle|\; \rho = \sum_i c_i\, |a_i\rangle\langle b_i|\,,\; c_i \in \mathbb{R} \right\} - 1 \\
&= 2\Gamma_{\mathcal{S}}(\rho) - 1.
\end{aligned} \tag{5.44}$$

This gives $R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) - 1$ as required. ∎

Combining Thm. 5.7 and 5.11, we have established a hierarchy of measures valid for any resource theory in the considered framework:

$$R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) - 1 \geq R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \tag{5.45}$$

where the second inequality is tight for all pure states.

It is known that in the resource theory of entanglement, not only the generalised robustness, but also the standard robustness of entanglement reduces to the vector atomic gauge $\Gamma_{\mathcal{V}}$ for pure states, i.e. $R^{\mathcal{S}_+}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = \Gamma_{\mathcal{V}}(|\psi\rangle)^2 - 1$ [94]. Very recently, this has been generalised also to resource theories of k-coherence and entanglement of Schmidt number k [BR11]. However, one can find that this does not hold for all quantum resources (see Sec. 6.4 for an explicit example in the resource theory of magic states) — it is then an interesting open question to characterise exactly when $R^{\mathcal{S}_+}_{\mathcal{S}_+}$ is equivalent to $\Gamma_{\mathcal{V}}$ on pure states. We remark that in any resource theory where $R^{\mathcal{S}_+}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = \Gamma_{\mathcal{V}}(|\psi\rangle)^2 - 1$ does hold, we have an extended hierarchy of quantifiers, all of which are equal on pure states:

[94] Vidal & Tarrach, PRA 1999

[BR11] Johnston et al., preprint 2018

$$\Gamma^{\cup}_{\mathcal{S}_+}(\rho) - 1 \geq R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) \geq \Gamma_{\mathcal{S}}(\rho) - 1 \geq R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \tag{5.46}$$

which follows since $\Gamma^{\cup}_{\mathcal{S}_+}$ is the largest convex function equal to $\Gamma^2_{\mathcal{V}}$ on pure states.

Proposition 5.12. *For any $P \in \mathbb{H}_+$ it holds that*

$$\Gamma^{\circ}_{\mathcal{S}_+\cup(-\mathbb{D})}(P) = \Gamma^{\circ}_{\mathcal{S}_+}(P). \tag{5.47}$$

Proof. By the positive semidefiniteness of P we have

$$\begin{aligned}
\Gamma^{\circ}_{\mathcal{S}_+\cup(-\mathbb{D})}(P) &= \sup\left\{ \langle P, Z\rangle \;\middle|\; Z \in \mathrm{conv}\left(\mathcal{S}_+ \cup (-\mathbb{D})\right) \right\} \\
&= \max\left\{0,\; \sup\left\{ \langle P, Z\rangle \;\middle|\; Z \in \mathcal{S}_+ \right\}\right\} \\
&= \Gamma^{\circ}_{\mathcal{S}_+}(P).
\end{aligned} \tag{5.48}$$

∎

One can use the result of Prop. 5.12 to give a bound on the value of $R^{\mathbb{D}}_{\mathcal{S}_+}$ using the Cauchy-Schwarz inequality for $\Gamma_{\mathcal{S}_+\cup(-\mathbb{D})}$, namely

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \geq \frac{1}{2}\left(\frac{\mathrm{Tr}\left(\rho^2\right)}{\Gamma^{\circ}_{\mathcal{S}_+}(\rho)} - 1\right). \tag{5.49}$$

However, a tighter bound can be obtained.

Proposition 5.13. *For any state $\rho \in \mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$ it holds that*

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \geq \frac{\mathrm{Tr}\left(\rho^2\right)}{\Gamma^{\circ}_{\mathcal{S}_+}(\rho)} - 1. \tag{5.50}$$

Proof. First, notice that we do not need to assume $\Gamma^{\circ}_{\mathcal{S}_+}(\rho) \neq 0$, since any positive semidefinite $\rho \in \mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$ satisfies $\mathrm{supp}(\rho) \subseteq \mathrm{span}(\mathcal{V})$ and so we cannot have that $\langle \rho, \sigma \rangle = 0 \ \forall \sigma \in \mathcal{S}_+$. From $\Gamma^{\circ}_{\mathcal{S}_+} = \gamma_{\mathcal{S}^{\circ}_+}$ we then have that $\frac{\rho}{\Gamma^{\circ}_{\mathcal{S}_+}(\rho)} \in \mathcal{S}^{\circ}_+$. Since $\rho \in \mathbb{H}_+$, this is then a feasible solution to the dual formulation of the generalised robustness (Eq. 5.38) and so we have

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \geq \left\langle \rho, \frac{\rho}{\Gamma^{\circ}_{\mathcal{S}_+}(\rho)} \right\rangle - 1 \tag{5.51}$$

as required. ∎

5.3 RESOURCE MONOTONES

A full characterisation of a resource theory requires, in addition to the chosen set of free states without a given resource, the choice of a relevant set of *free operations* which cannot generate the resource. For simplicity, in the following we will limit ourselves to the discussion of maps acting as linear operators on $\mathbb{H}$; the results can be generalised to maps between different spaces by suitably defining the set of free states in the input space as well as the set of free states in the output space (see e.g. [6]). We will in general consider maps which are completely positive, but not necessarily trace preserving.

[6] Gour, PRA 2017

The definition of what exactly constitutes the free operations is frequently dependent on the resource theory or the physical setting under consideration. The largest possible set of free operations are the *resource non-generating maps* $\mathcal{O}_{\mathrm{RNG}}$, which are all completely positive maps Φ such that $\sigma \in \mathcal{S}^{\star\star}_+ \Rightarrow \Phi(\sigma) \in \mathcal{S}^{\star\star}_+$. In particular, a trace-preserving resource non-generating operation acting on a free state always results in another free state. Often, a smaller subset of operations is considered: they are the *resource-free maps* $\mathcal{O}_{\mathrm{RF}}$, consisting of maps Λ which admit a Kraus decomposition of the form $\Lambda(\rho) = \sum_n K_n \rho K_n^{\dagger}$ such that $\sigma \in \mathcal{S}^{\star\star}_+ \Rightarrow K_n \sigma K_n^{\dagger} \in \mathcal{S}^{\star\star}_+ \ \forall\, n$. The physical motivation for the choice of

resource-free operations is that it guarantees that no resource can be created from a non-resource state in any possible measurement outcome. In the theory of entanglement, these are the separable operations [106], while in the theory of coherence they are called incoherent operations [15]. Even these operations are often considered to not reflect the physical restrictions sufficiently well, and much smaller subsets of physically relevant operations are used; as discussed before, in entanglement theory the set of operations of interest are the local operations and classical communication [52,107], and in coherence theory a common choice are the strictly incoherent operations [55] or even smaller subsets [56–58].

[106] Bennett et al., PRA 1999

[15] Baumgratz et al., PRL 2014

[52] Vedral et al., PRL 1997

[107] Chitambar et al., CMP 2014

[55] Winter & Yang, PRL 2016

[56] Chitambar & Gour, PRL 2016

[57] Marvian & Spekkens, PRA 2016

[58] de Vicente & Streltsov, JPA 2017

Identifying $\mathrm{conv}(\mathcal{S}_+)$ with the set of free states and $\mathcal{F}$ with a chosen set of free operations, we define a valid measure for the given resource to be a function $M : \mathbb{D} \to \mathbb{R}_+ \cup \{\infty\}$ which satisfies two basic criteria:

(C1) **Faithfulness**: $M(\rho) = 0$ if and only if $\rho \in \mathrm{conv}(\mathcal{S}_+)$.

(C2) **Monotonicity**: $M(\Theta(\rho)) \leq M(\rho)$ for all completely positive trace-preserving operations $\Theta \in \mathcal{F}$.

Often, additional conditions are imposed, the two most common ones being:

(C3) **Convexity**: $M\left(\sum_i t_i \rho_i\right) \leq \sum_i t_i M(\rho_i)$ for $t_i \in \mathbb{R}_+$, $\rho_i \in \mathbb{D}$, $\sum_i t_i = 1$.

(C4) **Strong monotonicity**: $\sum_i p_i M\left(\frac{\Theta_i(\rho)}{p_i}\right) \leq M(\rho)$ with $p_i = \mathrm{Tr}\,\Theta_i(\rho)$, where $\Theta_i \in \mathcal{F}$ are completely positive operations such that $\sum_i \Theta_i$ is trace preserving.

The choice of the free operations of interest $\mathcal{F}$ will, in general, not be unique — however, we stress that (strong) monotonicity under a given class of operations implies (strong) monotonicity under any subset of this class, and so it suffices to investigate monotonicity under larger sets of operations. Further, note that strong monotonicity implies monotonicity.

5.3.1 *Properties of atomic gauges*

We now verify the criteria for the candidate measures: $\Gamma_{\mathcal{S}}(\rho) - 1$, $\Gamma^{\cup}_{\mathcal{S}_+}(\rho) - 1$, $R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho)$, and $R^{\mathbb{D}}_{\mathcal{S}_+}(\rho)$.

Theorem 5.14. *A state $\rho \in \mathbb{D}$ belongs to the set $\mathrm{conv}(\mathcal{S}_+)$ if and only if*

$$\begin{aligned} 0 &= \Gamma_{\mathcal{S}_+}(\rho) - 1 = \Gamma_{\mathcal{S}}(\rho) - 1 = \Gamma_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}(\rho) - 1 = \Gamma^{\cup}_{\mathcal{S}_+}(\rho) - 1 \\ &= R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho) = R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \end{aligned} \tag{5.52}$$

That is, all of the considered measures are faithful.

Proof. It is helpful to recall that the gauge function of the set

$$\mathcal{U} = \left\{ |x\rangle\langle y| \;\middle|\; \||x\rangle\|_{\ell_2} \leq 1,\ \||y\rangle\|_{\ell_2} \leq 1 \right\} \tag{5.53}$$

is just the trace norm $\|\cdot\|_1$. Since $\mathcal{S}_+ \subseteq \mathcal{S}_+ \cup (-\mathcal{S}_+) \subseteq \mathcal{S} \subseteq \mathcal{U}$, $\mathcal{S}_+ \subseteq \mathcal{S}_+^{\cup} \subseteq \mathcal{U}$, and $\mathcal{S}_+ \subseteq \mathcal{S}_+ \cup (-\mathbb{D}) \subseteq \mathcal{U}$ it follows that $\Gamma_{\mathcal{U}}$ lower bounds all of the considered gauges, and so each of the gauges cannot be smaller than 1 when $\rho \in \mathbb{D}$ because $\|\rho\|_1 = 1$.

Then if $\rho \in \operatorname{conv}(\mathcal{S}_+)$, we have $\Gamma_{\mathcal{S}_+}(\rho) = 1$ by definition, and all the other gauges also have to be equal to 1 because they are upper bounded by $\Gamma_{\mathcal{S}_+}$ (Eq. (5.15)), which in turn means the robustnesses will be zero.

Conversely, assume that $R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) = 0$, which means that $\Gamma_{\mathcal{S}_+ \cup (-\mathbb{D})}(\rho) = 1$. This gives that $\rho \in \operatorname{conv}(\mathcal{S}_+ \cup (-\mathbb{D}))$, but since $\rho \in \mathbb{D}$ by assumption and $\operatorname{conv}(\mathcal{S}_+ \cup (-\mathbb{D})) \cap \mathbb{D} = \operatorname{conv}(\mathcal{S}_+)$, we necessarily have $\rho \in \operatorname{conv}(\mathcal{S}_+)$. Noting by Thm. 5.7 and the inequalities between the gauges that $R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) + 1$ lower bounds all other gauges in consideration, it follows that any of the gauges being equal to 1 implies that $R^{\mathbb{D}}_{\mathcal{S}_+}$ is zero, proving the claim. ∎

Proposition 5.15. *All of the considered measures are convex.*

Proof. Every gauge function of a convex set is convex. ∎

Lemma 5.16. *Let $\mathcal{C}$ denote either $\operatorname{conv}(\mathcal{S}_+ \cup (-\mathcal{S}_+))$ or $\operatorname{conv}(\mathcal{S}_+ \cup (-\mathbb{D}))$. Then, given a resource non-generating completely positive map $\Phi \in \mathcal{O}_{RNG}$ and an operator X, we have $X \in \mathcal{C} \Rightarrow \Phi(X) \in \mathbb{R}_+\mathcal{C}$.*

Proof. For $\mathcal{C} = \operatorname{conv}(\mathcal{S}_+ \cup (-\mathcal{S}_+))$, write $X = s\sigma^+ - (1-s)\sigma^-$ with $\sigma^\pm \in \operatorname{conv}(\mathcal{S}_+)$, $s \in [0,1]$. Then $\Phi(X) = s\Phi(\sigma^+) - (1-s)\Phi(\sigma^-) = s \operatorname{Tr}(\Phi(\sigma^+))\,\varsigma^+ + (1-s)\operatorname{Tr}(\Phi(\sigma^-))\,\varsigma^-$ for some $\varsigma^\pm \in \operatorname{conv}(\mathcal{S}_+)$ by linearity and the resource non-generating property of Φ. We then have $\Phi(X) = \operatorname{Tr}(\Phi(X))\,(s\varsigma^+ - (1-s)\varsigma^-)$ as required.

For $\mathcal{C} = \operatorname{conv}(\mathcal{S}_+ \cup (-\mathbb{D}))$, write $X = s\sigma - (1-s)\pi$ with $\sigma \in \operatorname{conv}(\mathcal{S}_+)$, $\pi \in \mathbb{D}$, $s \in [0,1]$. Then $\Phi(X) = s\Phi(\sigma) - (1-s)\Phi(\pi) = s \operatorname{Tr}(\Phi(\sigma))\,\varsigma - (1-s)\operatorname{Tr}(\Phi(\pi))\,\omega$ with $\varsigma \in \operatorname{conv}(\mathcal{S}_+)$ by linearity and the resource non-generating property of Φ, and $\omega \in \mathbb{D}$ as Φ is completely positive. This gives $\Phi(X) = \operatorname{Tr}(\Phi(X))\,(s\varsigma - (1-s)\omega)$ as required. ∎

Lemma 5.17. *Let $\mathcal{C}$ denote either $\operatorname{conv}(\mathcal{S})$ or $\operatorname{conv}(\mathcal{S}_+^{\cup})$. Then, given a resource-free completely positive map $\Lambda \in \mathcal{O}_{RF}$ and an operator X, we have $X \in \mathcal{C} \Rightarrow \Lambda(X) \in \mathbb{R}_+\mathcal{C}$.*

Proof. Let us consider $\mathcal{C} = \operatorname{conv}(\mathcal{S})$ first. Write any $\Lambda \in \mathcal{O}_{\mathrm{RF}}$ in its Kraus decomposition as $\Lambda(\rho) = \sum_n K_n \rho K_n^\dagger$. By definition of the resource-free operations,

we can note that for every pure state $|v_i\rangle \in \mathcal{V}$, we have $K_n |v_i\rangle = |v_i^n\rangle$ for some $|v_i^n\rangle \in \mathbb{R}_+\mathcal{V}$ as the complex phase can be absorbed into the state $|v_i^n\rangle$. By definition, any operator $X \in \mathrm{conv}(\mathcal{S})$ can be written as $X = \sum_i c_i |v_i\rangle\langle u_i|$ for some coefficients $c_i \in \mathbb{R}_+$ with $\sum_i c_i = 1$ and states $|v_i\rangle \in \mathcal{V}$, $|u_i\rangle \in \mathcal{V}$, which then gives

$$\begin{aligned}\Lambda(X) &= \sum_n K_n \left(\sum_i c_i |v_i\rangle\langle u_i| \right) K_n^\dagger \\ &= \sum_{i,n} c_i K_n |v_i\rangle\langle u_i| K_n^\dagger \\ &= \sum_{i,n} c_i |v_i^n\rangle\langle u_i^n|\end{aligned} \tag{5.54}$$

which is explicitly in $\mathcal{S}^{**}$.

Take now $\mathrm{conv}(\mathcal{S}_+^\cup)$. Noting that any state in $\mathrm{conv}(\mathcal{S}_+^\cup)$ can be written as $\rho = \sum_i c_i |w_i\rangle\langle w_i|$ with $|w_i\rangle \in \mathrm{conv}(\mathcal{V})$, $c_i \in \mathbb{R}_+$, and $\sum_i c_i = 1$, and similarly each $|w_i\rangle$ can be written as $|w_i\rangle = \sum_j d_{ij} |v_{ij}\rangle$ with $|v_{ij}\rangle \in \mathcal{V}$, $d_{ij} \in \mathbb{R}_+$, $\sum_j d_{ij} = 1$. This then gives

$$\begin{aligned}\Lambda(\rho) &= \sum_n K_n \left(\sum_{i,j,k} c_i d_{ij} d_{ik} |v_{ij}\rangle\langle v_{ik}| \right) K_n^\dagger \\ &= \sum_{i,j,k,n} c_i d_{ij} d_{ik} |v_{ij}^n\rangle\langle v_{ik}^n| \\ &= \sum_i c_i |q_i\rangle\langle q_i|\end{aligned} \tag{5.55}$$

where

$$|q_i\rangle = \sum_{j,n} d_{ij} |v_{ij}^n\rangle\,. \tag{5.56}$$

Since each $|q_i\rangle \in \mathcal{V}^{**}$, we have that $\Lambda(\rho) \in \mathcal{S}_+^{\cup **}$ as required. ■

Theorem 5.18. *The robustness measures $R^{\mathbb{D}}_{\mathcal{S}_+}$ and $R^{\mathcal{S}_+}_{\mathcal{S}_+}$ satisfy strong monotonicity under resource non-generating maps $\mathcal{O}_{RNG}$.*

Theorem 5.19. *The atomic gauges $\Gamma_{\mathcal{S}}$ and $\Gamma^{\cup}_{\mathcal{S}_+}$ satisfy strong monotonicity under resource-free maps $\mathcal{O}_{RF}$.*

Proof. Let $\mathcal{C}$ denote one of the sets: $(\mathcal{S}_+ \cup (-\mathcal{S}_+))$, $(\mathcal{S}_+ \cup (-\mathbb{D}))$, $\mathcal{S}$, or $\mathcal{S}_+^\cup$. Depending on the choice of the set, let $\mathcal{F}$ denote the relevant set of free operations — $\mathcal{O}_{\mathrm{RNG}}$ or $\mathcal{O}_{\mathrm{RF}}$. Now, let $\Theta = \sum_i \Phi_i \in \mathcal{F}$ denote a free CPTP map, where each $\Phi_i \in \mathcal{F}$ is a trace non-increasing, completely positive map.

By definition of the atomic gauge $\Gamma_{\mathcal{C}}$, for any $\rho \in \mathrm{dom}(\Gamma_{\mathcal{C}}) \cap \mathbb{D}$ we have that

$$\frac{\rho}{\Gamma_{\mathcal{C}}(\rho)} = X \in \mathrm{conv}(\mathcal{C}). \tag{5.57}$$

Then

$$\begin{aligned}\Phi_i(\rho) &= \Phi_i\left(\Gamma_{\mathcal{C}}(\rho)\, X\right) \\ &= \Gamma_{\mathcal{C}}(\rho)\, \Phi_i(X) \\ &= \Gamma_{\mathcal{C}}(\rho)\, Y_i\end{aligned} \tag{5.58}$$

for some $Y_i \in \operatorname{Tr} Y_i \operatorname{conv}(\mathcal{C})$ by Lemma 5.16 or Lemma 5.17. Now let $p_i = \operatorname{Tr}\left(\Phi_i(\rho)\right)$. Dividing Eq. (5.58) by p_i, we have that

$$\frac{\Phi_i(\rho)}{p_i} \in \frac{\Gamma_{\mathcal{C}}(\rho) \operatorname{Tr} Y_i}{p_i} \operatorname{conv}(\mathcal{C}) \tag{5.59}$$

and hence

$$\Gamma_{\mathcal{C}}\left(\frac{\Phi_i(\rho)}{p_i}\right) \leq \frac{\Gamma_{\mathcal{C}}(\rho) \operatorname{Tr} Y_i}{p_i}. \tag{5.60}$$

Notice now that, since $\Gamma_{\mathcal{C}}(\rho) \geq 1$ and $\sum_i \Phi_i$ is trace preserving, we have $1 = \operatorname{Tr}\rho \geq \operatorname{Tr} X = \operatorname{Tr}\sum_i Y_i$, which then gives

$$\sum_i p_i\, \Gamma_{\mathcal{C}}\left(\frac{\Phi_i(\rho)}{p_i}\right) \leq \Gamma_{\mathcal{C}}(\rho) \sum_i \operatorname{Tr} Y_i \leq \Gamma_{\mathcal{C}}(\rho) \tag{5.61}$$

and so strong monotonicity is satisfied. ■

Remark 5.20. The gauges $\Gamma_{\mathcal{S}}$ and $\Gamma^{\cup}_{\mathcal{S}_+}$ are, in general, not strongly monotonic under $\mathcal{O}_{\text{RNG}}$. An explicit counterexample has been given in the resource theory of coherence [108].

[108] Bu & Xiong, QIC 2017

Although the choice of the four measures $\Gamma_{\mathcal{S}}$, $\Gamma^{\cup}_{\mathcal{S}_+}$, $R^{\mathbb{D}}_{\mathcal{S}_+}$ and $R^{\mathcal{S}_+}_{\mathcal{S}_+}$ is somewhat natural in the framework of the gauge functions, there is of course no reason to limit oneself to these particular measures and the two types of free operations that we have considered. One can establish a similar sufficient condition for other gauge-based measures, which we will in particular employ in the succeeding section to characterise other classes of resource quantifiers.

Theorem 5.21. *Let $\Gamma_{\mathcal{C}}$ be the atomic gauge function corresponding to a compact set $\mathcal{C}$. Let $\mathcal{F}$ denote a chosen set of completely positive maps. We then have that if the convex cone $\mathcal{C}^{**}$ is closed under the operations $\mathcal{F}$, i.e. $\sigma \in \mathcal{C}^{**} \Rightarrow \Theta(\sigma) \in \mathcal{C}^{**}\ \forall\, \Theta \in \mathcal{F}$, then $\Gamma_{\mathcal{C}}$ is strongly monotonic under $\mathcal{F}$.*

The proof is the same as above.

Remark 5.22. Instead of choosing the resource quantifier as $M = \Gamma_{\mathcal{C}} - 1$ for one of the gauge functions $\Gamma_{\mathcal{C}}$, one can instead consider $M = f \circ \Gamma_{\mathcal{C}}$ for any monotonically non-decreasing function f on the interval $[1, \infty)$ s.t. $f(1) = 0$. Any choice of a convex f then preserves the convexity of the measure, and any choice of a concave f preserves the strong monotonicity — a common choice of a concave function is $f = \log$.

5.3.2 Dual characterisation and resource witnesses

In quantum resource theories, elements of the dual cone $\mathcal{S}_+^\star$ are often called *witnesses* of the given resource. The notion of witnesses is a fundamental concept that found a variety of uses in the characterisation, detection, and quantification of quantum resources [33,100,109,110].

[33] Horodecki et al., PLA 1996
[100] Brandão, PRA 2005
[109] Terhal, Theor. Comput. Sci. 2002
[110] Gühne & Tóth, Phys. Rep. 2009

The set of witnesses can also be described in terms of the polar set $\mathcal{S}_+^\circ$ — the equivalence is made explicit by noting that a witness can be obtained from any $S \in \mathcal{S}_+^\circ$ simply by considering $\mathbb{1} - S \in \mathcal{S}_+^\star$, and vice versa. This gives an intuitive interpretation of the polar gauge in this sense:

Proposition 5.23. *For any $X \in \mathbb{H}$ it holds that*

$$\Gamma^\circ_{\mathcal{S}_+}(X) = \inf\left\{\lambda \in \mathbb{R}_+ \;\middle|\; \lambda\mathbb{1} - X \in \mathcal{S}_+^\star\right\}. \tag{5.62}$$

Proof. Recalling that $\Gamma^\circ_{\mathcal{S}_+} = \Gamma_{\mathcal{S}_+^\circ} = \gamma_{\mathcal{S}_+^\circ}$, we have $\Gamma^\circ_{\mathcal{S}_+}(X) = \inf\left\{\lambda \in \mathbb{R}_+ \;\middle|\; X \in \lambda\mathcal{S}_+^\circ\right\}$ and the result follows since $\langle \mathbb{1}, \sigma\rangle = 1\ \forall \sigma \in \mathcal{S}_+$. ∎

Remark 5.24. Notice that if X is such that $\langle \sigma, X\rangle \leq 0\ \forall \sigma \in \mathcal{S}_+$, then $\Gamma^\circ_{\mathcal{S}_+}(X) = 0$. In particular, a positive semidefinite matrix can have $\Gamma^\circ_{\mathcal{S}_+}(X) = 0$ if and only if $X \in \mathcal{S}_+^\perp$, where $\mathcal{S}_+^\perp = \mathcal{S}_+^\star \cap (-\mathcal{S}_+)^\star$ is the orthogonal complement of $\mathcal{S}_+$. Note that $\mathcal{S}_+^\perp \cap \mathbb{H}_+ = \{0\}$ when $\operatorname{span}(\mathcal{V}) = \mathbb{C}^d$.

The crucial property of witnesses applied to quantum resources is that, since $\operatorname{conv}(\mathcal{S}_+)$ is a convex and closed set, by the strongly separating hyperplane theorem, for every $\rho \notin \operatorname{conv}(\mathcal{S}_+)$ there exists a witness $W \in \mathcal{S}_+^\star$ such that $\langle \rho, W\rangle < 0$, thus *detecting* a given resource. On the other hand, if $\rho \in \operatorname{conv}(\mathcal{S}_+)$, then we necessarily have $\langle \rho, W\rangle \geq 0\ \forall W \in \mathcal{S}_+^\star$. This leads to a natural formalism of witness-based measures, quantifying how much a given state can violate the condition $\langle \rho, W\rangle \geq 0$: one constructs a general witness-based measure as [100]

$$V^{\mathcal{C}}_{\mathcal{S}_+}(\rho) = \sup\left\{-\langle \rho, W\rangle \;\middle|\; W \in \mathcal{S}_+^\star \cap \mathcal{C}\right\} \tag{5.63}$$

where $\mathcal{C}$ is a set representing some additional constraints on the set of witnesses under consideration. This approach is particularly useful in experimental settings, allowing for the detection and quantification of resources without the need for full state tomography [111,112].

[111] Gühne et al., PRL 2007
[112] Eisert et al., NJP 2007

Many of the gauge functions considered in this work can be written in this form with a suitable choice of $\mathcal{C}$ (up to a constant), which can be seen from their dual representation. For example, the standard robustness corresponds to $V^{\mathcal{C}}_{\mathcal{S}_+}$ with the choice $\mathcal{C} = \mathcal{S}_+^\circ$, and the generalised robustness to $\mathcal{C} = \mathbb{D}^\circ = \left\{X \;\middle|\; X \leq \mathbb{1}\right\}$. More generally, we can establish the following equivalence.

Proposition 5.25. *If $\mathcal{C}$ can be expressed as $\mathcal{C} = \mathcal{X}^\circ$ for some subset $\mathcal{X} \subseteq \mathbb{H}_1$,*

then

$$V^{\mathcal{C}}_{\mathcal{S}_+}(\rho) = \frac{1}{2}\left(\Gamma_{\mathcal{S}_+\cup(-\mathcal{X})}(\rho) - 1\right) \tag{5.64}$$

for any $\rho \in \mathbb{D}$. *Likewise, if* $\mathcal{C} = \mathcal{Y}^\circ$ *for some* $\mathcal{Y} \subseteq -\mathbb{H}_1$, *then*

$$V^{\mathcal{C}}_{\mathcal{S}_+}(\rho) = \Gamma_{\mathcal{S}_+\cup\left(-\frac{1}{2}\mathcal{Y}\right)}(\rho) - 1. \tag{5.65}$$

Proof. Take the standard expression for the witness-based measure in Eq. (5.63) and apply the change of variables $W' = \mathbb{1} - 2W$ in the case of $\mathcal{X}$, or $W' = \mathbb{1} - W$ in the case of $\mathcal{Y}$. The result then follows from the dual characterisation of the gauge functions. ∎

One can note a similarity of the above forms to the generalised base gauges considered in section 5.2.

An example of a witness-based measure is obtained by taking $V^{\mathcal{C}}_{\mathcal{S}_+}$ with $\mathcal{C} = \left\{X \;\middle|\; \operatorname{Tr} X \leq d\right\}$ [113]. Noting that this corresponds to $\mathcal{C} = \mathcal{X}^\circ$ with $\mathcal{X} = \left\{\frac{\mathbb{1}}{d}\right\}$, we obtain

[113] Brandão & Vianna, IJQI 2006

$$V^{\mathcal{C}}_{\mathcal{S}_+}(\rho) = R^{\{\mathbb{1}/d\}}_{\mathcal{S}_+}(\rho) = \inf\left\{s \in \mathbb{R}_+ \;\middle|\; \rho + s\frac{\mathbb{1}}{d} \in \mathcal{S}_+^{\star\star}\right\} \tag{5.66}$$

which defines the *random robustness* [94]. Random robustness has full domain iff $\frac{\mathbb{1}}{d} \in \operatorname{relint}_{\mathbb{D}}(\mathcal{S}_+)$, and it corresponds to the gauge function $\frac{1}{2}\left(\Gamma_{\mathcal{S}_+\cup\{-\mathbb{1}/d\}} - 1\right)$. Alternatively, one can think of it as the atomic gauge function of $\mathcal{S}_+$ under a reparametrisation such that $\frac{\mathbb{1}}{d}$ is the origin. Note that the random robustness is not a monotone under the general free operations $\mathcal{O}_{\text{RNG}}$, but by Theorem 5.21 it can be noted to be faithful and strongly monotonic under unital resource non-generating operations (i.e. $\Phi \in \mathcal{O}_{\text{RNG}}$ such that $\Phi(\mathbb{1}) = \mathbb{1}$). By the set inclusion of the corresponding unit balls we get $R^{\{\mathbb{1}/d\}}_{\mathcal{S}_+}(\rho) \geq R^{\mathcal{S}_+}_{\mathcal{S}_+}(\rho)$.

[94] Vidal & Tarrach, PRA 1999

Another common choice is to consider $V^{\mathcal{C}}_{\mathcal{S}_+}$ with $\mathcal{C} = (-\mathbb{D})^\circ = \left\{X \;\middle|\; X \succeq -\mathbb{1}\right\}$, which by strong Lagrange duality is equal to the so-called *best free approximation* BFA, generalising the best separable approximation [114]:

[114] Lewenstein & Sanpera, PRL 1998

$$\operatorname{BFA}_{\mathcal{S}_+}(\rho) = \min_{\sigma\in\operatorname{conv}\mathcal{S}_+}\left\{1 - \lambda \;\middle|\; \rho \succeq \lambda\sigma\right\}. \tag{5.67}$$

Here, the optimal λ is the largest weight that a free state can take in a convex decomposition of $\rho = \lambda\sigma + (1-\lambda)\omega$ where $\sigma \in \operatorname{conv}(\mathcal{S}_+)$ and $\omega \in \mathbb{D}$. This expression is then equivalent to the gauge function $\Gamma_{\mathcal{S}_+\cup\left(\frac{1}{2}\mathbb{D}\right)} - 1$. Since the cone generated by the set $\mathcal{S}_+ \cup \left(\frac{1}{2}\mathbb{D}\right)$ is closed under resource non-generating operations, we have that the best free approximation is a faithful strong monotone under $\mathcal{O}_{\text{RNG}}$ by Theorem 5.21. Note, however, that it has some unusual properties for a resource monotone; for example, any resourceful pure state $|\psi\rangle \notin \mathcal{V}$ has $\operatorname{BFA}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = 1$, which fails to account for the fact that some pure states can be strictly more resourceful than others.

5.4 RELATIONS WITH DISTANCE-BASED MEASURES

One can define a faithful quantifier of any resource simply by considering the distance to the set of free states. That is, we can define

$$D_{\mathcal{S}_+}(\rho) \coloneqq \inf_{\sigma \in \operatorname{conv}(\mathcal{S}_+)} D(\rho, \sigma) \tag{5.68}$$

for some quasi-metric D such that $D(\Lambda(\rho), \Lambda(\sigma)) \leq D(\rho, \sigma)$ for arbitrary CPTP maps Λ [52,53]. Any such measure is then a (weak) monotone under resource non-generating operations. Although the representation is appealing from an intuitive point of view, the measures defined in this way are frequently difficult to evaluate and investigate in practice, and few results about their properties such as strong monotonicity are known.

[52] Vedral et al., PRL 1997

[53] Bengtsson & Życzkowski, 2007

The distance-based quantifiers do not fit into the gauge function formalism directly, but we can nevertheless obtain relations between the gauge-based and distance-based measures in several cases — we will consider some representative examples of such quantifiers.

5.4.1 *Trace distance and other gauge-based distances*

A commonly encountered case is when D is itself based on a gauge function — this includes, for instance, the fundamental measure of trace distance obtained for $D(\rho, \sigma) = \|\rho - \sigma\|_1$. In general, one can consider

$$D_{\mathcal{S}_+}(\rho) = \inf_{\sigma \in \operatorname{conv} \mathcal{S}_+} \Gamma_{\mathcal{C}}(\rho - \sigma) \tag{5.69}$$

where we will take $\mathcal{C} \subseteq \mathbb{H}$ for simplicity, but the same considerations apply to more general sets of complex matrices with the inner product $\operatorname{Re}\langle \cdot, \cdot \rangle$. To see the difference between the distance-based measures and gauge functions, we can consider the dual form of the general quantifier $D_{\mathcal{S}_+}$:

$$\begin{aligned} D_{\mathcal{S}_+}(\rho) &= \inf_{\sigma \in \operatorname{conv}(\mathcal{S}_+)} \sup_{W \in \mathcal{C}^\circ} \langle W, \rho - \sigma \rangle \\ &= \sup \left\{ \langle W, \rho \rangle - \sup_{\sigma \in \mathcal{S}_+} \langle W, \sigma \rangle \;\middle|\; W \in \mathcal{C}^\circ \right\} \end{aligned} \tag{5.70}$$

which follows by Sion's minimax theorem (or can be equivalently derived by considering the Lagrange dual of the original problem explicitly). We emphasise that the function $\mu_{\mathcal{S}_+}(W) \coloneqq \sup_{\sigma \in \mathcal{S}_+} \langle W, \sigma \rangle$ is, in general, different from the polar gauge $\Gamma^\circ_{\mathcal{S}_+}(W) = [\mu_{\mathcal{S}_+}(W)]_+$. This representation can nevertheless be helpful since the polar gauges, and as a result the function $\mu_{\mathcal{S}_+}$, can be relatively easy to characterise (see e.g. Ch. 6).

It is explicit from the above representation that, although distance measures can be based on gauge functions, they are not gauges themselves. An easy way to

modify any gauge-based distance measure $D_{\mathcal{S}_+}$ to turn it into a gauge function is simply to consider instead the distance to the cone generated by $\mathcal{S}_+$:

$$\begin{aligned} D'_{\mathcal{S}_+}(\rho) &\coloneqq \inf_{X\in\mathcal{S}_+^{\star\star}} \Gamma_{\mathcal{C}}(\rho - X) \\ &= \sup\left\{ \langle W,\rho\rangle - \sup_{X\in\mathcal{S}_+^{\star\star}} \langle W, X\rangle \;\middle|\; W\in\mathcal{C}^\circ \right\} \\ &= \sup\left\{ \langle W,\rho\rangle \;\middle|\; W\in\mathcal{C}^\circ \cap (-\mathcal{S}_+^\star)\right\} \\ &= \Gamma_{\mathcal{C}\cup\mathcal{S}_+^{\star\star}}(\rho). \end{aligned} \tag{5.71}$$

As an explicit example, consider first the trace distance $T_{\mathcal{S}_+}(\rho)$, which corresponds to $D_{\mathcal{S}_+}$ with the choice $\mathcal{C} = \left\{ X \;\middle|\; \|X\|_1 \leq 1\right\}$ and thus gives

$$T_{\mathcal{S}_+}(\rho) = \sup\left\{ \langle W,\rho\rangle - \mu_{\mathcal{S}_+}(W) \;\middle|\; \|W\|_\infty \leq 1\right\}. \tag{5.72}$$

An alternative formulation of the trace distance can be obtained by recalling that the trace norm $\|\cdot\|_1$ admits a representation as the base norm $\Gamma_{\mathbb{D}\cup(-\mathbb{D})}$ in the space of Hermitian matrices, allowing us to write

$$\begin{aligned} T_{\mathcal{S}_+}(\rho) &= \inf_{\sigma\in\mathrm{conv}(\mathcal{S}_+)} \sup\left\{ \langle W,\rho-\sigma\rangle \;\middle|\; W\in\mathbb{D}^\circ\cap(-\mathbb{D})^\circ\right\} \\ &= \inf_{\sigma\in\mathrm{conv}(\mathcal{S}_+)} \sup\left\{ \left\langle W',\rho-\sigma\right\rangle - \langle \mathbb{1},\rho\rangle + \langle \mathbb{1},\sigma\rangle \;\middle|\; W'\in\left(\frac{1}{2}\mathbb{D}\right)^\circ\cap\mathbb{D}^\star\right\} \\ &= 2\sup\left\{ \left\langle W',\rho\right\rangle - \Gamma^\circ_{\mathcal{S}_+}(W') \;\middle|\; \mathbb{1}\geq W'\geq 0\right\} \end{aligned} \tag{5.73}$$

where in the second line we make the change of variables $W' = W + \mathbb{1}$, and the last equality follows because $\mu_{\mathcal{S}_+}(W') = \Gamma^\circ_{\mathcal{S}_+}(W')$ for any $W'\in\mathbb{D}^\star = \mathbb{H}_+$. The trace distance can then be turned into a gauge function by considering $T'_{\mathcal{S}_+}(\rho) = \min_{X\in\mathcal{S}_+^{\star\star}} \|\rho - X\|_1$, a quantity known as the *modified trace distance*. We can relate it to the generalised robustness as

$$\begin{aligned} T'_{\mathcal{S}_+}(\rho) &= \sup\left\{ -\langle\rho, W\rangle \;\middle|\; W\in\mathbb{D}^\circ\cap(-\mathbb{D}^\circ)\cap\mathcal{S}_+^\star\right\} \\ &\leq \sup\left\{ -\langle\rho, W\rangle \;\middle|\; W\in\mathbb{D}^\circ\cap\mathcal{S}_+^\star\right\} \\ &= R^{\mathbb{D}}_{\mathcal{S}_+}(\rho). \end{aligned} \tag{5.74}$$

The advantage of using $T'_{\mathcal{S}_+}$ over $T_{\mathcal{S}_+}$ is, once again, that the modified trace distance is a gauge function, and therefore immediately inherits all of the desirable properties of gauges in resource quantification that we discussed in Sec. 5.3. As an example, a long-standing conjecture that the trace distance of entanglement is a strong monotone [115] was recently put to rest by employing results from the resource theory of coherence to show that the trace distance does *not* satisfy strong monotonicity under the relevant resource-free operations [116,117] — this gives an interesting example of a case where it becomes necessary to consider distances with respect to the unnormalised cone $\mathcal{S}_+^{\star\star}$ in order to ensure strong monotonicity.

[115] Eisert et al., JPA 2003

[116] Yu et al., PRA 2016

[117] Qiao et al., preprint 2017

5.4.2 Geometric measures

The family of the so-called *geometric measures* with respect to $\mathcal{S}_+$ is defined by noting that, for pure states, the quantity $\Gamma^\circ_{\mathcal{V}}(|\psi\rangle)$ effectively quantifies the resource contained in a state by measuring the largest possible overlap with a free state [118,119], reaching its maximal value 1 only on the set of free states. One can then define a resource quantifier simply as

[118] Shimony, Ann. NY Ac. 1995

[119] Wei & Goldbart, PRA 2003

$$G_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) \coloneqq 1 - \Gamma^\circ_{\mathcal{V}}(|\psi\rangle)^2. \tag{5.75}$$

This function can now be extended to all mixed states through the convex roof, giving

$$\begin{aligned} G_{\mathcal{S}_+}(\rho) &\coloneqq \left(1 - \Gamma^{\circ 2}_{\mathcal{V}}\right)^{\cup}(\rho) = 1 - \Gamma^{\circ\cap}_{\mathcal{S}_+}(\rho) \\ &= 1 - \max\left\{ \sum_i p_i \, \Gamma^\circ_{\mathcal{V}}(|\psi_i\rangle)^2 \;\middle|\; \rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i|, \; p_i \in \mathbb{R}_+, \; \sum_i p_i = 1 \right\}. \end{aligned} \tag{5.76}$$

Since $G_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = 1 - \max_{|\eta\rangle\in\mathcal{V}} |\langle\psi|\eta\rangle|^2 = 1 - \sup_{\sigma\in\operatorname{conv}\mathcal{S}_+} F(|\psi\rangle\langle\psi|, \sigma)$, one could expect this measure to be related to the fidelity-based distance quantifier:

$$G'_{\mathcal{S}_+}(\rho) \coloneqq 1 - \max_{\sigma\in\operatorname{conv}(\mathcal{S}_+)} F(\rho, \sigma). \tag{5.77}$$

Surprisingly, the two quantities are in fact equal, which can be seen as follows.

Theorem 5.26 (Streltsov et al. [120]). *For any $\rho \in \mathbb{D}$,*

[120] Streltsov et al., NJP 2010

$$\max_{\sigma\in\operatorname{conv}\mathcal{S}_+} F(\rho, \sigma) = \max\left\{ \sum_i p_i \max_{\sigma\in\operatorname{conv}\mathcal{S}_+} F(|\psi_i\rangle\langle\psi_i|, \sigma) \;\middle|\; \rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \right\} \tag{5.78}$$

where the maximisation is over all pure-state convex decompositions of ρ. In particular,

$$G_{\mathcal{S}_+}(\rho) = G'_{\mathcal{S}_+}(\rho). \tag{5.79}$$

Proof. By Uhlmann's theorem [28], the left-hand side can be expressed as

[28] Uhlmann, Rep. Math. Phys. 1976

$$\max_{\sigma\in\mathcal{S}_+} F(\rho, \sigma) = \max_{\sigma\in\mathcal{S}_+} \max_{\Psi_\sigma, \Psi_\rho} \left|\langle\Psi_\rho|\Psi_\sigma\rangle\right|^2, \tag{5.80}$$

where the internal maximisation on the right-hand side is over all purifications of ρ and σ. Furthermore, we can restrict ourselves to a *fixed* purification of σ [121]. If $\sigma = \sum_i q_i \, |\eta_i\rangle\langle\eta_i|$ is a decomposition of σ into pure states $|\eta_i\rangle \in \mathcal{V}$, we choose $|\Psi_\sigma\rangle \coloneqq \sum_i \sqrt{q_i} \, |\eta_i\rangle \, |i\rangle$. Any purification of ρ over the same system can then be expanded as $|\Psi_\rho\rangle = \sum_i \sqrt{p_i} \, |\psi_i\rangle \, |i\rangle$ with respect to the orthonormal basis $\{|i\rangle\}$ on the purifying system. Here, $\{p_i, |\psi_i\rangle\}$ forms a pure-state decomposition of

[121] Jozsa, J. Mod. Opt. 1994

ρ. Since $\langle\Psi_\sigma|\Psi_\rho\rangle = \sum_i \sqrt{p_i q_i}\,\langle\eta_i|\psi_i\rangle$, we deduce that

$$\begin{aligned}\max_{\sigma\in\mathcal{S}_+} F(\rho,\sigma) &= \max\left\{\left|\sum_i \sqrt{p_i q_i}\,\langle\eta_i|\psi_i\rangle\right|^2 \;\middle|\; \rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|\,,\right.\\ &\qquad\left.\omega = \sum_i q_i\,|\eta_i\rangle\langle\eta_i|\,,\ |\eta_i\rangle\in\mathcal{V}\right\} \qquad (5.81)\\ &= \max\left\{\sum_i p_i\,|\langle\eta_i|\psi_i\rangle|^2 \;\middle|\; \rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|\,,\ |\eta_i\rangle\in\mathcal{V}\right\}\\ &= \max\left\{\sum_i p_i \max_{\sigma\in\mathcal{M}_m} F(|\psi_i\rangle\langle\psi_i|,\sigma) \;\middle|\; \rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|\right\}.\end{aligned}$$

■

The above result is rather remarkable in that it reduces the quantification of the convex roof–based measure $G_{\mathcal{S}_+}$ to a much simpler convex optimisation problem.

$\Gamma^{\circ\cap}_{\mathcal{S}_+}$ is concave and therefore is larger on $\mathbb{D}$ than any convex function equal to it on the extremal points of the set, including in particular the gauge $\Gamma^{\circ}_{\mathcal{S}_+}$ itself. We thus obtain the bound

$$G_{\mathcal{S}_+}(\rho) \le 1 - \Gamma^{\circ}_{\mathcal{S}_+}(\rho) \qquad (5.82)$$

which holds for any density matrix ρ. From Prop. 5.13 we then have that

$$R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \ge \frac{\mathrm{Tr}(\rho^2)}{\Gamma^{\circ}_{\mathcal{S}_+}(\rho)} - 1 \ge \frac{\mathrm{Tr}(\rho^2)}{1 - G_{\mathcal{S}_+}(\rho)} - 1. \qquad (5.83)$$

for any $\rho \in \mathrm{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$. This in particular establishes the relation

$$R^{\mathbb{D}}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) \ge \frac{1}{1 - G_{\mathcal{S}_+}(|\psi\rangle\langle\psi|)} - 1 = \frac{G_{\mathcal{S}_+}(|\psi\rangle\langle\psi|)}{1 - G_{\mathcal{S}_+}(|\psi\rangle\langle\psi|)} \qquad (5.84)$$

for all $|\psi\rangle \in \mathrm{span}(\mathcal{V})$, which generalises a known property of the robustness of entanglement [122].

[122] Cavalcanti, PRA 2006

5.4.3 *Relative entropy*

Let $D(\rho\|\sigma)$ denote the quantum relative entropy, defined as

$$D(\rho\|\sigma) = \begin{cases}\mathrm{Tr}\,(\rho\log\rho - \rho\log\sigma) & \text{if } \mathrm{supp}(\rho)\subseteq\mathrm{supp}(\sigma),\\ \infty & \text{otherwise.}\end{cases} \qquad (5.85)$$

This quantity is neither symmetric nor does it satisfy the triangle inequality, but it nevertheless can be treated as a quasi-metric and used as a measure to faithfully

quantify a given quantum resource [52,53]. A quantifier of interest is then the *relative entropy with respect to* $\mathcal{S}_+$,

[52] Vedral et al., PRL 1997

[53] Bengtsson & Życzkowski, 2007

$$S_{\mathcal{S}_+}(\rho) = \inf_{\sigma \in \operatorname{conv} \mathcal{S}_+} D(\rho\|\sigma). \tag{5.86}$$

Note that $\operatorname{dom}(S_{\mathcal{S}_+}) = \operatorname{dom}(R^{\mathbb{D}}_{\mathcal{S}_+})$.

One can obtain useful relations between this quantity and atomic gauges, generalising results for entanglement and coherence theories found for example in [99,123,124]. The connection between the two frameworks is made through quantities called the max- and min-relative entropy, defined as [99,125]

[99] Datta, IEEE TIT 2009

[123] Zhu et al., JPA 2017

[124] Rana et al., PRA 2017

[125] Dupuis et al., 2012

$$\begin{aligned} D_{\max}(\rho\|\sigma) &= \inf\left\{ s \in \mathbb{R}_+ \;\middle|\; \rho \preceq 2^s \sigma \right\} = \log \inf\left\{ s \in \mathbb{R}_+ \;\middle|\; \rho \preceq s\sigma \right\} \\ D_{\min}(\rho\|\sigma) &= -2 \log F(\rho, \sigma) \end{aligned} \tag{5.87}$$

We make note of the difference between this quantity and a different $D_{\min}$ defined in [99], which we will later encounter as the hypothesis testing relative entropy D_H. From the definition of the generalised robustness $R^{\mathbb{D}}_{\mathcal{S}_+}$, it then follows that

$$\inf_{\sigma \in \operatorname{conv}(\mathcal{S}_+)} D_{\max}(\rho\|\sigma) = \log\left(1 + R^{\mathbb{D}}_{\mathcal{S}_+}(\rho)\right), \tag{5.88}$$

making the $D_{\max}$-based measure strongly monotonic under resource non-generating operations. Using the relation $D_{\min}(\rho\|\sigma) \leq D(\rho\|\sigma) \leq D_{\max}(\rho\|\sigma)$ [99,125], we then have

$$G_{\mathcal{S}_+}(\rho) \leq -2\log(1 - G_{\mathcal{S}_+}(\rho)) \leq S_{\mathcal{S}_+}(\rho) \leq \log\left(1 + R^{\mathbb{D}}_{\mathcal{S}_+}(\rho)\right) \leq \log \Gamma_{\mathcal{S}}(\rho) \tag{5.89}$$

where the first inequality comes from the fact that $-2\log(x) \geq 1 - x \;\forall x \in [0,1]$, and the last inequality from Theorem 5.7. This establishes a quantitative relation between some non-convex resource monotones and the relative entropy. In the resource theory of quantum coherence, it has been conjectured that the inequality $S_{\mathcal{S}_+}(\rho) \leq \Gamma_{\mathcal{S}}(\rho) - 1$ also holds [124,126]. Following [124], we can use the fact that $x \geq \log(1+x) \;\forall x \geq 1$ to similarly establish

[126] Rana et al., PRA 2016

$$S_{\mathcal{S}_+}(\rho) \leq R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \leq \Gamma_{\mathcal{S}}(\rho) - 1 \qquad \forall\, \rho \text{ s.t. } R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \geq 1 \tag{5.90}$$

$$S_{\mathcal{S}_+}(\rho) \leq R^{\mathbb{D}}_{\mathcal{S}_+}(\rho) \log e \leq (\Gamma_{\mathcal{S}}(\rho) - 1)\log e \qquad \forall\, \rho, \tag{5.91}$$

and it is an interesting open question whether one can tighten the relation.

Gauges for physical resources

6

We now show how to apply the formalism introduced in the previous chapter to two representative quantum resources — entanglement and coherence — as well as the recently established resource theory of magic states, demonstrating the universality of the results. We introduce new gauge-based quantifiers as well as show that many known measures of the three resources fit into the gauge function framework, allowing for their characterisation in this formalism. An important point to note here is that the properties of the measures such as strong monotonicity under relevant free operations, faithfulness, as well as quantitative bounds and relations, all follow straightforwardly from our discussion in Ch. 5. Since these properties are often not easy to show explicitly for a given measure, one can benefit from exploiting the convex geometry of the quantifiers.

This chapter is based on results published in [BR4]. We begin with the discussion of quantum coherence, as the introduced measures will form a basis for our investigation of bipartite entanglement.

[BR4] Regula, JPA 2018

6.1 k-COHERENCE

Let us recall the setting of the resource theory of quantum coherence. The set of free pure states in this resource theory, which we will denote as $\mathcal{B}$, corresponds to state vectors which have only one coefficient in a fixed orthonormal basis $\{|i\rangle\}$, that is, each free pure state vector is a normalised scalar multiple of a basis vector. The free density matrices, defined in the usual way as the convex hull of $\left\{|\psi\rangle\langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{B}\right\}$, are called the *incoherent* states. Any state which cannot be written as a convex combination of incoherent pure states is then *coherent*, and can be used as a resource.

In some applications, however, not every coherent state is useful for the considered physical tasks [BR5,18,19,104,127,128] — in these circumstances, a fine-grained quantification of quantum coherence is necessary. Given a fixed orthonormal basis $\{|i\rangle\}$, the *coherence rank* $\mathrm{CR}(|\psi\rangle)$ is defined as the cardinality of the state vector $|\psi\rangle$, that is, the number of non-zero coefficients in this basis. The set of free pure states is then given by $\mathcal{B}^k \coloneqq \left\{|\phi\rangle \;\middle|\; \mathrm{CR}(|\psi\rangle) \leq k,\ \langle\psi|\psi\rangle = 1\right\}$ and we extend this to mixed states as done previously; for the simplicity of notation, we will hereafter work with the convex hull directly:

$$\mathcal{I}_+^k \coloneqq \mathrm{conv}\left\{|\psi\rangle\langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{B}^k\right\}. \tag{6.1}$$

[BR5] Regula et al., NJP 2018
[18] Killoran et al., PRL 2016
[19] Theurer et al., PRL 2017
[104] Ringbauer et al., preprint 2017
[127] Chin, PRA 2017
[128] Chin, JPA 2017

Any state $\rho \in \mathcal{I}_+^k$ is then a free state in the resource theory of $k+1$-coherence, with all states $\rho \notin \mathcal{I}_+^k$ being $k+1$-coherent. Note that $k = 1$, that is the quantification of 2-coherence, corresponds to the standard notion of quantum coherence, where the only resource-free states are convex combinations of the basis projectors; also, note that $\mathcal{I}_+^d = \mathbb{D}$. Let us also define the corresponding set $\mathcal{I}^k := \operatorname{conv}\left\{ |\psi\rangle\langle\phi| \;\middle|\; |\psi\rangle \in \mathcal{B}^k, |\phi\rangle \in \mathcal{B}^k \right\}$ as before.

6.1.1 $k = 1$ *and* $k = d$

The resource theory of quantum coherence ($k = 1$) has been investigated in many works, and a variety of quantifiers have been considered [55,61,62,82,129,130]. Applying the formalism of this work to the states $\mathcal{I}_+^k$ with $k = 1$, one can notice that we have:

[55] Winter & Yang, PRL 2016
[61] Streltsov et al., PRL 2015
[62] Napoli et al., PRL 2016
[82] Piani et al., PRA 2016
[129] Yuan et al., PRA 2015
[130] Qi et al., JPA 2017

$$\begin{aligned}\Gamma_{\mathcal{B}^1}(|\psi\rangle) &= \inf\left\{ \sum_i c_i \;\middle|\; |\psi\rangle = \sum_i c_i |v_i\rangle,\ c_i \in \mathbb{R}_+,\ |v_i\rangle \in \mathcal{B}^1 \right\} \\ &= \inf\left\{ \sum_i |c_i| \;\middle|\; |\psi\rangle = \sum_i c_i |i\rangle,\ c_i \in \mathbb{C} \right\} \\ &= \||\psi\rangle\|_{\ell_1}.\end{aligned} \tag{6.2}$$

Following in a similar way for the other quantifiers, we obtain several known norms and gauges for Hermitian matrices, many of which have been used as coherence measures:

Gauge	Also known as
$\Gamma_{\mathcal{B}^1}$	vector ℓ_1 norm $\|\cdot\|_{\ell_1}$
$\Gamma^\circ_{\mathcal{B}^1}$	vector infinity norm $\|\cdot\|_{\ell_\infty}$
$\Gamma_{\mathcal{I}^1}$	element-wise ℓ_1 norm $\|\cdot\|_{\ell_1}$
$\Gamma^\circ_{\mathcal{I}^1}$	element-wise max norm $\|\cdot\|_{\ell_\infty}$
$\Gamma^\circ_{\mathcal{I}^1_+}$	maximum of diagonal elements and 0
$\Gamma^\circ_{\mathcal{I}^1_+ \cup (-\mathcal{I}^1_+)}$	maximum of absolute diagonal elements
$R^{\mathcal{I}^1_+}_{\mathcal{I}^1_+}$	(standard robustness of coherence)
$R^{\mathbb{D}}_{\mathcal{I}^1_+}$	generalised robustness of coherence [82]
$\Gamma^{\cup}_{\mathcal{I}^1_+} - 1$	coherence concurrence [129, 130]
$1 - \Gamma^{\circ\cap}_{\mathcal{I}^1_+}$	geometric measure of coherence[61]

Note that the set $\mathcal{I}_+^1$ only spans the set of diagonal matrices in the given basis, which means that the measure $R^{\mathcal{I}^1_+}_{\mathcal{I}^1_+}$ is not a practically useful measure of 2-coherence as it is infinite for $\rho \notin \mathcal{I}_+^1$.

For comparison, we also characterise the quantities obtained for $k = d$, i.e. when $\mathcal{I}_+^d = \mathbb{D}$:

Gauge	Also known as
$\Gamma_{\mathcal{B}^d}$	vector ℓ_2 norm $\|\cdot\|_{\ell_2}$
$\Gamma^\circ_{\mathcal{B}^d}$	vector ℓ_2 norm $\|\cdot\|_{\ell_2}$
$\Gamma_{\mathcal{I}^d}$	trace (nuclear) norm $\|\cdot\|_1$
$\Gamma^\circ_{\mathcal{I}^d}$	operator (spectral) norm $\|\cdot\|_\infty$
$\Gamma^\circ_{\mathcal{I}^d_+}$	maximum of the eigenvalues and 0
$\Gamma^\circ_{\mathcal{I}^d_+ \cup (-\mathcal{I}^d_+)}$	spectral (numerical) radius [131]
$R^{\mathcal{I}^d_+}_{\mathcal{I}^d_+}$	trace of positive part −1
$R^{\mathbb{D}}_{\mathcal{I}^d_+}$	trace of positive part −1
$\Gamma^{\cup}_{\mathcal{I}^d_+} - 1$	trace norm −1
$1 - \Gamma^{\circ\cap}_{\mathcal{I}^d_+}$	1− trace

[131] Horn & Johnson, 2012

By the trace of the positive part of a Hermitian matrix M we understand the quantity $\mathrm{Tr}(\Pi_+ M)$ with Π_+ denoting the orthogonal projection onto the subspace spanned by the non-negative eigenvalues of M. Further, we recall that $\Gamma_{\mathcal{I}^d_+ \cup (-\mathcal{I}^d_+)}$ corresponds to the trace norm itself for all Hermitian matrices [84].

[84] Hartkämper & Neumann, 1974

6.1.2 $1 \leq k \leq d$

We will now consider the extension of the atomic gauge formalism to arbitrary k. Several of the measures in this resource theory have been defined in the literature already: they are the geometric measure of k-coherence [BR5] and the generalised robustness $R^{\mathbb{D}}_{\mathcal{I}^k_+}$ [104]. All of the other measures that we have considered in Chapter 5 can be defined analogously. We will consider some explicit examples.

[BR5] Regula et al., NJP 2018

[104] Ringbauer et al., preprint 2017

The sets $\mathcal{I}^k_+$ for $k \geq 2$ have been shown to span the whole space $\mathbb{H}$ [104], and so $R^{\mathcal{I}^k_+}_{\mathcal{I}^k_+}$ can be used as a quantifier for $k \geq 2$. The other considered measures $\Gamma_{\mathcal{B}^k}$, $\Gamma_{\mathcal{I}^k}$, $\Gamma^{\cup}_{\mathcal{I}^k_+}$, and $R^{\mathbb{D}}_{\mathcal{I}^k_+}$ have no domain problems because $\mathrm{span}(\mathcal{I}^k) = \mathbb{C}^d \ \forall k$.

The atomic vector norm on the set $\mathcal{B}^k$, called the k-support norm $\|\cdot\|_{(k)}$, was introduced in [132] in the context of machine learning and optimisation. The norm is in fact exactly computable for any k.

[132] Argyriou et al., 2012

Definition 6.1 (Argyriou et al. [132])**.** Let $\mathscr{P}_k$ be the set of all possible subsets of $\{1, \ldots, d\}$ of at most k elements, and let $\mathrm{esupp}(|x\rangle) \coloneqq \left\{ i \;\middle|\; x_i \neq 0 \right\}$ where x_i is the ith coefficient of $|x\rangle$ in the given basis. The *k-support norm*[1] is defined as the atomic norm for the set $\mathcal{B}^k$ and is given

by

$$\begin{aligned}\Gamma_{\mathcal{B}^k}(|x\rangle) &= \||x\rangle\|_{(k)} \\ &\coloneqq \inf\left\{ \sum_{I\in\mathscr{P}_k} \||v_I\rangle\|_{\ell_2} \;\middle|\; \operatorname{esupp}(|v_I\rangle) \subseteq I,\ \sum_{I\in\mathscr{P}_k} |v_I\rangle = |x\rangle \right\} \\ &= \sqrt{\sum_{i=1}^{k-r-1} |x_i^{\downarrow}|^2 + \frac{1}{r+1}\left(\sum_{i=k-r}^{d} |x_i^{\downarrow}|\right)^2} \end{aligned} \tag{6.3}$$

where $x^{\downarrow}$ denotes the coefficients of $|x\rangle$ sorted in non-increasing order by magnitude ($|x_1^{\downarrow}| \geq |x_2^{\downarrow}| \geq \ldots$), and r is the unique integer in $\{1, \ldots, k-1\}$ satisfying

$$|x_{k-r-1}^{\downarrow}| > \frac{1}{r+1}\sum_{i=k-r}^{d} |x_i^{\downarrow}| \geq |x_{k-r}^{\downarrow}| \tag{6.4}$$

or 0 if no such integer exists.

It is explicit from Eq. (6.3) that the k-support norm interpolates between the ℓ_1 and ℓ_2 norms for vectors — indeed, we have that $\|\cdot\|_{(1)} = \|\cdot\|_{\ell_1}$ and $\|\cdot\|_{(d)} = \|\cdot\|_{\ell_2}$.

The polar (dual norm) of the k-support norm is given by

$$\||x\rangle\|_{(k)}^{\circ} = \sqrt{\sum_{i=1}^{k} |x_i^{\downarrow}|^2}, \tag{6.5}$$

that is, the ℓ_2 norm of the k largest coefficients. In general, we have that $\||x\rangle\|_{\ell_1} \geq \||x\rangle\|_{(k)} \geq \||x\rangle\|_{\ell_2} \geq \||x\rangle\|_{(k)}^{\circ} \geq \||x\rangle\|_{\ell_\infty}$.

We remark that an alternative way to derive the exact formula for the k-support norm is to start with the dual norm $\|\cdot\|_{(k)}^{\circ}$, which is easier to compute explicitly [BR5], and apply a duality result from [134] which immediately gives the form in Eq. (6.3). This approach was taken in [105] for a related problem in entanglement theory.

[BR5] Regula et al., NJP 2018

[134] Mudholkar & Freimer, Proc. Amer. Math. Soc. 1985

[105] Johnston & Kribs, Houst. J. Math. 2015

[1] The terminology of "k-support" comes from the fact that the set of non-zero coefficients of a matrix is frequently referred to as "support" in statistical learning and optimisation literature [133]. Since this differs from the usual definition of "support" in quantum information, to avoid confusion we use the notation of esupp to refer to the "element-wise support".

The atomic norm $\Gamma_{\mathcal{I}^k}$ on the set $\mathcal{I}^k$ is the *(k,k)-trace norm*, defined as [85,86]:

[85] Richard et al., 2014

[86] Richard et al., preprint 2014

$$\begin{aligned}\|\rho\|_{(k,k)} &= \inf\left\{\sum_i \||a_i\rangle\|_{(k)}\,\||b_i\rangle\|_{(k)} \;\middle|\; \rho = \sum_i |a_i\rangle\langle b_i|\right\}\\ &= \inf\left\{\sum_{\substack{I\in\mathscr{P}_k\\ J\in\mathscr{P}_k}} \|Z_{I,J}\|_1 \;\middle|\; \operatorname{esupp}(Z_{I,J})\subseteq I\times J,\ \rho=\sum_{I,J} Z_{I,J}\right\} \qquad (6.6)\\ \|\rho\|^{\circ}_{(k,k)} &= \sup\left\{|\langle a|\rho|b\rangle| \;\middle|\; |a\rangle, |b\rangle \in \mathcal{V}^k\right\} \qquad (6.7)\end{aligned}$$

where $\operatorname{esupp}(M)$ is defined for a matrix with elements M_{ij} as $\operatorname{esupp}(M) = \{(i,j) \mid M_{ij}\neq 0\}$. This norm interpolates between the entrywise ℓ_1 norm for $\|\cdot\|_{(1,1)}$ and the standard trace norm for $\|\cdot\|_{(d,d)}$. The (k,k)-trace norm is therefore a natural generalisation of the ℓ_1-norm of coherence, a fundamental measure of quantum coherence [15,126], to the formalism of k-coherence.

[15] Baumgratz et al., PRL 2014

[126] Rana et al., PRA 2016

While $\|\cdot\|_{(k,k)}$ is in general NP-hard to compute exactly, a numerical algorithm to approximate it was introduced in [85]. Note also that the generalised robustness of k-coherence $R^{\mathbb{D}}_{\mathcal{I}^k_+}$ was shown to be computable with a semidefinite program (SDP) [104]. It follows easily that the same property holds for the standard robustness $R^{\mathcal{I}^k_+}_{\mathcal{I}^k_+}$, and since this quantity is well-defined for any $k\geq 2$, we obtain tight bounds for the (k,k)-trace norm as

[104] Ringbauer et al., preprint 2017

$$R^{\mathcal{I}^k}_{\mathcal{I}^k}(\rho) \geq \|\rho\|_{(k,k)} - 1 \geq R^{\mathbb{D}}_{\mathcal{I}^k}(\rho). \qquad (6.8)$$

Interestingly, it has been shown that $R^{\mathcal{I}^k}_{\mathcal{I}^k}(|\psi\rangle\langle\psi|) = \||\psi\rangle\|^2_{(k)} - 1$ for any pure state [BR11], which means that both of the above inequalities are tight for pure states.

[BR11] Johnston et al., preprint 2018

The convex roof extension of the k-support norm $\Gamma^{\cup}_{\mathcal{I}^k_+}$ is the natural convex roof measure in this formalism. By Theorem 5.19, both the convex roof–extended k-support norm as well as the (k,k)-trace norm are strongly monotonic under k-incoherent operations [104]. We remark that another convex roof–based measure was defined in [127], based on an entanglement monotone called k-concurrence; it is not a generalisation of the ℓ_1 norm in the present framework. A quantitative comparison of the k-concurrence and an equivalent entanglement monotone based on the k-support norm will be considered in the next section.

[127] Chin, PRA 2017

6.2 BIPARTITE ENTANGLEMENT

Let $|\psi\rangle \in \mathbb{C}^{d_A}\otimes\mathbb{C}^{d_B}$ denote a bipartite pure state shared between two parties A and B, with the dimensions of the corresponding spaces denoted as d_A and d_B. A pure state is called separable or a product state (in the $A|B$ bipartition) if it can be written as $|\psi\rangle = |\psi_A\rangle\otimes|\psi_B\rangle$, and entangled otherwise. We denote the set

of all product state vectors as $\mathcal{V}$, and we define $\mathcal{S}_+$ to be the set of all separable density matrices defined through the convex hull. The *Schmidt rank* of a state $|\psi\rangle$ is defined as

$$\mathrm{SR}(|\psi\rangle) = \min\left\{ r \in \mathbb{N} \;\middle|\; |\psi\rangle = \sum_{i=1}^{r} \lambda_i |v_i\rangle\,,\ |v_i\rangle \in \mathcal{V} \right\}. \tag{6.9}$$

Then $|\psi\rangle$ is unentangled if and only if $\mathrm{SR}(|\psi\rangle) = 1$, and in general we have $1 \leq \mathrm{SR}(|\psi\rangle) \leq \min(d_A, d_B)$. A crucial property of bipartite entanglement is that every state can be expressed in the so-called Schmidt decomposition as

$$|\psi\rangle = \sum_{i=1}^{\mathrm{SR}(|\psi\rangle)} \lambda_i |a_i\rangle \otimes |b_i\rangle \tag{6.10}$$

where $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ form orthonormal bases for $\mathbb{C}^{d_A}$ and $\mathbb{C}^{d_B}$, respectively, and the terms λ_i are called *Schmidt coefficients*.

As previously, we now want to consider situations where not all entangled states are resourceful; that is, for our particular task, a state with at least a given Schmidt rank k is necessary, and we would like to quantify the entanglement corresponding to a particular Schmidt rank [135–138]. We define the Schmidt vector $\xi(|\psi\rangle)$ of a state $|\psi\rangle$ to be the vector consisting of its Schmidt coefficients including the zero terms, so that it is always $\min(d_A, d_B)$-dimensional. Since the Schmidt rank of a state corresponds to the cardinality of the Schmidt vector just as the coherence rank corresponds to the cardinality of the state vector in the given basis, in the case of pure states, the resource theory of bipartite entanglement of Schmidt rank k can be seen to be analogous to the resource theory of k-coherence applied to the Schmidt vector. Indeed, this relation between the two resource theories has been used to relate the corresponding quantifiers [123,139,140].

[135] Terhal & Horodecki, PRA 2000

[136] Chruściński & Kossakowski, CMP 2009

[137] Johnston & Kribs, JMP 2010

[138] Sentís et al., PRL 2016

[123] Zhu et al., JPA 2017

[139] Chen et al., PRA 2016

[140] Zhu et al., PRA 2017

We now define the relevant sets of free pure state vectors and free density matrices as, respectively:

$$\begin{aligned} \mathcal{V}^k &= \left\{ |\psi\rangle \;\middle|\; \mathrm{SR}(|\psi\rangle) \leq k, \||\psi\rangle\|_{\ell_2} = 1 \right\} \\ \mathcal{S}_+^k &= \mathrm{conv}\left\{ |\psi\rangle\langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{V}^k \right\} \end{aligned} \tag{6.11}$$

and define the corresponding set

$$\mathcal{S}^k = \mathrm{conv}\left\{ |\psi\rangle\langle\phi| \;\middle|\; |\psi\rangle \in \mathcal{V}^k,\ |\phi\rangle \in \mathcal{V}^k \right\}. \tag{6.12}$$

$\mathcal{S}_+^1$ is then the set of separable states, and $\mathcal{S}_+^d = \mathbb{D}$. A mixed state $\rho \in \mathcal{S}_+^k$ is said to have Schmidt number at most k [135]. The dual cone $\mathcal{S}_+^{k\star}$ defines the set of the so-called k-block positive operators [141].

[141] Skowronek et al., JMP 2009

Note that for any k, we have that $\mathrm{span}(\mathcal{V}^k) = \mathbb{C}^d$ and $\mathrm{span}(\mathcal{S}_+^k) = \mathbb{H}$ [142], which in particular means that all of the considered functions are finite for any

[142] Życzkowski et al., PRA 1998

density matrix, and that the relevant symmetric gauges ($\Gamma_{\mathcal{V}^k}, \Gamma_{\mathcal{S}_+^k \cup (-\mathcal{S}_+^k)}, \Gamma_{\mathcal{S}^k}$) all define valid norms.

The set of operations of interest in the theory of entanglement — LOCC — is a strict subset of both the resource non-generating (separability-preserving) and the resource-free (separable) operations [107], which by the results of Sec. 5.3 means that all of the considered measures are strong monotones under LOCC. We stress that the measures introduced herein are monotonic not only under free operations which do not generate entanglement, but also under a larger set of operations which do not generate entanglement of Schmidt rank $k+1$.

[107] Chitambar et al., CMP 2014

6.2.1 $k = 1$

For $k = 1$, we obtain many familiar measures and quantifiers in this formalism:

Gauge	Also known as
$\Gamma_{\mathcal{V}^1}$	sum of Schmidt coefficients
$\Gamma^{\circ}_{\mathcal{V}^1}$	largest Schmidt coefficient [118]
$\Gamma_{\mathcal{S}^1}$	greatest cross norm (projective tensor norm) [143]
$\Gamma^{\circ}_{\mathcal{S}^1}$	Schmidt operator norm (injective tensor norm) [137]
$\Gamma^{\circ}_{\mathcal{S}_+^1}$	—
$\Gamma^{\circ}_{\mathcal{S}_+^1 \cup (-\mathcal{S}_+^1)}$	product numerical radius [144]
$R^{\mathcal{S}_+^1}_{\mathcal{S}_+^1}$	robustness of entanglement [94]
$R^{\mathbb{D}}_{\mathcal{S}_+^1}$	generalised robustness of entanglement [94,101]
$\Gamma^{\cup}_{\mathcal{S}_+^1} - 1$	(2×) convex roof–extended negativity [145]
$1 - \Gamma^{\circ\cap}_{\mathcal{S}_+^1}$	geometric measure of entanglement [119]

[118] Shimony, Ann. NY Ac. 1995
[143] Rudolph, JPA 2000
[137] Johnston & Kribs, JMP 2010
[144] Gawron et al., JMP 2010
[94] Vidal & Tarrach, PRA 1999
[101] Steiner, PRA 2003
[145] Lee et al., PRA 2003
[119] Wei & Goldbart, PRA 2003

where by "—" we denote functions which, to our knowledge, have not been previously defined in the literature explicitly or do not correspond to known quantities, although they can of course be defined as detailed before.

The vector gauge $\Gamma_{\mathcal{V}^1}$ is the ℓ_1 norm of the Schmidt vector, corresponding to the sum of Schmidt coefficients and constituting a convex relaxation of the Schmidt rank of a pure state.

The negativity of a state is defined as $N(\rho) = \frac{1}{2}\left(\left\|\rho^{T_B}\right\|_1 - 1\right)$ [92,142], where ρ^{T_B} denotes the partial transpose. The negativity of pure states is precisely [92]

[92] Vidal & Werner, PRA 2002
[142] Życzkowski et al., PRA 1998

$$N(|\psi\rangle\langle\psi|) = \sum_{j<k} \lambda_j \lambda_k = \frac{\Gamma_{\mathcal{V}^1}(|\psi\rangle)^2 - 1}{2}, \tag{6.13}$$

which means that the convex roof–based quantifier $\Gamma^{\cup}_{\mathcal{S}^1}$ is (twice) the *convex roof–extended negativity*. This function was proposed as an alternative generalisation of concurrence to systems beyond two qubits [145], and it was further

suggested as the measure most suitable to characterise the so-called monogamy relations of entanglement between qudits [146–149], which the concurrence fails to satisfy. Here we also see that the convex roof–extended negativity arises as the natural gauge-based generalisation of the concurrence. The faithfulness and strong monotonicity of $\Gamma^{\cup}_{\mathcal{S}^1}$ under separable operations follow from Theorem 5.19.

[146] Ou & Fan, PRA 2007
[147] Kim et al., PRA 2009
[148] Choi & Kim, PRA 2015
[149] Tian et al., Sci. Rep. 2016

The atomic norm $\Gamma_{\mathcal{S}^1}$ has been first considered in the context of entanglement quantification by Rudolph [143,150]. In fact, both $\Gamma_{\mathcal{S}^1}$ and $\Gamma_{\mathcal{V}^1}$ are what is known as *projective tensor norms* [87,151], but defined on different underlying spaces — $\Gamma_{\mathcal{V}^1}$ on the tensor product of the Hilbert spaces, or equivalently on the tensor product $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ both endowed with the ℓ_2 norm; and $\Gamma_{\mathcal{S}^1}$ on the tensor product of the corresponding spaces of trace-class linear operators, or equivalently $\mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$ endowed with the trace norm. Specifically, we can write

[143] Rudolph, JPA 2000
[150] Rudolph, JMP 2001
[87] Jameson, 1987
[151] Pisier, 1986

$$\Gamma_{\mathcal{V}^1}(|\psi\rangle) = \inf\left\{ \sum_i \left\| \left|x_A^{(i)}\right\rangle \right\|_{\ell_2} \left\| \left|x_B^{(i)}\right\rangle \right\|_{\ell_2} \;\middle|\; |\psi\rangle = \sum_i \left|x_A^{(i)}\right\rangle \otimes \left|x_B^{(i)}\right\rangle, \; \left|x_A^{(i)}\right\rangle \in \mathbb{C}^{d_A}, \; \left|x_B^{(i)}\right\rangle \in \mathbb{C}^{d_B} \right\} \tag{6.14}$$

and

$$\Gamma_{\mathcal{S}^1}(\rho) = \min\left\{ \sum_i \left\| X_A^{(i)} \right\|_1 \left\| X_B^{(i)} \right\|_1 \;\middle|\; \rho = \sum_i X_A^{(i)} \otimes X_B^{(i)}, \; X_A^{(i)} \in \mathbb{C}^{d_A \times d_A}, \; X_B^{(i)} \in \mathbb{C}^{d_B \times d_B} \right\}. \tag{6.15}$$

This characterisation will allow for a straightforward generalisation of the quantities later. Also, notice that the base norm $\Gamma_{\mathcal{S}^1_+ \cup (-\mathcal{S}^1_+)}$ can be expressed simply by replacing the optimisation over arbitrary complex matrices in the above with an optimisation over Hermitian matrices only [93], that is,

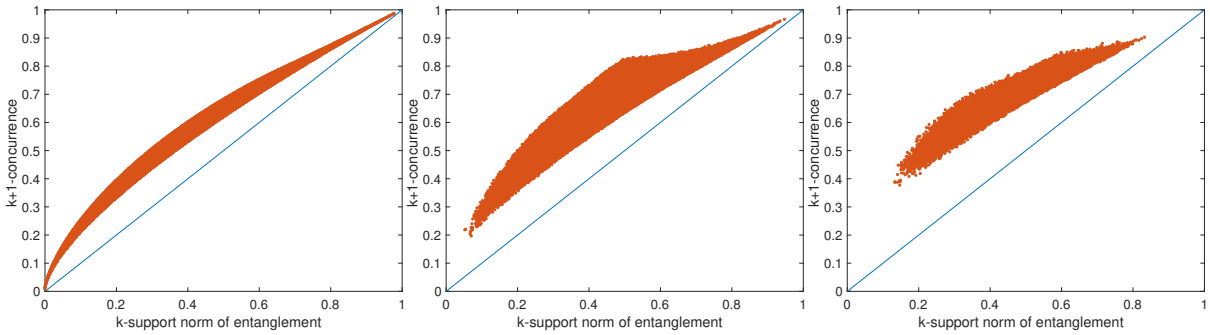
[93] Rudolph, QIP 2005

$$\Gamma_{\mathcal{S}^1}(\rho) = \min\left\{ \sum_i \left\| X_A^{(i)} \right\|_1 \left\| X_B^{(i)} \right\|_1 \;\middle|\; \rho = \sum_i X_A^{(i)} \otimes X_B^{(i)}, \; X_A^{(i)} \in \mathbb{H}^{d_A \times d_A}, \; X_B^{(i)} \in \mathbb{H}^{d_B \times d_B} \right\}. \tag{6.16}$$

6.2.2 $k \geq 1$

The generalisation of the robustness measures to the set of bipartite states with a given Schmidt number k was considered by Clarisse [102], although exact quantitative expressions were not previously known, even for pure states. Many of the other gauge-based quantities were later described by Johnston and Kribs in [105,137,153]. Similarly to the case of coherence, we obtain a hierarchy of quantifiers, each corresponding to a different level of bipartite entanglement.

[102] Clarisse, JPA 2006
[105] Johnston & Kribs, Houst. J. Math. 2015
[137] Johnston & Kribs, JMP 2010
[153] Johnston & Kribs, QIC 2011



(a) $d_A = d_B = 3,\ k = 2$ **(b)** $d_A = d_B = 4,\ k = 2$ **(c)** $d_A = d_B = 6,\ k = 3$

Figure 6.1: Comparison of the different Schmidt rank k entanglement measures for 10^6 randomly generated pure states (uniformly distributed with respect to the Haar measure). On the x axis is the quantity $\Gamma_{\mathcal{V}^k}(|\psi\rangle)^2 - 1 = \|\xi(|\psi\rangle)\|_{(k)}^2 - 1$ (with a normalisation factor $\frac{k}{d_A - k}$), and on the y axis is the $k+1$-concurrence monotone from [152].

The quantity $\Gamma_{\mathcal{V}^k}$ is nothing but the k-support norm of the Schmidt vector, $\|\xi(|\psi\rangle)\|_{(k)}$. This gives a natural generalisation of the convex roof–extended negativity to a measure of Schmidt rank k entanglement:

$$\Gamma^{\cup}_{\mathcal{S}^k}(\rho) = \inf\left\{ \sum_i p_i \|\xi(|\psi_i\rangle)\|^2_{(k)} \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|,\ p_i \in \mathbb{R}_+,\ \sum_i p_i = 1 \right\}. \tag{6.17}$$

Another common convex roof–based measure of Schmidt rank k entanglement is Gour's $k+1$-concurrence [152] — a comparison between the values of $\Gamma_{\mathcal{V}^k}$ and the $k+1$-concurrence on pure states can be found in Fig. 6.1.

[152] Gour, PRA 2005

The choice of $\Gamma_{\mathcal{V}^k}$ as a quantifier of pure-state entanglement of Schmidt rank $k+1$ has the advantage that several natural resource measures such as the robustness $R^{\mathbb{D}}_{\mathcal{S}^k_+}$ reduce to it on all pure states. In fact, it has been shown in [BR11] that the robustness $R^{\mathcal{S}^k_\ddagger}_{\mathcal{S}^k_+}$ of a pure state is upper bounded by $\Gamma_{\mathcal{V}^k}(|\psi\rangle)^2 - 1$, which implies that $R^{\mathcal{S}^k_\ddagger}_{\mathcal{S}^k_+}(|\psi\rangle\langle\psi|) = R^{\mathbb{D}}_{\mathcal{S}^k_+}(|\psi\rangle\langle\psi|)$. This generalises a known relation between the robustness of entanglement and negativity [94].

[BR11] Johnston et al., preprint 2018

[94] Vidal & Tarrach, PRA 1999

The norm $\Gamma_{\mathcal{S}^k}$ generalises the greatest cross norm to a faithful quantifier of entanglement of a given Schmidt rank [105]. One can also express $\Gamma_{\mathcal{S}^k}$ as the nuclear norm

[105] Johnston & Kribs, Houst. J. Math. 2015

$$\Gamma_{\mathcal{S}^k}(\rho) = \inf\left\{ \sum_i \|\xi(|x_i\rangle)\|_{(k)} \|\xi(|y_i\rangle)\|_{(k)} \;\middle|\; \rho = \sum_i |x_i\rangle\langle y_i| \right\}. \tag{6.18}$$

We remark that, just as in the case of coherence, the robustnesses $R^{\mathcal{S}^k_\ddagger}_{\mathcal{S}^k_+}$ and $R^{\mathbb{D}}_{\mathcal{S}^k_+}$ provide tight upper and lower bounds for this norm. Note also that the polar gauge $\Gamma^\circ_{\mathcal{S}^k}$ can be computed exactly in small dimensions, and in general bounded by semidefinite programs [153], allowing for an efficient characterisation of k-block positive operators [137].

[153] Johnston & Kribs, QIC 2011

[137] Johnston & Kribs, JMP 2010

The quantity $1 - \Gamma^{\circ\cap}_{\mathcal{S}^k_+}$ provides a generalisation of the geometric measure of entanglement, and in fact corresponds to a family of convex roof–based monotones introduced by Vidal [154]. These monotones can be computed exactly for special cases such as isotropic or Werner states [155].

[154] Vidal, PRL 1999

[155] Girard & Gour, PRA 2017

6.2.3 *Remarks about the resource theory of negative partial transpose*

The set of states with positive partial transpose is $\mathsf{PPT}_+ = \left\{\rho \in \mathbb{D} \;\middle|\; \rho^{T_B} \in \mathbb{H}_+\right\}$. This set is not a convex hull of rank-one terms, and therefore does not precisely fit into the formalism considered here; however, many of our considerations, including in particular the robustness quantifiers, apply to it very straightforwardly.

Recall that $\mathcal{S}^1_+ \subseteq \mathsf{PPT}_+$ and note that the separable operations (resource-free operations in the resource theory of bipartite entanglement) are a subset of the so-called PPT-preserving operations (resource non-generating operations in the resource theory corresponding to PPT states) [34,156]. What this means in particular is that, by considering the resource theory in which PPT_+ is the set of free states, any monotone under PPT-preserving operations is an entanglement monotone under separable operations — including the quantities $R^{\mathsf{PPT}_+}_{\mathsf{PPT}_+}$ and $R^{\mathbb{D}}_{\mathsf{PPT}_+}$.

[34] Horodecki, PLA 1997

[156] Rains, IEEE TIT 2001

A different choice of the reference set of free states is often made, taking instead the set of unit trace Hermitian matrices with a positive partial transpose, $\mathsf{PPT} = \left\{X \in \mathbb{H}_1 \;\middle|\; X^{T_B} \in \mathbb{H}_+\right\}$. The gauge function formalism straightforwardly applies to such cases as well — indeed, the negativity of any mixed state, defined as $N(\rho) = \frac{1}{2}\left(\left\|\rho^{T_B}\right\|_1 - 1\right)$, can be written simply as $N(\rho) = R^{\mathsf{PPT}}_{\mathsf{PPT}}(\rho)$ [81,92]. The function $R^{\mathsf{PPT}_+}_{\mathsf{PPT}_+}$ was in fact mentioned in the original work of Vidal and Werner [92] as an alternative to negativity, leading to the relation $R^{\mathcal{S}^1_+}_{\mathcal{S}^1_+}(\rho) \geq R^{\mathsf{PPT}_+}_{\mathsf{PPT}_+}(\rho) \geq N(\rho)$ by the set inclusion of their unit balls. As we remarked before, it is known that $R^{\mathcal{S}^1_+}_{\mathcal{S}^1_+}(|\psi\rangle\langle\psi|) = 2N(|\psi\rangle\langle\psi|)$ [92]. More recently, other gauge functions in the resource theory of negative partial transpose have been employed to provide bounds on distillable entanglement [157], and we will encounter further examples in Ch. 8.

[81] Plenio & Virmani, QIC 2007

[92] Vidal & Werner, PRA 2002

[157] Wang & Duan, PRA 2016

These considerations generalise also to multipartite entanglement in the form of quantifiers such as the genuine multipartite negativity [158,159], as we will consider explicitly in the next section. Furthermore, we will explicitly investigate the quantitative and operational properties of measures associated with the sets PPT_+ and PPT in Chapter 8.

[158] Jungnitsch et al., PRL 2011

[159] Hofmann et al., JPA 2014

6.3 k-PARTITE ENTANGLEMENT

Let us now consider a system consisting of $n \geq 2$ parties, $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}$. Similarly to the hierarchy of the Schmidt rank in the case of bipartite entangle-

ment, one can define a hierarchy of multipartite entanglement. We will take $\mathcal{P}^k$ with $1 \leq k \leq n$ to be the set of pure states which are k-producible — that is, they can be expressed as [160]

[160] Gühne et al., NJP 2005

$$|\psi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdot \otimes |\phi_m\rangle \tag{6.19}$$

where each of the m states $|\phi_i\rangle$ consists of at most k parties. Since such states are at most k-partite entangled, they are precisely the free states in the resource theory of $k+1$-partite entanglement. The convex hull of the corresponding set of matrices $\mathcal{M}_+^k = \operatorname{conv}\left\{ |a\rangle\langle a| \;\middle|\; |a\rangle \in \mathcal{P}^k \right\}$ then defines the set of k-producible mixed states, and we analogously define $\mathcal{M}^k = \operatorname{conv}\left\{ |a\rangle\langle b| \;\middle|\; |a\rangle, |b\rangle \in \mathcal{P}^k \right\}$. Note that $\mathcal{M}_+^1$ is the set of fully separable states and $\mathcal{M}_+^n = \mathbb{D}$.

Unlike the theory of bipartite entanglement, our understanding of the theory of multipartite entanglement is still very limited [7,161]. While one can straightforwardly generalise the quantities defined previously for bipartite entanglement — for instance, as the robustness of k-partite entanglement [103], geometric measure of k-partite entanglement [162], and product numerical radius [163] — their quantification and characterisation is in general a much more difficult task.

[7] Horodecki et al., RMP 2009
[161] Walter et al., preprint 2016
[103] Cavalcanti et al., PRA 2005
[162] Blasone et al., PRA 2008
[163] Puchała et al., Lin. Alg. Appl. 2011

In the case of $k = 1$, the corresponding norm $\Gamma_{\mathcal{M}^1}$ generalises the greatest cross norm [150]. This quantity along with its dual norm $\Gamma^\circ_{\mathcal{M}^1}$ have been investigated in detail in [164], including the infinite-dimensional case. Again, the norms $\Gamma_{\mathcal{P}^1}$ and $\Gamma_{\mathcal{M}^1}$ are both projective tensor norms, and can be written as

[150] Rudolph, JMP 2001
[164] Arveson, J. Funct. Anal. 2009

$$\begin{aligned}
\Gamma_{\mathcal{P}^1}(|\psi\rangle) = \inf\Bigg\{ \sum_i \left\| \left| a_1^{(i)} \right\rangle \right\|_{\ell_2} \left\| \left| a_2^{(i)} \right\rangle \right\|_{\ell_2} \cdots \left\| \left| a_n^{(i)} \right\rangle \right\|_{\ell_2} \;\Bigg|\; & \left| a_j^{(i)} \right\rangle \in \mathbb{C}^{d_j}, \\
& |\psi\rangle = \sum_i \left| a_1^{(i)} \right\rangle \otimes \left| a_2^{(i)} \right\rangle \otimes \cdots \otimes \left| a_n^{(i)} \right\rangle \Bigg\} \\
\Gamma_{\mathcal{M}^1}(\rho) = \inf\Bigg\{ \sum_i \left\| X_1^{(i)} \right\|_1 \left\| X_2^{(i)} \right\|_1 \cdots \left\| X_n^{(i)} \right\|_1 \;\Bigg|\; & \rho = \sum_i X_1^{(i)} \otimes X_2^{(i)} \otimes \cdots \otimes X_n^{(i)}, \\
& X_j^{(i)} \in \mathbb{C}^{d_j \times d_j} \Bigg\}
\end{aligned} \tag{6.20}$$

following [150,164], where d_j denotes the dimension of the Hilbert space of the jth system. Note that the base norm $\Gamma_{\mathcal{M}_+^1 \cup (-\mathcal{M}_+^1)}$ can be expressed by replacing $X_j^{(i)} \in \mathbb{C}^{d_j \times d_j}$ with $X_j^{(i)} \in \mathbb{H}^{d_j \times d_j}$ in Eq. (6.20) [93].

[93] Rudolph, QIP 2005

Generalising this norm to the fine-grained classification of k-partite entanglement, one obtains a hierarchy of norms akin to the one defined for Schmidt rank k entanglement. Explicitly, we have:

$$\begin{aligned}
\Gamma_{\mathcal{M}^k}(\rho) &= \inf\left\{ \sum_i c_i \;\middle|\; \rho = \sum_i c_i |v_i\rangle\langle w_i|,\ |v_i\rangle, |w_i\rangle \in \mathcal{P}^k \right\} \\
&= \inf\left\{ \sum_i \Gamma_{\mathcal{P}^k}(|x_i\rangle)\Gamma_{\mathcal{P}^k}(|y_i\rangle) \;\middle|\; \rho = \sum_i |x_i\rangle\langle y_i| \right\} \\
\Gamma^\circ_{\mathcal{M}^k}(\rho) &= \sup\left\{ |\langle a|\rho|b\rangle| \;\middle|\; |a\rangle, |b\rangle \in \mathcal{P}^k \right\}
\end{aligned} \tag{6.21}$$

and analogously for the other definitions. The computation of these quantities is of course not easy, especially without being able to rely on tools such as the bipartite Schmidt decomposition, although let us remark that generalisations of the Schmidt decomposition have been proposed [165–167] and one of them in particular has been related to the quantification of the $\Gamma_{\mathcal{M}^1}$ norm for some special cases of states [167].

[165] Carteret et al., JMP 2000
[166] Acín et al., PRL 2000
[167] Sokoli & Alber, JPA 2014

A case of particular interest is the so-called *genuine* n-partite entanglement, corresponding to states which are not separable across any bipartition; in other words, the resource theory whose free states are constituted by the $(n-1)$-producible states $\mathcal{M}_+^{n-1}$. This case is particularly simple to consider for pure states, where it suffices to minimise the bipartite atomic gauge over all possible $2^{n-1}-1$ bipartitions of the system:

$$\Gamma_{\mathcal{P}^{n-1}}(|\psi\rangle) = \min_{\substack{\text{bipartition} \\ (A|B)}} \Gamma_{\mathcal{V}^1}^{(A|B)}(|\psi\rangle) \tag{6.22}$$

with $\Gamma_{\mathcal{V}^1}^{(A|B)}$ denoting the bipartite atomic gauge $\Gamma_{\mathcal{V}^1}$ introduced in Sec. 6.2 computed in the given bipartition $(A|B)$. Since we have shown $\Gamma_{\mathcal{V}^1}^2 - 1$ to be equal to twice the bipartite negativity, it means that $\Gamma_{\mathcal{P}^{n-1}}^2 - 1$ is in fact equal to twice the so-called renormalised genuine multiparticle negativity (RGMN) [159]. The convex roof extension of $\Gamma_{\mathcal{P}^{n-1}}^2 - 1$ then generalises the convex roof–extended negativity to a measure of genuine multipartite entanglement.

[159] Hofmann et al., JPA 2014

The RGMN itself, computable with a semidefinite program [159], can be used to provide lower bounds for the measures. Furthermore, we have that the quantifiers such as the generalised robustness of genuine multipartite entanglement are equal to twice the RGMN on pure states, generalising the known relation of the negativity and robustness for $n = 2$ [92,101] and allowing for an efficient quantification of genuine multipartite entanglement for pure states.

[92] Vidal & Werner, PRA 2002
[101] Steiner, PRA 2003

The properties of the quantities introduced above such as faithfulness, strong monotonicity under free operations (which includes LOCC as a subset), quantitative bounds, and their relation to witnesses of k-partite entanglement all follow from the results of Ch. 5. We remark that since for each k, $\mathcal{P}^k$ and $\mathcal{M}_+^k$ span the whole space, the considered measures all have full domain and $\Gamma_{\mathcal{P}^k}$, $\Gamma_{\mathcal{M}^k}$, and $\Gamma_{\mathcal{M}_+^k \cup (-\mathcal{M}_+^k)}$ all define valid norms.

[23] Veitch et al., NJP 2014
[83] Howard & Campbell, PRL 2017

6.4 MAGIC STATES

[168] Veitch et al., NJP 2012
[169] Howard et al., Nature 2014
[170] Delfosse et al., PRX 2015
[171] Bravyi et al., PRX 2016
[172] Raussendorf et al., PRA 2017
[173] Ahmadi et al., PRA 2018

As the final example of the application of the atomic gauge formalism, we briefly consider the resource theory of magic states, recently characterised in [23,83,168–173]. Here, the free pure states are the so-called stabiliser states, consisting of the eigenvectors of Heisenberg-Weyl operators, and states outside of their convex hull are called *magic states*. Since there is always a finite number of pure stabiliser states (although it scales superexponentially with the number of involved

subsystems [174]), the set of free density matrices is given by

[174] Gross, JMP 2006

$$\mathcal{W}_+ = \operatorname{conv}\{|v_1\rangle\langle v_1|, \ldots, |v_k\rangle\langle v_k|\} \tag{6.23}$$

and forms a polytope in the space of Hermitian matrices, which means that the quantification of many measures in the atomic gauge formalism will reduce to solving a linear or semidefinite programming problem [83,173]. We will make this explicit by expressing the optimisation problems in simplified forms.

[83] Howard & Campbell, PRL 2017

[173] Ahmadi et al., PRA 2018

To begin, let us define the non-balanced set $\mathcal{T}' = \{|v_1\rangle, \ldots, |v_k\rangle\}$, where $\{|v_i\rangle\}$ are the pure stabiliser states with any complex phase. Since we require the set of free states to be balanced, the set of interest $\mathcal{T} \subset \mathbb{C}^d$ is given by $\mathcal{T}'$ symmetrised around the origin. The atomic gauge function corresponding to this set can be obtained by considering the $d \times k$ matrix whose columns correspond to the free states, i.e. $T = \big(|v_1\rangle, \ldots, |v_k\rangle\big)$, which then gives

$$\begin{aligned}
\Gamma_{\mathcal{T}}(|\psi\rangle) &= \min\left\{ \sum_i c_i \;\middle|\; |\psi\rangle = \sum_i c_i |w_i\rangle,\ |w_i\rangle \in \mathcal{T},\ c_i \in \mathbb{R}_+ \right\} \\
&= \min\left\{ \sum_i |c_i| \;\middle|\; |\psi\rangle = \sum_i c_i |v_i\rangle,\ |v_i\rangle \in \mathcal{T}',\ c_i \in \mathbb{C} \right\} \\
&= \min_{x \in \mathbb{C}^k} \left\{ \|x\|_{\ell_1} \;\middle|\; Tx = |\psi\rangle \right\}
\end{aligned} \tag{6.24}$$

where we do not use the Dirac notation for $x \in \mathbb{C}^k$ to differentiate the two different spaces $\mathbb{C}^k$ and $\mathbb{C}^d$. This quantity can be thought of as a convex relaxation of the stabiliser rank, quantifying the least number of stabiliser states one needs to superpose to express the state $|\psi\rangle$ [171].

[171] Bravyi et al., PRX 2016

Defining the set $\mathcal{W} = \left\{ |w_i\rangle\langle w_j| \;\middle|\; |w_i\rangle, |w_j\rangle \in \mathcal{T} \right\}$, it is straightforward to see from the above characterisation that the nuclear gauge of the set $\mathcal{W}$ is given by

$$\Gamma_{\mathcal{W}}(\rho) = \min_{X \in \mathbb{C}^{k\times k}} \left\{ \|X\|_{\ell_1} \;\middle|\; TXT^\dagger = \rho \right\}. \tag{6.25}$$

By the discussion in Sec. 5.3, this defines a faithful monotone of magic. Since the set $\mathcal{T}$ spans the space $\mathbb{C}^d$, we have that $\Gamma_{\mathcal{W}}$ is a valid norm for all complex matrices. Again in a similar way, we get the atomic gauges corresponding to the robustness measures as

$$2R^{\mathcal{W}_+}_{\mathcal{W}_+}(\rho) + 1 = \Gamma_{\mathcal{W}_+ \cup (-\mathcal{W}_+)}(\rho) = \min_{x \in \mathbb{R}^k} \left\{ \|x\|_{\ell_1} \;\middle|\; T\operatorname{diag}(x)T^\dagger = \rho \right\} \tag{6.26}$$

$$R^{\mathbb{D}}_{\mathcal{W}_+}(\rho) + 1 = \min_{x \in \mathbb{R}^k_+} \left\{ \|x\|_{\ell_1} \;\middle|\; T\operatorname{diag}(x)T^\dagger \succeq \rho \right\}. \tag{6.27}$$

Although no longer a linear program, the generalised robustness has the advantage that, once again, its quantification reduces to computing the gauge $\Gamma_{\mathcal{T}}$ for pure states by Theorem 5.8. An alternative characterisation of the above quantities in terms of generalised Bloch vectors can be obtained by following [83],

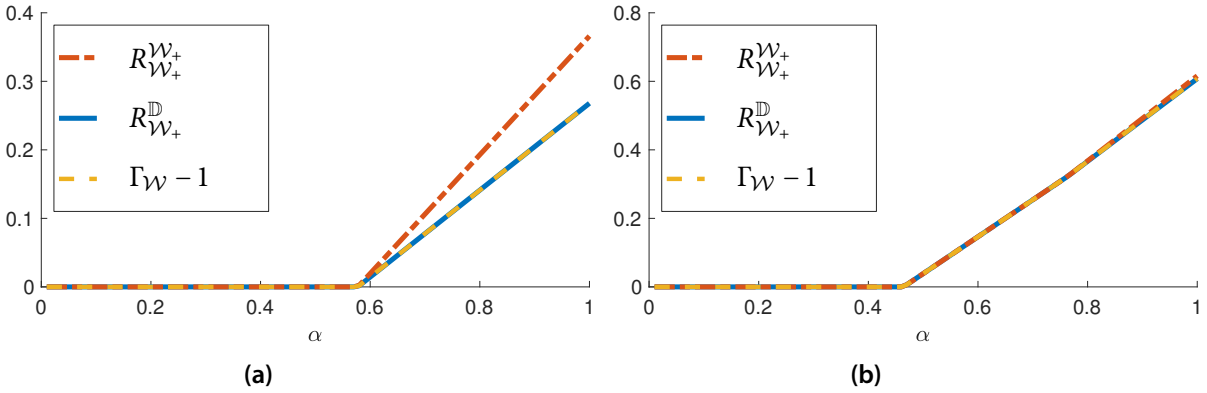


Figure 6.2: Quantitative comparison of the measures of magic for: (a) the one-qubit state $(1-\alpha)\frac{\mathbb{1}}{2} + \alpha\,|T\rangle\langle T|$, (b) the two-qubit state $(1-\alpha)\frac{\mathbb{1}}{4} + \alpha\,|T\otimes T\rangle\langle T\otimes T|$.

where we remark that in [83] the quantity $\Gamma_{\mathcal{W}_+\cup(-\mathcal{W}_+)}$ itself was referred to as the robustness of magic.

[83] Howard & Campbell, PRL 2017

The resource theory of magic states is an example of a resource theory where the standard robustness $R^{\mathcal{W}_+}_{\mathcal{W}_+}$ is in general strictly larger than the other measures on pure states: as an explicit example, consider the one-qubit pure state $|T\rangle\langle T| = \frac{1}{2}\left(\mathbb{1} + \frac{\sigma_x+\sigma_y+\sigma_z}{\sqrt{3}}\right)$ [175] with σ_i being the Pauli matrices. The standard robustness of this state can be computed exactly as $R^{\mathcal{W}_+}_{\mathcal{W}_+}(|T\rangle\langle T|) = \frac{1}{2}(\sqrt{3}-1)$ [83], while a similar calculation for the other measures yields

[175] Bravyi & Kitaev, PRA 2005

$$\Gamma_{\mathcal{T}}(|T\rangle)^2 - 1 = 2 - \sqrt{3} \approx 0.268 < R^{\mathcal{W}_+}_{\mathcal{W}_+}(|T\rangle\langle T|) = \frac{\sqrt{3}-1}{2} \approx 0.366 \qquad (6.28)$$

where we recall that $R^{\mathbb{D}}_{\mathcal{W}_+}(|T\rangle\langle T|) = \Gamma_{\mathcal{W}}(|T\rangle\langle T|) - 1 = \Gamma_{\mathcal{T}}(|T\rangle)^2 - 1$. However, the difference appears to become less pronounced with increasing dimension, with the two-qubit state $|T\otimes T\rangle$ having standard robustness of approx. 0.616 and generalised robustness of approx. 0.607. A comparison of the measures for a class of mixed states is plotted in Fig. 6.2.

We remark that a very similar formalism applies to any resource theory where the set of free states is finite (polytopic). For example, the resource theory of quantum coherence ($k = 1$) can be obtained by taking T to consist of the vectors of the reference orthonormal basis, although care has to be taken about the effective domains of the measures as discussed before.

Part III

Operational characterisation

One-shot distillation of quantum coherence

7

The standard approach to the problem of distillation in quantum information theory, both in the resource theories of entanglement [37–39] and coherence [55,176], is to consider the asymptotic limit — that is, assume that we have access to an unbounded number of independent and identically distributed (i.i.d.) copies of a quantum system. In a realistic setting, however, the resources are finite and the number of i.i.d. prepared states is necessarily limited. More importantly, it is very difficult to perform coherent state manipulations over large numbers of systems. Therefore, it becomes crucial to be able to characterise how well we can distill maximally coherent states from a finite number of copies of the prepared states. The study of such non-asymptotic scenarios has garnered great interest in quantum information theory [177–184], including work in the one-shot theory of entanglement distillation [185–187]. More recently, one-shot results in the resource theory of coherence [188,189] and more general quantum resource theories [6,190] were obtained.

In this chapter, we develop the framework for non-asymptotic coherence distillation, in which one has access only to a single copy of a quantum system and allows for a finite accuracy, reflecting the realistic restrictions on state transformations. In this non-asymptotic setting, the characterisation of resource distillation can be understood as the study of the trade-off between the achievable rate of distillation and the realistic restrictions on state transformations, including the number of accessible i.i.d. copies of a given state as well as the allowed error tolerance. In particular, we establish an exact expression for the one-shot distillable coherence under MIO and DIO, which can be efficiently computed as a semidefinite program. Interestingly, we show that the two quantities are in fact the same, demonstrating that MIO and DIO have the same power in the task of coherence distillation, and together with recent results in coherence dilution [188,191] shedding light on the asymptotic reversibility of state transformations under DIO. Further, we generalise two fundamental quantifiers of coherence, the robustness of coherence [62] and the modified trace distance of coherence [116], establishing a family of measures of coherence which we show to have an operational application in characterizing the maximal fidelity of distillation under different sets of operations. We derive exact expressions for these measures for all pure states, leading to a complete characterisation of pure-state one-shot distillation of coherence and showing that all the considered sets of operations — IO, SIO, DIO, MIO — have exactly the same power in such a task. Furthermore, we develop a general framework of *probabilistic* distillation of quantum coherence, in which the distillation will succeed only with some probability. We investigate

[37] Bennett et al., PRL 1996

[38] Bennett et al., PRA 1996

[39] Rains, PRA 1999

[55] Winter & Yang, PRL 2016

[176] Chitambar et al., PRL 2016

[177] Wang & Renner, PRL 2012

[178] Renes & Renner, IEEE TIT 2011

[179] Tomamichel & Hayashi, IEEE TIT 2013

[180] Berta et al., CMP 2011

[181] Leung & Matthews, IEEE TIT 2015

[182] Datta & Hsieh, IEEE TIT 2013

[183] Wang et al., IEEE TIT 2018

[184] Fang et al., preprint 2017

[185] Brandão & Datta, IEEE TIT 2011

[186] Buscemi & Datta, JMP 2010

[187] Buscemi & Datta, IEEE TIT 2013

[188] Zhao et al., PRL 2018

[189] Bu et al., PRL 2017

[6] Gour, PRA 2017

[190] Gour et al., preprint 2017

[191] Chitambar, PRA 2018

[62] Napoli et al., PRL 2016

[116] Yu et al., PRA 2016

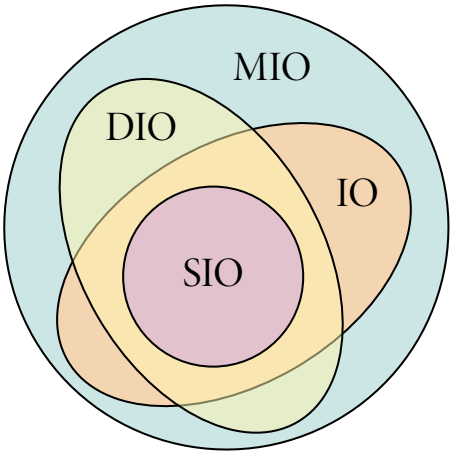


Figure 7.1: The hierarchy of free operations. The diagram presents a schematic representation of the relations between the different sets of operations. Based on [60].

[60] Chitambar & Gour, PRA 2016

probabilistic distillation under different classes of free operations, highlighting differences in their capabilities and establishing their fundamental limitations in state transformations. In particular, we achieve a complete characterisation of probabilistic coherence distillation with pure input states.

This chapter is based on results contained in the works [BR6,BR7,BR8].

[BR6] Regula et al., PRL 2018

[BR7] Fang et al., preprint 2018

[BR8] Regula et al., preprint 2018

7.1 A FAMILY OF COHERENCE QUANTIFIERS

Let us recall the basic structure of the resource theory of quantum coherence. We use Δ to denote the diagonal map (fully dephasing channel) in the basis $\{|i\rangle\}$. We will denote by $\mathcal{I}$ the set of density matrices which are diagonal in this basis, i.e. $\rho \in \mathbb{D}$ such that $\rho = \Delta(\rho)$, and by $\mathcal{I}^{**}$ the cone of diagonal positive semidefinite matrices which are not necessarily normalised. The largest possible set of such free operations are the maximally incoherent operations, which are given by completely positive and trace-preserving maps $\mathcal{E}$ such that $\mathcal{E}(\rho) \in \mathcal{I}$ for every $\rho \in \mathcal{I}$. The incoherent operations are those for which there exists a Kraus decomposition into incoherent Kraus operators, that is, $\{K_\ell\}$ such that $K_\ell \rho K_\ell^\dagger \in \mathcal{I}^{**}$ for all ℓ and all $\rho \in \mathcal{I}$. The strictly incoherent operations are operations for which both $\{K_\ell\}$ and $\{K_\ell^\dagger\}$ are sets of incoherent operators. Finally, the dephasing-covariant incoherent operations are maps $\mathcal{E}$ such that $[\Delta, \mathcal{E}] = 0$, or equivalently $\mathcal{E}(|i\rangle\langle i|) \in \mathcal{I}$ and $\Delta(\mathcal{E}(|i\rangle\langle j|)) = 0$. The inclusions between the sets of operations are visualised in Fig. 7.1.

A natural example of a coherence quantifier is the *trace distance of coherence*, given by [15,126]

$$T_{\mathcal{I}}(\rho) = \min\left\{ \|\rho - \sigma\|_1 \;\middle|\; \sigma \in \mathcal{I} \right\}. \tag{7.1}$$

[15] Baumgratz et al., PRL 2014

[126] Rana et al., PRA 2016

Although the trace distance is a fundamental quantity in quantum information theory [25,26], the trace distance of coherence was found to violate the property of strong monotonicity under incoherent operations [116], which is considered as one of the requirements that a valid measure of coherence should satisfy [15,16]. Therefore, an alternative measure called the *modified trace distance of coherence* satisfying strong monotonicity under IO was proposed [116]:

[25] Nielsen & Chuang, 2011
[26] Wilde, 2017
[116] Yu et al., PRA 2016
[15] Baumgratz et al., PRL 2014
[16] Streltsov et al., RMP 2017

$$\begin{aligned} T'_{\mathcal{I}}(\rho) &= \min\left\{ \|\rho - \lambda\sigma\|_1 \;\middle|\; \sigma \in \mathcal{I},\ \lambda \geq 0 \right\} \\ &= \min\left\{ \|\rho - X\|_1 \;\middle|\; X \in \mathcal{I}^{**} \right\}. \end{aligned} \tag{7.2}$$

We have seen in Sec. 5.4.1 that this is a valid gauge function. Noting that strong duality holds, we can then consider the Lagrange dual of the above expression to obtain a characterisation of the modified trace distance of coherence as

$$T'_{\mathcal{I}}(\rho) = \max\left\{ \langle \rho, W \rangle \;\middle|\; -\mathbb{1} \leq W \leq \mathbb{1},\ \Delta(W) \leq 0 \right\}. \tag{7.3}$$

We can then extend the above to a family of quantifiers given by SDPs of the form

$$T^{(m)}_{\mathcal{I}}(\rho) = \max\left\{ \langle \rho, W \rangle \;\middle|\; -\mathbb{1} \leq W \leq m\mathbb{1},\ \Delta(W) \leq 0 \right\} \tag{7.4}$$

similarly to the approach of Brandão [100] for entanglement measures. Here, we will take $m \in \mathbb{N}$, although it can be treated as a continuous parameter in general. The fact that each such measure is a valid coherence monotone is a simple consequence of the characterisation of $T^{(m)}_{\mathcal{I}}$ as a gauge function, and by a direct application of Theorem 5.21 we have the following.

[100] Brandão, PRA 2005

Proposition 7.1. *For each $m \geq 1$, $T^{(m)}_{\mathcal{I}}$ is a faithful and convex coherence measure satisfying strong monotonicity under MIO.*

Note that $m = 1$ gives $T'_{\mathcal{I}}(\rho)$. For $m = d-1$, one can notice that the constraint $W \leq (d-1)\mathbb{1}$ is redundant: the other constraints ensure that the smallest eigenvalue of W is at least -1 and that the trace of W is at most 0, together implying that there cannot exist an eigenvalue of W which is larger than $d-1$. Therefore, we get

$$\begin{aligned} T^{(d-1)}_{\mathcal{I}}(\rho) &= \max\left\{ \langle \rho, W \rangle \;\middle|\; -\mathbb{1} \leq W,\ \Delta(W) \leq 0 \right\} \\ &= R^{\mathbb{D}}_{\mathcal{I}}(\rho), \end{aligned} \tag{7.5}$$

where $R^{\mathbb{D}}_{\mathcal{I}}(\rho)$ is the robustness of coherence [62,82]. This shows that $T^{(m)}_{\mathcal{I}}$ can be thought of as a family of measures interpolating between the modified trace distance and the robustness of coherence. Notice that we clearly also have $T^{(m)}_{\mathcal{I}}(\rho) = R^{\mathbb{D}}_{\mathcal{I}}(\rho)$ for any $m > d-1$. For $m = 0$, the optimisation is over negative semidefinite matrices, which means that the maximum is reached at $W = 0$ and so $T^{0}_{\mathcal{I}}(\rho) = 0$ for all ρ. In general we have that $0 \leq T^{(m)}_{\mathcal{I}}(\rho) \leq m\ \forall \rho \in \mathbb{D}$. The dual form of the measures can be obtained as follows.

[62] Napoli et al., PRL 2016
[82] Piani et al., PRA 2016

Proposition 7.2. *The measures $T_{\mathcal{S}}^{m}$ can be equivalently expressed as*

$$T_{\mathcal{I}}^{(m)}(\rho) = \min_{X \in \mathcal{I}^{**}} \; m \operatorname{Tr}(\rho - X)_{+} + \operatorname{Tr}(\rho - X)_{-} \,, \tag{7.6}$$

where $(\rho - X)_{+}$ (respectively, $(\rho - X)_{-}$) denotes the positive (negative) part of the Hermitian operator $\rho - X$.

Proof. For any self-adjoint operator X, let $\{X \geq 0\}$ (respectively, $\{X \leq 0\}$) denote the orthogonal projection operator onto the span of the eigenvectors corresponding to non-negative (non-positive) eigenvalues of X. The positive and negative parts of X are then given by $X_{+} = \{X \geq 0\}X$ and $X_{-} = -\{X \leq 0\}X$, such that $X = X_{+} - X_{-}$.

By strong Lagrange duality we have

$$T_{\mathcal{I}}^{m}(\rho) = \min\left\{ m \operatorname{Tr} A + \operatorname{Tr} B \;\middle|\; \rho - X = A - B,\ A, B \geq 0,\ X \in \mathcal{I}^{**} \right\}. \tag{7.7}$$

We will now show that for each feasible X, the optimal value of the optimisation problem

$$\min\left\{ m \operatorname{Tr} A + \operatorname{Tr} B \;\middle|\; \rho - X = A - B,\ A, B \geq 0 \right\} \tag{7.8}$$

is given by $m \operatorname{Tr}(\rho - X)_{+} + \operatorname{Tr}(\rho - X)_{-}$. To see this, note that on one hand we can take $A = (\rho - X)_{+}$ and $B = (\rho - X)_{-}$, and on the other hand by strong Lagrange duality we have

$$\begin{aligned} &\min\left\{ m \operatorname{Tr} A + \operatorname{Tr} B \;\middle|\; \rho - X = A - B,\ A, B \geq 0 \right\} \\ &= \max\left\{ \langle \rho - X, W \rangle \;\middle|\; -\mathbb{1} \leq W \leq m\mathbb{1} \right\} \end{aligned} \tag{7.9}$$

for which $W = m\{\rho \geq X\} - \{\rho \leq X\}$ is a feasible solution. ∎

To characterise the values of the family of quantifiers $T_{\mathcal{I}}^{(m)}$ on pure states, we consider the following quantity:

$$\| |\psi\rangle \|_{[m]} = \min_{|\psi\rangle = |x\rangle + |y\rangle} \| |x\rangle \|_{\ell_1} + \sqrt{m}\, \| |y\rangle \|_{\ell_2} \,, \tag{7.10}$$

which we (for reasons that will become clear later) call the *m-distillation norm*. One can immediately notice from the inequality $\|\cdot\|_{\ell_2} \leq \|\cdot\|_{\ell_1} \leq \sqrt{d}\, \|\cdot\|_{\ell_2}$ that for any normalised pure state we have $\| |\psi\rangle \|_{[1]} = \| |\psi\rangle \|_{\ell_2}$ and $\| |\psi\rangle \|_{[d]} = \| |\psi\rangle \|_{\ell_1}$. Further, we can obtain the following result.

Theorem 7.3. *For any pure state $|\psi\rangle$ and any $m \geq 1$,*

$$T_{\mathcal{I}}^{(m-1)}(\psi) = \| |\psi\rangle \|_{[m]}^{2} - 1. \tag{7.11}$$

Proof. The proof follows by noting that each $T_{\mathcal{I}}^{(m-1)}$ can be viewed as the robustness $R_{\mathcal{Q}_m}^{\mathbb{D}}$, where $\mathcal{Q}_m \coloneqq \mathcal{I} \cup \frac{1}{m}\mathbb{D}$, and applying Theorem 5.8. The details are as follows.

Notice that with the change of variables $W' = W + \mathbb{1}$ we have

$$\begin{aligned} T_{\mathcal{I}}^{(m-1)}(\rho) &= \max\left\{ \langle \rho, W\rangle \;\middle|\; -\mathbb{1} \preceq W \preceq (m-1)\mathbb{1},\ \Delta(W) \preceq 0 \right\} \\ &= \max\left\{ \left\langle \rho, W'\right\rangle \;\middle|\; 0 \preceq W' \preceq m\mathbb{1},\ \Delta(W') \preceq \mathbb{1} \right\} - 1 \\ &= R_{\mathcal{Q}_m}^{\mathbb{D}}(\rho). \end{aligned} \tag{7.12}$$

Let us then define the sets

$$\begin{aligned} \mathcal{B} &\coloneqq \left\{ |\psi\rangle \;\middle|\; |\psi\rangle\langle\psi| \in \mathcal{I} \right\} \\ \mathcal{N}_m &\coloneqq \left\{ |\psi\rangle \;\middle|\; \||\psi\rangle\|_{\ell_2} = \frac{1}{\sqrt{m}} \right\} \\ \mathcal{V}_m &\coloneqq \mathcal{B} \cup \mathcal{N}_m \end{aligned} \tag{7.13}$$

using which we get

$$\mathcal{Q}_m = \left\{ |\psi\rangle\langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{V}_m \right\}. \tag{7.14}$$

Notice that the atomic gauge function of the set $\mathcal{V}_m$ can be given as

$$\begin{aligned} \Gamma_{\mathcal{V}_m}(|\psi\rangle) &= (\Gamma_{\mathcal{B}} \mathbin{\dot{\nabla}} \Gamma_{\mathcal{N}_m})(|\psi\rangle) \\ &= \inf_{|\psi\rangle = |x\rangle + |y\rangle} \Gamma_{\mathcal{B}}(|x\rangle) + \Gamma_{\mathcal{N}_m}(|y\rangle) \\ &= \inf_{|\psi\rangle = |x\rangle + |y\rangle} \||x\rangle\|_{\ell_1} + \sqrt{m}\,\||y\rangle\|_{\ell_2} \\ &= \||\psi\rangle\|_{[m]}\ . \end{aligned} \tag{7.15}$$

The result is then a direct application of Theorem 5.8. ∎

A property of the m-distillation norm which will be crucial in the characterisation of coherence distillation is that it can, in fact, be computed exactly. In particular, the following holds.

Theorem 7.4. *For a pure state $|\psi\rangle$, let $\psi^{\downarrow}_{1:k}$ denote the vector consisting of the k largest (by magnitude) coefficients of $|\psi\rangle$, and analogously let $\psi^{\downarrow}_{k+1:d}$ denote the vector of the $d-k$ smallest coefficients of $|\psi\rangle$, with $\psi^{\downarrow}_{1:0}$ being the zero vector. Then, for any pure state $|\psi\rangle$ and any integer $m \in \{1,\dots,d\}$ we have*

$$\||\psi\rangle\|_{[m]} = \left\|\psi^{\downarrow}_{1:m-k^\star}\right\|_{\ell_1} + \sqrt{k^\star}\left\|\psi^{\downarrow}_{m-k^\star+1:d}\right\|_{\ell_2},$$

where

$$k^\star = \operatorname*{arg\,min}_{1\le k\le m} \frac{1}{k}\left\|\psi^{\downarrow}_{m-k+1:d}\right\|^2_{\ell_2}. \tag{7.16}$$

Proof. Without loss of generality, assume that $|\psi\rangle$ has non-negative coefficients arranged in non-increasing order. Let

$$\begin{aligned} L_1 &:= \left\| \psi^{\downarrow}_{1:m-k^\star} \right\|_{\ell_1} \\ L_2 &:= \left\| \psi^{\downarrow}_{m-k^\star+1:d} \right\|_{\ell_2} \end{aligned} \tag{7.17}$$

and take the feasible solutions

$$\begin{aligned} |x\rangle &= \left(\psi_1 - \frac{L_2}{\sqrt{k^\star}}, \ldots, \psi_{m-k^\star} - \frac{L_2}{\sqrt{k^\star}}, 0, \ldots, 0 \right)^T \\ |y\rangle &= \left(\underbrace{\frac{L_2}{\sqrt{k^\star}}, \ldots, \frac{L_2}{\sqrt{k^\star}}}_{m-k^\star \text{ times}}, \psi_{m-k^\star+1}, \ldots, \psi_d \right)^T . \end{aligned} \tag{7.18}$$

Since $|\psi\rangle = |x\rangle + |y\rangle$, we have

$$\begin{aligned} \| |\psi\rangle \|_{[m]} &\leq \| |x\rangle \|_{\ell_1} + \sqrt{m} \, \| |y\rangle \|_{\ell_2} \\ &= L_1 - (m-k^\star) \frac{L_2}{\sqrt{k^\star}} + \sqrt{m} \sqrt{(m-k^\star)\frac{L_2^2}{k^\star} + L_2^2} \\ &= L_1 - (m-k^\star) \frac{L_2}{\sqrt{k^\star}} + \sqrt{\frac{m(m-k^\star)}{k^\star} L_2^2 + m L_2^2} \\ &= L_1 - (m-k^\star) \frac{L_2}{\sqrt{k^\star}} + \frac{m}{\sqrt{k^\star}} L_2 \\ &= L_1 + \sqrt{k^\star} L_2 . \end{aligned} \tag{7.19}$$

Now, by strong Lagrange duality we have

$$\| |\psi\rangle \|_{[m]} = \max \left\{ |\langle \psi | x \rangle| \;\middle|\; \| |x\rangle \|_{\ell_\infty} \leq 1, \; \| |x\rangle \|_{\ell_2} \leq \sqrt{m} \right\} . \tag{7.20}$$

Choosing

$$|x\rangle = \left(\underbrace{1, \ldots, 1}_{m-k^\star \text{ times}} , \frac{\sqrt{k^\star}}{L_2} \psi_{m-k^\star+1}, \ldots, \frac{\sqrt{k^\star}}{L_2} \psi_d \right)^T \tag{7.21}$$

we have that

$$\begin{aligned} \| |x\rangle \|_{\ell_2} &= \sqrt{m - k^\star + \frac{k^\star}{L_2^2} L_2^2} = \sqrt{m} \\ \| |x\rangle \|_{\ell_\infty} &= \max \left\{ 1, \; \frac{\sqrt{k^\star}}{L_2} \psi_{m-k^\star+1} \right\} . \end{aligned} \tag{7.22}$$

To see that $\| |x\rangle \|_{\ell_\infty}$ is upper bounded by 1, consider first the case when $k^\star = 1$. We then have

$$\frac{\sqrt{k^\star}}{L_2} \psi_{m-k^\star+1} = \frac{\psi_m}{\psi_m + \left\| \psi^{\downarrow}_{m+1:d} \right\|_{\ell_2}^2} \leq 1. \tag{7.23}$$

Assume now that $k^\star > 1$ and consider the following chain of equivalent inequalities:

$$\begin{aligned}
\frac{\left\|\psi^{\downarrow}_{m-k^\star+1:d}\right\|^2_{\ell_2}}{k^\star} &\leq \frac{\left\|\psi^{\downarrow}_{m-k^\star+2:d}\right\|^2_{\ell_2}}{k^\star-1} \\
\psi^2_{m-k^\star+1} + \left\|\psi^{\downarrow}_{m-k^\star+2:d}\right\|^2_{\ell_2} &\leq \frac{k^\star}{k^\star-1}\left\|\psi^{\downarrow}_{m-k^\star+2:d}\right\|_{\ell_2} \\
\psi^2_{m-k^\star+1} &\leq \frac{1}{k^\star-1}\left(1-\left\|\psi^{\downarrow}_{1:m-k^\star+1}\right\|^2_{\ell_2}\right) \\
k^\star\psi^2_{m-k^\star+1} &\leq \left(1-\left\|\psi^{\downarrow}_{1:m-k^\star}\right\|^2_{\ell_2}\right) \\
\sqrt{k^\star}\,\psi_{m-k^\star+1} &\leq L_2
\end{aligned} \tag{7.24}$$

where the first line follows by definition of $k^\star$, and in the third and fifth lines we have used the fact that $|\psi\rangle$ is a normalised pure state. This means that

$$\frac{\sqrt{k^\star}}{L_2}\psi_{m-k^\star+1} \leq \frac{L_2}{L_2} = 1 \tag{7.25}$$

and so $|x\rangle$ is indeed a feasible solution to the dual optimisation problem. We then have

$$\begin{aligned}
\||\psi\rangle\|_{[m]} &\geq |\langle\psi|x\rangle| \\
&= L_1 + \sqrt{k^\star}L_2
\end{aligned} \tag{7.26}$$

which completes the proof. ∎

The above result generalises the recent result of Johnston et al. [192], where an explicit formula for the modified trace distance $T^{(1)}_{\mathcal{I}}$ was obtained for all pure states. Notice in particular that, if all coefficients of $|\psi\rangle$ satisfy $|\psi_i| \leq \frac{1}{\sqrt{m}}$, then $k^\star = m$ is optimal and we have $\||\psi\rangle\|_{[m]} = \sqrt{m}$, which means that $T^{(m-1)}_{\mathcal{I}}(\psi)$ reaches its maximum value $m-1$. As a consequence, $T^{(m-1)}_{\mathcal{I}}$ does not, in general, admit a unique maximiser in the form of the maximally coherent state, which in [192] was considered as a possible indication that this quantity is not suitable as a coherence quantifier. In the following, however, we will instead demonstrate its operational usefulness in the characterisation of the fidelity of one-shot coherence distillation.

[192] Johnston et al., preprint 2017

7.2 ONE-SHOT COHERENCE DISTILLATION

We will denote by $\Psi_m = |\Psi_m\rangle\langle\Psi_m|$ the m-dimensional maximally coherent state $|\Psi_m\rangle = \sum_{i=1}^m \frac{1}{\sqrt{m}}|i\rangle$ in the reference basis. Recall that the *distillable coherence* $C^\infty_{d,\mathrm{IO}}(\rho)$ is the asymptotic rate at which Ψ_2 can be obtained per copy of a given state ρ via incoherent operations. Winter and Yang [55] showed that the distillable coherence of an arbitrary mixed state coincides with the *relative entropy*

[55] Winter & Yang, PRL 2016

of coherence $C_r(\rho) := \min_{\sigma\in\mathcal{I}} D(\rho\|\sigma) = S(\Delta(\rho)) - S(\rho)$ introduced in [14], where the quantum relative entropy is given as $D(\rho\|\sigma) := \operatorname{Tr}\rho(\log\rho - \log\sigma)$ For any state ρ, the distillable coherence is then given by $C^{\infty}_{d,\mathrm{IO}}(\rho) = C_r(\rho)$.

[14] Aberg, preprint 2006

We now consider the non-asymptotic setting. For any quantum state ρ, the fidelity of coherence distillation under the class of operations $\mathcal{O}$ is defined by

$$F_{\mathcal{O}}(\rho, m) := \max_{\Lambda\in\mathcal{O}} \langle \Lambda(\rho), \Psi_m \rangle . \tag{7.27}$$

The one-shot ε-error distillable coherence is then defined as the maximum over all distillation rates achievable under the given class of operations with an error tolerance of $\varepsilon \in [0,1)$, that is,

$$C^{(1),\varepsilon}_{d,\mathcal{O}}(\rho) := \log \max\left\{ m \in \mathbb{N} \;\middle|\; F_{\mathcal{O}}(\rho, m) \geq 1-\varepsilon \right\}. \tag{7.28}$$

As a consequence, the asymptotic distillable coherence can be given as

$$C^{\infty}_{d,\mathcal{O}}(\rho) = \lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{1}{n} C^{(1),\varepsilon}_{d,\mathcal{O}}(\rho^{\otimes n}). \tag{7.29}$$

One of our main results is that the one-shot distillable coherence can be computed exactly. Using the shorthand notation

$$\lfloor x \rfloor_{\log} := \log \lfloor 2^x \rfloor , \tag{7.30}$$

we can obtain the following result.

Theorem 7.5. *If* $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}\}$*, then for any state* $\rho \in \mathbb{D}$*, the fidelity of coherence distillation and one-shot* ε*-error distillable coherence can both be written as the following semidefinite programs:*

$$F_{\mathcal{O}}(\rho, m) = \max\left\{ \langle G, \rho\rangle \;\middle|\; 0 \leq G \leq \mathbb{1},\ \Delta(G) = \frac{1}{m}\mathbb{1} \right\}, \tag{7.31}$$

$$C^{(1),\varepsilon}_{d,\mathcal{O}}(\rho) = \left\lfloor -\log\min\left\{ m \in \mathbb{R} \;\middle|\; \langle G, \rho\rangle \geq 1-\varepsilon,\ 0 \leq G \leq \mathbb{1},\ \Delta(G) = m\mathbb{1} \right\} \right\rfloor_{\log}$$

Proof. Consider first the class MIO. Denote J as the Choi matrix of operation Λ. Then

$$\langle \Lambda(\rho), \Psi_m \rangle = \left\langle \operatorname{Tr}_A \left(J[\rho^T \otimes \mathbb{1}_m] \right), \Psi_m \right\rangle = \left\langle \operatorname{Tr}_B \left(J[\mathbb{1}_d \otimes \Psi_m] \right), \rho^T \right\rangle . \tag{7.32}$$

Denote U_π as the unitary that permutes the basis $\{|i\rangle\}$ to $\{|\pi(i)\rangle\}$. Noting that Ψ_m is invariant under all U_π, we can exploit this symmetry and consider only

operations such that $J = \sum \mathbb{1}_d \otimes U_\pi J \mathbb{1}_d \otimes U_\pi^\dagger$. Since any operator invariant under all U_π can be written as $a\mathbb{1}_m + b(\Psi_m - \frac{1}{m}\mathbb{1}_m)$ for some $a, b \in \mathbb{R}$, we can write

$$J = R \otimes \mathbb{1}_m + Q \otimes (\Psi_m - \frac{1}{m}\mathbb{1}_m) \tag{7.33}$$

for some Hermitian operators R, Q. Now, since Λ is a trace-preserving map, we have $\mathrm{Tr}_B J = \mathbb{1}_d$, and thus $R = \frac{1}{m}\mathbb{1}_d$.

The condition that $\Lambda \in \mathrm{MIO}$ then corresponds to $\Lambda(|i\rangle\langle i|) \in \mathcal{I}\ \forall i$. Since $\Lambda(|i\rangle\langle i|) = \frac{1}{m}\mathbb{1} + \langle i|Q|i\rangle \otimes (\Psi_m - \frac{1}{m}\mathbb{1}_m)$, the condition can only be satisfied if $\Delta(Q) = 0$. Thus, the optimal operation has the structure

$$J = Q \otimes (\Psi_m - \frac{1}{m}\mathbb{1}_m) + \mathbb{1}_d \otimes \frac{1}{m}\mathbb{1}_m,\ \Delta(Q) = 0. \tag{7.34}$$

It is then easy to check that $J \geq 0$ if and only if $-\frac{1}{m-1}\mathbb{1}_d \leq Q \leq \mathbb{1}_d$. By direct calculation, $\langle \Lambda(\rho), \Psi_m \rangle = (1 - \frac{1}{m})\left\langle Q, \rho^T \right\rangle + \frac{1}{m}$, which gives

$$F_{\mathrm{MIO}}(\rho, m) = \max\left\{ \left(1 - \frac{1}{m}\right)\langle Q, \rho\rangle + \frac{1}{m} \;\middle|\; -\frac{1}{m-1}\mathbb{1}_d \leq Q \leq \mathbb{1}_d, \Delta(Q) = 0 \right\}. \tag{7.35}$$

Performing the substitution $G \coloneqq \frac{m-1}{m}Q + \frac{1}{m}\mathbb{1}_d$, we get

$$F_{\mathrm{MIO}}(\rho, m) = \max\left\{ \langle G, \rho\rangle \;\middle|\; 0 \leq G \leq \mathbb{1}_d, \Delta(G) = \frac{1}{m}\mathbb{1}_d \right\}. \tag{7.36}$$

The optimal operation is therefore given as

$$\Lambda(\rho) = (\mathrm{Tr}\, G\rho)\Psi_m + (1 - \mathrm{Tr}\, G\rho)\frac{\mathbb{1}_m - \Psi_m}{m-1},\ 0 \leq G \leq \mathbb{1}_d,\ \Delta(G) = \frac{1}{m}\mathbb{1}_d. \tag{7.37}$$

Following the definition of one-shot distillable coherence, we have the SDP for $C^{(1),\varepsilon}_{d,\mathrm{MIO}}(\rho)$.

Now, notice that the optimal operation for MIO in Eq. (7.34) is actually DIO — this can be seen explicitly by noting that $\Lambda(|i\rangle\langle j|) \propto (\Psi_m - \frac{1}{m}\mathbb{1}_m)$, whose diagonal is zero. Thus we have

$$F_{\mathrm{MIO}}(\rho, m) = F_{\mathrm{DIO}}(\rho, m) \tag{7.38}$$

and therefore the rates of one-shot distillation are also equal. ■

The result reveals a fundamental relation between different sets of operations in the resource theory of coherence, showing that MIO and DIO have the same power in the task of coherence distillation. This correspondence is in fact surprising: not only is DIO a strict subset of MIO, it is also known that MIO is strictly more powerful than DIO in state transformations [57,60], that there

[57] Marvian & Spekkens, PRA 2016

[60] Chitambar & Gour, PRA 2016

exist entropic coherence monotones under DIO which are not monotones under MIO [60], and that the two sets can exhibit different operational capabilities in tasks such as coherence dilution [188]. Furthermore, since MIO constitutes the largest class of free operations in the resource theory of coherence, the result is of practical relevance as it shows that using DIO is sufficient to achieve the best rates of distillation achievable under any class of free operations.

The result also has important consequences beyond the one-shot regime, in particular for the *asymptotic* reversibility of state transformations in the resource theory of coherence — that is, the question whether the amount of coherence which can be distilled from a number of copies of a state ρ (distillable coherence C_d^∞) is the same as the amount of coherence needed to prepare the same number of copies (coherence cost C_c^∞) in the asymptotic limit of an arbitrarily large number of i.i.d. copies. It is known that the resource theory of coherence is reversible under MIO [3,188], but irreversible under IO as we have $C^\infty_{d,\mathrm{IO}}(\rho) < C^\infty_{c,\mathrm{IO}}(\rho)$ in general [55]. Recently, it has also been shown that $C^\infty_{c,\mathrm{DIO}}(\rho) = C^\infty_{c,\mathrm{MIO}}(\rho) = C_r(\rho)$ [188,191]. Our result in Thm. 7.5 in particular shows that $C^{(1),\varepsilon}_{d,\mathrm{DIO}}(\rho) = C^{(1),\varepsilon}_{d,\mathrm{MIO}}(\rho)$ and therefore $C^\infty_{d,\mathrm{DIO}}(\rho) = C^\infty_{d,\mathrm{MIO}}(\rho) = C_r(\rho)$, complementing the claims of Ref. [188] and strengthening the asymptotic results of Ref. [191] by showing their applicability even in the one-shot case. The fact that state transformations are indeed reversible under DIO and the maximal set of operations MIO is not necessary for full reversibility contrasts with other resource theories such as entanglement, where the maximal set of operations is the only one known to provide asymptotic reversibility [193–196].

[60] Chitambar & Gour, PRA 2016

[188] Zhao et al., PRL 2018

[3] Brandão & Gour, PRL 2015

[55] Winter & Yang, PRL 2016

[191] Chitambar, PRA 2018

[193] Vidal & Cirac, PRL 2001

[194] Vidal et al., PRL 2002

[195] Brandão & Plenio, Nat Phys 2008

[196] Wang & Duan, PRL 2017

7.2.1 *Zero-error distillation*

In the case of exact distillation, that is $\varepsilon = 0$, the problem can be given as follows.

$$C^{(1),0}_{d,\mathcal{O}}(\rho) = \left\lfloor -\log \min\left\{ m \in \mathbb{R} \mid 0 \leq G \leq \mathbb{1},\ \langle G, \rho\rangle = 1,\ \Delta(G) = m\mathbb{1} \right\} \right\rfloor_{\log}. \tag{7.39}$$

Let us look at the structure of this SDP.

> **Lemma 7.6.** *Any feasible G in the SDP (7.39) has the form $G = \Pi_\rho + H$, where Π_ρ is the projector onto the support of ρ, $0 \leq H \leq \mathbb{1}$, and $\mathrm{supp}(H) \in \ker(\rho)$.*

Proof. Write ρ in its spectral decomposition as $\rho = \sum_{i=1}^r \lambda_i |\psi_i\rangle\langle\psi_i|$. We then have

$$\sum_{i=1}^r \lambda_i = \mathrm{Tr}\,\rho = 1 = \langle G, \rho\rangle = \sum_{i=1}^r \lambda_i \langle G, |\psi_i\rangle\langle\psi_i|\rangle \tag{7.40}$$

and so $\langle\psi_i|G|\psi_i\rangle = 1$ for each $i \in \{1, \ldots, r\}$ since $0 \leq G \leq \mathbb{1}$. The constraint $G \leq \mathbb{1}$ then implies that $\Pi_\rho G \Pi_\rho = \Pi_\rho$ as well as $\Pi_\rho G(\mathbb{1} - \Pi_\rho) = 0$ and $(\mathbb{1} - \Pi_\rho) G \Pi_\rho = 0$, since otherwise one could always find a pure state such that $\langle\phi|G|\phi\rangle > 1$. ∎

This immediately gives us the following fundamental restriction on coherence distillation under all sets of free operations.

Corollary 7.7. *Zero-error coherence distillation from a full-rank state is impossible.*

Proof. For any full-rank state ρ, the only feasible solution is $G = \mathbb{1}$ with $m = 1$. ∎

Corollary 7.8. *The zero-error one-shot distillable coherence can be expressed as*

$$C_{d,\mathcal{O}}^{(1),0}(\rho) = \left\lfloor -\log \min \left\{ m \;\middle|\; \Pi_\rho \preceq G \preceq \mathbb{1},\ \Delta(G) = m\mathbb{1} \right\} \right\rfloor_{\log}. \tag{7.41}$$

Proof. We have seen that any feasible solution of (7.39) satisfies $G \geq \Pi_\rho$, and on the other hand any $G \geq \Pi_\rho$ satisfies $G \geq 0$ and $\langle G, \rho \rangle \geq 1$, so the feasible sets of the two problems are equal. ∎

Let us now define the relaxation

$$\widetilde{C}(\rho) \coloneqq -\log \min \left\{ m \;\middle|\; \Pi_\rho \preceq G \preceq \mathbb{1},\ \Delta(G) \preceq m\mathbb{1} \right\}. \tag{7.42}$$

Noticing that $\Delta(\Pi_\rho) \preceq \Delta(\Pi_\rho + H)$ for any $H \geq 0$, we have that $G = \Pi_\rho$ is optimal, giving

$$\widetilde{C}(\rho) = -\log \left\| \Delta(\Pi_\rho) \right\|_\infty \tag{7.43}$$

which straightforwardly leads to

$$\widetilde{C}(\rho \otimes \omega) = \widetilde{C}(\rho) + \widetilde{C}(\omega) \tag{7.44}$$

for any $\rho, \omega \in \mathbb{D}$. This gives a bound on the asymptotic zero-error distillable coherence:

Proposition 7.9. *If $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}\}$, then*

$$\lim_{n\to\infty} \frac{1}{n} C_{d,\mathcal{O}}^{(1),0}(\rho^{\otimes n}) \leq -\log \left\| \Delta(\Pi_\rho) \right\|_\infty. \tag{7.45}$$

As we will see, this bound is tight for all pure ρ.

7.2.2 *Hypothesis testing and pure-state distillation*

We will now show that the quantities introduced in Theorem 7.5 admit alternative characterisations. In particular, we will express the fidelity of distillation as a measure related to the family $T_{\mathcal{I}}^{(m)}$ introduced before, and the one-shot distillable coherence as a quantum hypothesis testing problem. To do so, we will need

to optimise over a larger set of matrices than the incoherent states $\mathcal{I}$: namely, the set $\mathcal{J} \coloneqq \left\{ X \;\middle|\; \mathrm{Tr}(X) = 1,\ \Delta(X) = X \right\}$ of unit-trace diagonal Hermitian matrices, and analogously the set $\mathcal{J}^{**}$ of unnormalised diagonal matrices. We then define the quantities

$$\begin{aligned} T_{\mathcal{J}}^{(m)}(\rho) &\coloneqq \min_{X \in \mathcal{J}^{**}} \; m \,\mathrm{Tr}\,(\rho - X)_+ + \mathrm{Tr}\,(\rho - X)_- \\ &= \max\left\{ \langle \rho, W \rangle \;\middle|\; -\mathbb{1} \leq W \leq m\mathbb{1},\ \Delta(W) = 0 \right\}, \end{aligned} \tag{7.46}$$

in analogy with the measures $T_{\mathcal{I}}^{(m)}$. Following the proof of Prop. 7.1 one can easily see that $T_{\mathcal{J}}^{(m)}$ are also faithful strong monotones under MIO. Let us now consider the *hypothesis testing relative entropy* D_H^ε [177,179], defined as

[177] Wang & Renner, PRL 2012

[179] Tomamichel & Hayashi, IEEE TIT 2013

$$D_H^\varepsilon(\rho \| \sigma) \coloneqq -\log \min\left\{ \langle M, \sigma \rangle \;\middle|\; 0 \leq M \leq \mathbb{1},\ 1 - \langle M, \rho \rangle \leq \varepsilon \right\} \tag{7.47}$$

In the setting of quantum hypothesis testing, one is interested in distinguishing between two quantum states — ρ and σ — by performing a test measurement $\{M, \mathbb{1} - M\}$ where $0 \leq M \leq \mathbb{1}$. The probability of incorrectly accepting state σ as true (type-I error) is given by $\langle \mathbb{1} - M, \rho \rangle$, and the probability of incorrectly accepting state ρ as true (type-II error) is given by $\langle M, \sigma \rangle$ [197]. The quantity $D_H^\varepsilon(\rho \| \sigma)$ then characterises the minimum type-II error while constraining the type-I error to be no greater than ε. Alternatively, $D_H^\varepsilon(\rho \| \sigma)$ can be viewed as the operator-smoothed version of min-relative entropy [99,186]. We remark that $-\log$ is a convex function on its effective domain $\mathbb{R}_{++}$, and so the extension obtained by setting $-\log(x) = \infty \ \forall x \leq 0$ is a convex function on $\mathbb{R}$.

[197] Hayashi, 2016

[99] Datta, IEEE TIT 2009

[186] Buscemi & Datta, JMP 2010

Proposition 7.10. *The fidelity of coherence distillation and one-shot ε-error distillable coherence under $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}\}$ admit a characterisation as*

$$\begin{aligned} F_{\mathcal{O}}(\rho, m) &= \frac{1}{m}\left(T_{\mathcal{J}}^{(m-1)}(\rho) + 1 \right), \\ C_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) &= \left\lfloor \min_{X \in \mathcal{J}} D_H^\varepsilon(\rho \| X) \right\rfloor_{\log}. \end{aligned} \tag{7.48}$$

Proof. To begin with, note that

$$\inf\left\{ k \;\middle|\; \Delta(W) = k\mathbb{1} \right\} = \sup\left\{ \langle M, \Delta(W) \rangle \;\middle|\; \mathrm{Tr}(M) \leq 1 \right\} \tag{7.49}$$

by strong Lagrange duality. The formula for $F_{\mathcal{O}}(\rho, m)$ is then a simple change

of variables. For the rate of one-shot distillation, we obtain

$$\begin{aligned} C_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) &= \left\lfloor -\log \min_{\substack{\langle \rho, G\rangle \geq 1-\varepsilon \\ 0\leq G\leq \mathbb{1}}} \inf\left\{ k \;\middle|\; \Delta(G) = k\mathbb{1}\right\} \right\rfloor_{\log} \\ &= \left\lfloor -\log \min_{\substack{\langle \rho, G\rangle \geq 1-\varepsilon \\ 0\leq G\leq \mathbb{1}}} \sup_{\mathrm{Tr}(M)\leq 1} \langle M, \Delta(G)\rangle \right\rfloor_{\log} \\ &= \left\lfloor -\log \sup_{\mathrm{Tr}(M)\leq 1} \min_{\substack{\langle \rho, G\rangle \geq 1-\varepsilon \\ 0\leq G\leq \mathbb{1}}} \langle \Delta(M), G\rangle \right\rfloor_{\log} \\ &= \left\lfloor \inf_{\substack{\mathrm{Tr}(M)\leq 1 \\ M=\Delta(M)}} -\log \min_{\substack{\langle \rho, G\rangle \geq 1-\varepsilon \\ 0\leq G\leq \mathbb{1}}} \langle M, G\rangle \right\rfloor_{\log} \\ &= \left\lfloor \inf_{M\in\mathcal{J}^{\circ\circ}} D_H^{\varepsilon}(\rho\|M) \right\rfloor_{\log} \end{aligned} \tag{7.50}$$

where the third equality follows by Sion's minimax theorem and the self-duality of the diagonal map Δ. Notice now that the problem is always feasible (take $G = \mathbb{1}, k = 1$ in the first line) and bounded (take $M = \mathbb{1}/d$ in the last line, showing that the optimal value cannot be ∞), and in particular we can optimise only over $M \in \mathcal{J}$. ■

Although the optimisation in the above problems is over matrices which are not necessarily positive semidefinite, one can show that if one of the problems admits a positive semidefinite optimal solution, then so does the other. This occurs in particular in the case $m = d$; in fact, not only does the fidelity of distillation for $m = d$ simplify to an optimisation over $\mathcal{I}$, but it has previously been shown that $F_{\mathrm{DIO}}(\rho, d) = F_{\mathrm{IO}}(\rho, d) = F_{\mathrm{SIO}}(\rho, d)$ [189], which combined with Prop. 7.10 shows that $F_{\mathcal{O}}(\rho, d)$ is the same for any $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}, \mathrm{SIO}, \mathrm{IO}\}$. However, the case of interest is when we can optimise only over $\mathcal{I}$ for all values of m, as this will allow us to simplify the computation of the quantities. We will now show that this is true for all pure states.

[189] Bu et al., PRL 2017

We first notice that each $T_{\mathcal{I}}^{(m)}$ provides an upper bound on the corresponding $T_{\mathcal{J}}^{(m)}$, giving $F_{\mathrm{MIO}}(\rho, m) \leq \frac{1}{m}\left(T_{\mathcal{I}}^{(m-1)}(\rho) + 1\right)$. To show that this bound is in fact tight for all pure states, we consider different sets of operations — SIO as well as IO. Pure state transformations under IO and SIO are known to be fully characterised by majorisation relations [55,60,198,199], which we can use to obtain the following result.

[55] Winter & Yang, PRL 2016

[60] Chitambar & Gour, PRA 2016

[198] Nielsen, PRL 1999

[199] Zhu et al., PRA 2018

Theorem 7.11. *For any pure state $|\psi\rangle$ and any integer $m \geq 1$*

$$\begin{aligned} F_{\mathrm{SIO}}(\psi, m) &= F_{\mathrm{IO}}(\psi, m) = F_{\mathrm{DIO}}(\psi, m) = F_{\mathrm{MIO}}(\psi, m) \\ &= \frac{1}{m} \||\psi\rangle\|_{[m]}^2 . \end{aligned} \tag{7.51}$$

Proof. Using the fact that SIO $\subset$ DIO $\subset$ MIO, the inclusion $\mathcal{I} \subset \mathcal{J}$, as well as Theorem 7.3, we have

$$F_{\mathrm{SIO}}(\psi, m) \leq F_{\mathrm{DIO}}(\psi, m) \leq F_{\mathrm{MIO}}(\psi, m) \leq \frac{1}{m}(T_{\mathcal{I}}^{(m-1)}(\psi) + 1) = \frac{1}{m} \||\psi\rangle\|_{[m]}^2 . \tag{7.52}$$

We will now show that $F_{\mathrm{SIO}}(\psi, m) \geq \frac{1}{m} \||\psi\rangle\|_{[m]}^2$, thus concluding the proof. To this end, we use the fact that the deterministic transformation from $|\psi\rangle$ to another pure state $|\eta\rangle$ is possible by SIO or IO if and only if the vector of squared coefficients of $|\psi\rangle$ is majorised by the vector of squared coefficients of $|\eta\rangle$ [55,56,198,199], that is,

$$\sum_{i=1}^{k} \psi_i^2 \leq \sum_{i=1}^{k} \eta_i^2 \quad \forall k \in \{1, \ldots, d\} \tag{7.53}$$

where we have assumed without loss of generality that the coefficients of $|\psi\rangle$ and $|\eta\rangle$ are non-negative and arranged in non-increasing order.

Assuming that $m \leq d$ (since otherwise we have $\||\psi\rangle\|_{[m]} = \||\psi\rangle\|_{[d]}$, so we can take $m = d$ and proceed analogously), let $k^\star = \arg\min_{1 \leq k \leq m} \frac{1}{k} \left\| \psi^{\downarrow}_{m-k+1:d} \right\|_{\ell_2}^2$ and $L_2 = \left\| \psi^{\downarrow}_{m-k^\star+1:d} \right\|_{\ell_2}$. Consider then the m-dimensional state

$$|\eta\rangle = \left(\psi_1, \ldots, \psi_{m-k^\star}, \underbrace{\frac{L_2}{\sqrt{k^\star}}, \ldots, \frac{L_2}{\sqrt{k^\star}}}_{k^\star \text{ times}} \right)^T . \tag{7.54}$$

To show that $|\psi\rangle$ can be transformed into $|\eta\rangle$ by IO/SIO, recall from (7.24) that

$$\psi_{m-k^\star+1}^2 \leq \frac{L_2^2}{k^\star}, \tag{7.55}$$

which means that the squared coefficients of $|\eta\rangle$ majorise the squared coefficients of $|\psi\rangle$. This then gives

$$\begin{aligned} F_{\mathrm{SIO}}(\psi, m) &\geq |\langle \Psi_m | \eta \rangle|^2 \\ &= \frac{1}{m} \left(\left\| \psi^{\downarrow}_{1:m-k^\star} \right\|_{\ell_1} + \sqrt{k^\star} \left\| \psi^{\downarrow}_{m-k^\star+1:d} \right\|_{\ell_2} \right)^2 \\ &= \frac{1}{m} \||\psi\rangle\|_{[m]}^2 \end{aligned} \tag{7.56}$$

by Theorem 7.4, as required. ∎

[55] Winter & Yang, PRL 2016

[56] Chitambar & Gour, PRL 2016

[198] Nielsen, PRL 1999

[199] Zhu et al., PRA 2018

This extends the operational equivalence between MIO and DIO in coherence distillation to the strictly smaller set SIO, and has several important consequences. Firstly, it shows that the one-shot distillable coherence of pure states under any of the classes of operations $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}, \mathrm{SIO}, \mathrm{IO}\}$ is exactly the same, and in fact can be expressed as the quantum hypothesis testing problem

$$\begin{aligned} C_{d,\mathcal{O}}^{(1),\varepsilon}(\psi) &= \left\lfloor \min_{\sigma \in \mathcal{I}} D_H^{\varepsilon}(\psi \| \sigma) \right\rfloor_{\log} \\ &= \left\lfloor -\log \min_{\substack{\langle \psi, G \rangle \geq 1-\varepsilon \\ 0 \leq G \leq \mathbb{1}}} \|\Delta(G)\|_{\infty} \right\rfloor_{\log} . \end{aligned} \tag{7.57}$$

Secondly, we can use the properties of the m-distillation norm to obtain exact formulas for the one-shot distillable coherence. In particular, noting that $\||\psi\rangle\|_{[m]} = \sqrt{m}$ (or equivalently $F_{\mathcal{O}}(\psi, m) = 1$) if and only if $\||\psi\rangle\|_{\ell_\infty} \leq \frac{1}{\sqrt{m}}$, we see that the zero-error distillable coherence is given by

$$C_{d,\mathcal{O}}^{(1),0}(\psi) = \log \left\lfloor \||\psi\rangle\|_{\ell_\infty}^{-2} \right\rfloor . \tag{7.58}$$

Relating the m-distillation norm with the fidelity of distillation also allows us to more easily make quantitative statements about the distillability of pure states on average. For example, in Ref. [192] it was shown that the proportion of pure states with respect to the Haar measure for which $\||\psi\rangle\|_{[2]} = 1$ is given by $1 - d\, 2^{1-d}$, which sharply tends to 1 as d increases. In light of our results, this then shows that, with growing dimension, only an exponentially small fraction of pure states are one-shot undistillable, while a significant majority of pure states satisfy $C_{d,\mathcal{O}}^{(1),0}(\psi) \geq 1$ and therefore allow for a zero-error one-shot distillation of at least one bit of coherence.

[192] Johnston et al., preprint 2017

7.3 PROBABILISTIC DISTILLATION OF COHERENCE

Deterministic protocols such as the ones studied in the previous section may not always be the most suitable choice, particularly when the maximal fidelity of distillation achievable in such protocols is not sufficient for the desired applications. In such cases, one might instead opt for *probabilistic* protocols, which could yield a higher amount of coherence, albeit only with some non-unit probability. This framework can be considered as a more general characterisation of distillation in non-asymptotic regimes, where the allowed error is described by two key parameters with practical relevance: the success probability of the one-shot distillation process, and the fidelity between the extracted state and the target state $|\Psi_m\rangle$. To have a systematic understanding of coherence distillation with finite resources and be able to implement practical schemes for this task, it is crucial to describe and optimise the fundamental relations between these two parameters.

The basic task of probabilistic distillation can be understood as follows (see also Fig. 7.2). For any given quantum state ρ held by a single party A, we aim to transform this state to an m-dimensional maximally coherent state (target state) $|\Psi_m\rangle := m^{-1/2} \sum_{i=1}^{m} |i\rangle$ with high fidelity. A single-bit classical flag register L is used to indicate whether the transformation succeeds or not. If the flag is in the 0 state, it tells that the distillation process succeeded and the output state σ has fidelity at least $1 - \varepsilon$ with the target state. Otherwise, the process has failed, and we discard the unwanted output state ω. Since the output state behaves probabilistically, there is an inherent trade-off between the success probability and the transformation fidelity. Our goal is then to maximise the success probability while keeping the transformation infidelity within some tolerance ε.

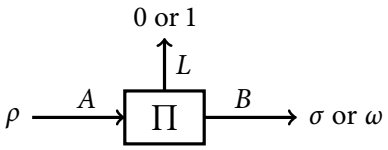


Figure 7.2: General framework of probabilistic coherence distillation with classical register L and quantum registers A, B.

Formally, for any triplet (ρ, m, ε) with a given initial state ρ, target state dimension m, and infidelity tolerance ε, the maximal success probability of coherence distillation under the operation class $\mathcal{O} \in \{\mathrm{SIO}, \mathrm{IO}, \mathrm{DIO}, \mathrm{MIO}\}$ is denoted as $P_{\mathcal{O}}(\rho \to \Psi_m, \varepsilon)$, where $\Psi_m := |\Psi_m\rangle\langle\Psi_m|$. This is given by the maximal value of p such that there exists a transformation $\Pi_{A\to LB} \in \mathcal{O}$ satisfying the constraints

$$\begin{aligned} &\Pi_{A\to LB}(\rho) = p\,|0\rangle\langle 0|_L \otimes \sigma + (1-p)\,|1\rangle\langle 1|_L \otimes \omega, \\ &F(\sigma, \Psi_m) \geq 1 - \varepsilon. \end{aligned} \tag{7.59}$$

If the distillation fails, we can perform a free operation to make the unwanted state ω completely mixed without changing the success probability. Thus, without loss of generality, we can take $\omega = \mathbb{1}/m$. Exploiting the fact that the target state Ψ_m is invariant under the twirling operation $\mathcal{T}(\rho) = \frac{1}{d!} \sum_{i=1}^{d!} P_i \rho P_i$, where P_i are all the permutations on the input system and d is the input dimension, we can also fix the optimal output state as $\sigma = \Psi_m^{\varepsilon}$ where

$$\Psi_m^{\varepsilon} := (1-\varepsilon)\Psi_m + \varepsilon \frac{\mathbb{1} - \Psi_m}{m-1}. \tag{7.60}$$

More precisely, for any optimal output state σ, we can further perform the free operation $\mathcal{T}$, which gives a new output state $\mathcal{T}(\sigma)$ always in the form of $a\Psi_m + b(\mathbb{1} - \Psi_m)/(m-1)$, where we can choose $a = 1 - \varepsilon$ and $b = \varepsilon$ while keeping the fidelity with the target state and the optimal success probability unchanged. This allows us to write

$$P_{\mathcal{O}}(\rho \to \Psi_m, \varepsilon) = P_{\mathcal{O}}(\rho \to \Psi_m^{\varepsilon}, 0), \tag{7.61}$$

meaning that the maximal success probability of coherence distillation is the same as the maximal success probability of transforming the given state to the target Ψ_m^ε with fidelity one.

7.3.1 Computing the maximum distillation probability

We now set out to find efficiently computable expressions for the maximal distillation probability. Consider a generalisation of the set $\mathcal{O}$ to the class $\mathcal{O}_{\text{sub}}$ of subnormalised quantum operations, that is, completely positive and trace-nonincreasing maps. Using this notation, we can conveniently express the maximal success probability as follows (see also [200,201]).

[200] Ishizaka & Plenio, PRA 2005

[201] Buscemi & Gour, PRA 2017

Lemma 7.12. *For any triplet (ρ, m, ε) and operation class $\mathcal{O} \in \{\text{SIO}, \text{IO}, \text{DIO}, \text{MIO}\}$, the maximal success probability $P_{\mathcal{O}}(\rho \to \Psi_m, \varepsilon)$ is given by*

$$\max\left\{ p \in \mathbb{R}_+ \mid \mathcal{E}(\rho) = p \cdot \Psi_m^\varepsilon,\ \mathcal{E} \in \mathcal{O}_{\text{sub}} \right\}. \tag{7.62}$$

Proof. For any quantum operation $\Pi_{A\to LB}$ such that $\Pi_{A\to LB}(\rho) = |0\rangle\langle 0|_L \otimes \mathcal{E}_0(\rho) + |1\rangle\langle 1|_L \otimes \mathcal{E}_1(\rho)$ where $\mathcal{E}_0$ and $\mathcal{E}_1$ are two subnormalised operations, it is not difficult to see that $\Pi_{A\to LB} \in \mathcal{O}$ if and only if $\mathcal{E}_0, \mathcal{E}_1 \in \mathcal{O}_{\text{sub}}$ and $\mathcal{E}_0 + \mathcal{E}_1$ is trace preserving. Thus, finding the optimal solution in the optimisation (7.59) is equivalent to finding the optimal subnormalised operations $\mathcal{E}_0$ and $\mathcal{E}_1$ such that $\mathcal{E}_0(\rho) = p \cdot \Psi_m^\varepsilon$, $\mathcal{E}_1(\rho) = (1-p)\mathbb{1}/m$ and $\mathcal{E}_0 + \mathcal{E}_1$ trace-preserving. Since we can always take $\mathcal{E}_1(\rho) = (\operatorname{Tr}\rho - \operatorname{Tr}\mathcal{E}_0(\rho))\mathbb{1}/m$ without compromising the success probability, the maximal success probability of coherence distillation is only dependent on $\mathcal{E}_0$, and the result follows. ∎

In particular, notice from the above Lemma that

$$P_{\mathcal{O}}(\rho \to \Psi_m, \varepsilon)^{-1} = \Gamma_{\mathcal{S}_\rho}(\Psi_m^\varepsilon) \tag{7.63}$$

where $\mathcal{S}_\rho := \left\{ \mathcal{E}(\rho) \,\middle|\, \mathcal{E} \in \mathcal{O}_{\text{sub}} \right\}$ is the set of all the output operators of ρ under the operation class $\mathcal{O}_{\text{sub}}$. The result simplifies the optimisation of the maximal success probability via subnormalised free operations, providing a geometric interpretation for the maximal success probability as a gauge function. This justifies our intuition that the closer the state ρ to Ψ_m^ε, the less we need to expand the set $\mathcal{S}_\rho$, and thus the larger success probability we can obtain.

By further exploiting the symmetry of Ψ_m^ε, we can compute the maximal success probability under MIO/DIO via the following semidefinite programs.

Theorem 7.13. *For any triplet (ρ, m, ε), the maximal success probability*

of distillation under MIO and DIO are

$$
\begin{aligned}
P_{\mathrm{MIO}}(\rho \to \Psi_m, \varepsilon) = \underset{C,G}{\text{maximise}} \quad & \langle G, \rho \rangle \\
\text{subject to} \quad & \Delta(G) = m\Delta(C), && \text{(7.64 a)} \\
& 0 \preceq C \preceq G \preceq \mathbb{1}, && \text{(7.64 b)} \\
& \langle C, \rho \rangle \geq (1-\varepsilon)\langle G, \rho \rangle && \text{(7.64 c)}
\end{aligned}
$$

$$
\begin{aligned}
P_{\mathrm{DIO}}(\rho \to \Psi_m, \varepsilon) = \underset{C,G}{\text{maximise}} \quad & \langle G, \rho \rangle \\
\text{subject to} \quad & \text{Eqs. (7.64 a,7.64 b,7.64 c)}, \\
& G = \Delta(G) \\
= \underset{C}{\text{maximise}} \quad & m\langle \Delta(C), \rho \rangle \\
\text{subject to} \quad & 0 \preceq C \preceq m\Delta(C) \preceq \mathbb{1}, \\
& \langle C, \rho \rangle \geq m(1-\varepsilon)\langle \Delta(C), \rho \rangle .
\end{aligned}
$$

Proof. The approach is very similar to Theorem 7.5, so we only sketch the full proof. For any optimal subnormalised quantum operation $\mathcal{E}$ in Lemma 7.12, the operation $\widetilde{\mathcal{E}} = \mathcal{T} \circ \mathcal{E}$ is also optimal since Ψ_m^{ε} is invariant under the twirling operation $\mathcal{T}$. Denoting $J_{\mathcal{N}}$ as the Choi matrix of the operation $\mathcal{N}$, we then have $J_{\widetilde{\mathcal{E}}} = C \otimes \Psi_m + D \otimes (\mathbb{1} - \Psi_m)$ for some operators C and D. Taking this form of Choi matrix into the conditions of the optimisation in Lemma 7.12, we obtain the desired forms of the SDPs. ∎

These SDPs provide us with an efficient way to numerically calculate the maximal success probability for general triplets (ρ, m, ε). They also allow us to obtain fundamental results about the capabilities of the different sets of operations for probabilistic coherence distillation.

In this respect, let us also consider the choice of IO or SIO as the free operations. It is known that these two sets of operations have the same power in pure-state transformations, completely characterised by majorisation relations. This yields [56,140,154,202,203]

$$
P_{(\mathrm{S})\mathrm{IO}}(\varphi \to \Psi_m, 0) = \begin{cases} 0 & \text{if } \operatorname{rank}\Delta(\varphi) < m, \\ \min\limits_{k\in[1,m]} \dfrac{m}{k} \displaystyle\sum_{i=m-k+1}^{d} \varphi_i^2 & \text{otherwise,} \end{cases} \tag{7.65}
$$

where we have assumed without loss of generality that the coefficients of φ are non-negative and arranged in non-increasing order. Notice in particular that IO and SIO can never increase the coherence rank of a pure state — it is known that DIO also obeys this restriction, MIO does in fact allow for the rank to increase [56], suggesting that MIO is a much stronger class. It is therefore surpris-

[56] Chitambar & Gour, PRL 2016

[140] Zhu et al., PRA 2017

[154] Vidal, PRL 1999

[202] Lo & Popescu, PRA 2001

[203] Du et al., Quantum Inf Comput 2015

ing that MIO and DIO have exactly the same power in the task of deterministic coherence distillation (as we have shown in Sec. 7.2), and that the two sets of operations lead to the same asymptotic transformation rates for all states [191]. In the following, we will instead explicitly show crucial differences between MIO and DIO when one goes beyond deterministic transformations, highlighting the increased capabilities of MIO in probabilistic distillation, as well as establishing the limitations on coherence distillation in general.

[191] Chitambar, PRA 2018

Theorem 7.14. *For any triplet $(\rho, m, 0)$ with a full-rank state ρ, it holds that $P_{\mathrm{MIO}}(\rho \to \Psi_m, 0) = 0$.*

Proof. Using the notation from Thm. 7.13, we know that $G - C \geq 0$ and $\mathrm{Tr}(G - C)\rho = 0$. Since ρ is full rank, we have $G = C$. Together with $\Delta(G) = m\Delta(C)$, we have $G = C = 0$ and $P_{\mathrm{MIO}}(\rho \to \Psi_m, 0) = 0$. ∎

The results establishes a no-go theorem for coherence distillation, showing that no class of free operations preserving incoherent states can allow to distill any perfect coherence from a full-rank state, even probabilistically. Note that any generic density matrix has full rank, and so does Ψ_m^ε for any $\varepsilon > 0$. Thus

$$|P_{\mathrm{MIO}}(\Psi_m^\varepsilon \to \Psi_m, 0) - P_{\mathrm{MIO}}(\Psi_m \to \Psi_m, 0)| = 1, \tag{7.66}$$

even though Ψ_m^ε can be arbitrarily close to Ψ_m, implying that the maximal success probability is not continuous with respect to the input state. In practical terms, any amount of depolarizing noise will result in a full-rank state, so in a scenario where the coherent state Ψ_m is stored in a quantum memory exposed to depolarizing noise, it is impossible to recover it perfectly using free operations with any non-zero probability.

Theorem 7.15. *For any triplet $(\varphi, m, 0)$ with a coherent pure state $|\varphi\rangle = \sum_{i=1}^n \varphi_i |i\rangle$, $\varphi_i \neq 0$, it holds that $P_{\mathrm{MIO}}(\varphi \to \Psi_m, 0) > 0$.*

Proof. We will first establish the general bound

$$P_{\mathrm{MIO}}(\varphi \to \Psi_m, 0) \geq \frac{n^2}{\sum_{i=1}^n |\varphi_i|^{-2}} \left\| \frac{n-m}{n-1} \widetilde{\varphi} + \frac{n(m-1)}{n-1} \Delta(\widetilde{\varphi}) \right\|_\infty^{-1} \tag{7.67}$$

where

$$|\widetilde{\varphi}\rangle := \frac{1}{\sqrt{s}} \sum_{i=1}^n \frac{\varphi_i}{|\varphi_i|^2} |i\rangle \quad \text{with} \quad s = \sum_{j=1}^n |\varphi_j|^{-2}. \tag{7.68}$$

To this end, let us choose

$$\begin{aligned} C &= c\widetilde{\varphi}, \\ G &= c\widetilde{\varphi} + \frac{(m-1)c}{n-1} \left(n\Delta(\widetilde{\varphi}) - \widetilde{\varphi} \right), \end{aligned} \tag{7.69}$$

where

$$c = \left\| \frac{n-m}{n-1}\widetilde{\varphi} + \frac{n(m-1)}{n-1}\Delta(\widetilde{\varphi}) \right\|_\infty^{-1}. \tag{7.70}$$

We check the constraints in (7.64) one by one. The first condition (7.64 a) trivially holds by the construction. The last condition (7.64 c) holds since $\langle\varphi|n\Delta(\widetilde{\varphi}) - \widetilde{\varphi}|\varphi\rangle = 0$, which implies that $\langle\varphi|C|\varphi\rangle = \langle\varphi|G|\varphi\rangle$. We now move on to the second condition in Eq. (7.64 b). Clearly $C \geq 0$ and furthermore $G \geq C$ as follows from $\varphi \leq n\Delta(\varphi)$. To show that $G \leq \mathbb{1}$, just observe that

$$\begin{aligned} \|G\|_\infty &= c \left\| \widetilde{\varphi} + \frac{m-1}{n-1}\left(n\Delta(\widetilde{\varphi}) - \widetilde{\varphi}\right) \right\|_\infty \\ &= c \left\| \frac{n-m}{n-1}\widetilde{\varphi} + \frac{n(m-1)}{n-1}\Delta(\widetilde{\varphi}) \right\|_\infty \\ &= 1. \end{aligned} \tag{7.71}$$

Hence, C, G as defined above form a valid ansatz for the semidefinite program (7.64) and

$$P_{\mathrm{MIO}}(\varphi, m, 0) \geq \langle G, \varphi\rangle = \frac{n^2 c}{s}, \tag{7.72}$$

which yields the lower bound in (7.67).

Finally, we will show that

$$P_{\mathrm{MIO}}(\varphi \to \Psi_m, 0) \geq \frac{n^2}{m\left(\sum_{i=1}^n |\varphi_i|^{-2}\right)} > 0. \tag{7.73}$$

To this end, it suffices to show that $c \geq 1/m$, i.e. that $c^{-1} \leq m$. This can be done using the triangle inequality:

$$\begin{aligned} c^{-1} &= \left\| \widetilde{\varphi} + \frac{m-1}{n-1}\left(n\Delta(\widetilde{\varphi}) - \widetilde{\varphi}\right) \right\|_\infty \\ &\leq \|\widetilde{\varphi}\|_\infty + (m-1)\left\| \frac{n\Delta(\widetilde{\varphi}) - \widetilde{\varphi}}{n-1} \right\|_\infty \\ &\leq 1 + (m-1) \\ &= m, \end{aligned} \tag{7.74}$$

where we have used the fact that $\frac{n\Delta(\widetilde{\varphi})-\widetilde{\varphi}}{n-1}$ is a valid density matrix. ∎

In contrast to Thm. 7.14, the above result shows that it is always possible to probabilistically distill a maximally coherent state of arbitrary dimension via MIO from any pure state. In particular, the bound in Eq. (7.67) shows that

$$P_{\mathrm{MIO}}(\Psi_n \to \Psi_m, 0) \geq \begin{cases} 1 & \text{if } m \leq n, \\ \frac{n-1}{m-1} & \text{if } m > n. \end{cases} \tag{7.75}$$

Observe that instead $P_{\text{DIO}}(\Psi_n \to \Psi_m, 0) = 0$ for $m > n$. This tells us that, as the dimension n increases, there are n-dimensional density matrices ρ_n such that $P_{\text{MIO}}(\rho_n \to \Psi_{n+1}, 0) \to 1$ while $P_{\text{DIO}}(\rho_n \to \Psi_{n+1}, 0) = 0$ for all n, i.e., P_{MIO} and P_{DIO} can exhibit an arbitrarily large gap. This shows that in the probabilistic distillation scenario, MIO can be much more powerful than DIO in general, in a stark contrast with the case of deterministic coherence distillation (Sec. 7.2).

The relations between the capabilities of DIO and other classes of operations is made precise by the following result, characterizing the fundamental problem of distilling coherence from pure input states.

Theorem 7.16. *For any pure state φ and any m, we have*

$$P_{\text{DIO}}(\varphi \to \Psi_m, 0) = P_{\text{(S)IO}}(\varphi \to \Psi_m, 0). \tag{7.76}$$

Due to the length and technicality of the proof, we defer it to Appendix A.

We can therefore see that DIO does not provide any operational advantage over SIO and IO in pure-state probabilistic coherence distillation. The result is not a priori obvious, since the class DIO is known to be strictly larger than SIO, with fundamental quantitative differences between the two types of operations [56,108]. Putting together the results of Theorems 7.14 and 7.16, we have shown that a large gap in the operational capabilities of operations in the one-shot resource theory of quantum coherence exists between MIO and DIO, but not between DIO and SIO/IO — this can be compared with the case of deterministic distillation, where we have shown that all sets of operations $\mathcal{O} \in \{\text{SIO}, \text{IO}, \text{DIO}, \text{MIO}\}$ allow for the same achievable rate of distillation from pure states.

[56] Chitambar & Gour, PRL 2016

[108] Bu & Xiong, QIC 2017

In the task of distilling maximally coherent qubit states Ψ_2, we can extend the above result and characterise analytically the maximal probability of distillation under DIO and MIO with an arbitrary infidelity ε. In this particular case, MIO does not provide any advantage over DIO.

Proposition 7.17. *For $\mathcal{O} \in \{DIO, MIO\}$ and any pure state φ with $\varphi_1 \geq \ldots \geq \varphi_n > 0$, it holds that*

$$P_{\mathcal{O}}(\varphi \to \Psi_2, \varepsilon) = \begin{cases} 1 & \text{if } \varepsilon \geq \varepsilon_0(\varphi_1), \\ 2(1-\varphi_1^2)\left(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon}\right)^2 & \text{otherwise,} \end{cases}$$

$$\text{where } \varepsilon_0(\varphi_1) := \begin{cases} 0 & \text{if } \varphi_1 \leq \frac{1}{\sqrt{2}}, \\ \frac{1}{2} - \varphi_1\sqrt{1-\varphi_1^2} & \text{otherwise.} \end{cases}$$

The full proof is presented in Appendix A. Using this analytical result, we can give a concrete example to show that the probabilistic distillation framework

can outperform the deterministic one. Suppose we need to distill a maximally coherent qubit state Ψ_2 from the input state $|\varphi\rangle = (3|0\rangle + |1\rangle)/\sqrt{10}$ with acceptable fidelity at least 0.9. The input state becomes useless in the deterministic scenario, since the maximal fidelity achievable via deterministic protocols is given by 0.8. However, probabilistic operations allow us to achieve the required distillation fidelity with 1/2 success probability, demonstrating an explicit operational advantage of probabilistic distillation. In another scenario, if the acceptable fidelity is 0.8, we can gain higher distillation fidelity by compromising some success probability even though deterministic protocols are sufficient to accomplish the task. Such a setting may be dubbed "gambling with coherence", adapting terminology from [38, 202].

7.3.2 *Relations between the probability and fidelity of distillation*

For any given input state ρ and target state dimension m, the maximal success probability is only dependent on the transformation fidelity. The higher the fidelity we require from our output state, the lower the probability that we will succeed. Intuitively, one would expect that the success probability will smoothly decrease as the fidelity increases; however, we find a phenomenon which breaks this expectation. Specifically, we will now show that the maximal success probability may happen to vanish if the fidelity goes beyond some threshold. This can be thought of as similar in spirit to a phase transition in thermodynamics or to the strong converse theorem in channel coding theory [204–206], which says that the coding success probability goes to zero if the coding rate exceeds the capacity of the channel. Note that this phenomenon will not occur in the case of distillation from pure input states under MIO due to Theorem 7.14. In the following result, we completely characterise this property for pure input states under DIO.

[204] Wolfowitz, 1978

[205] Ogawa & Nagaoka, IEEE TIT 1999

[206] Winter, IEEE TIT 1999

Proposition 7.18. *For any pure state $|\varphi\rangle = \sum_{i=1}^{n} \varphi_i |i\rangle$ with nonzero coefficients φ_i, it holds that*

$$P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \begin{cases} > 0 & \text{if } n \geq m \text{ or if } n < m \text{ and } \varepsilon \geq 1 - \frac{n}{m}, \\ = 0 & \text{if } n < m \text{ and } \varepsilon < 1 - \frac{n}{m}. \end{cases}$$

Proof. For the first argument, if $n \geq m$, we know that

$$P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \geq P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) > 0, \tag{7.77}$$

where the second inequality follows from Eq. (7.65).

Note that if $P_{\mathcal{O}}(\sigma_1 \to \sigma_2, 0) = 1$, then $P_{\mathcal{O}}(\rho \to \sigma_2, 0) \geq P_{\mathcal{O}}(\rho \to \sigma_1, 0)$ since we can first transform ρ to σ_1 perfectly and then to get σ_2. For the second argument, if $\varepsilon \geq 1 - \frac{n}{m}$, we have

$$P_{\mathrm{DIO}}(\Psi_n \to \Psi_m^{\varepsilon}, 0) = P_{\mathrm{DIO}}(\Psi_n \to \Psi_m, \varepsilon) = 1. \tag{7.78}$$

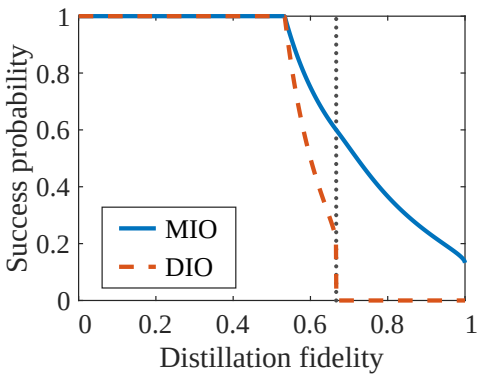


Figure 7.3: An example of the threshold in the trade-off between the fidelity ($F = 1-\varepsilon$) and success probability of distillation (p) for the transformation $(|0\rangle+3|1\rangle)/\sqrt{10} \to \Psi_3$. The auxiliary dotted line $F = 2/3$ is perpendicular to the horizontal axis.

The first equality follows from the fact that $P_{\mathcal{O}}(\rho \to \Psi_m, \varepsilon) = P_{\mathcal{O}}(\rho \to \Psi_m^{\varepsilon}, 0)$. The second equality follows from Lemma 7.19 below. Then

$$\begin{aligned} P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) &= P_{\mathrm{DIO}}(\varphi \to \Psi_m^{\varepsilon}, 0) \\ &\geq P_{\mathrm{DIO}}(\varphi \to \Psi_n, 0) \\ &\geq P_{\mathrm{SIO}}(\varphi \to \Psi_n, 0) \\ &> 0. \end{aligned} \tag{7.79}$$

The first inequality follows from Eq. (7.78). The last inequality follows from Eq. (7.65). If $\varepsilon \leq 1 - \frac{n}{m}$, we have $P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \leq P_{\mathrm{DIO}}(\Psi_n \to \Psi_m, \varepsilon) = 0$, where the second equality follows from Lemma 7.19 below. ■

Lemma 7.19. *For any integer $n \leq m$, it holds that*

$$P_{\mathrm{DIO}}(\Psi_n \to \Psi_m, \varepsilon) = \begin{cases} 1, & \varepsilon \geq 1 - \frac{n}{m}, \\ 0, & \varepsilon < 1 - \frac{n}{m}. \end{cases} \tag{7.80}$$

The result shows that if the input dimension is larger than the target state dimension, there is always a trade-off as expected between the maximal success probability and the transformation fidelity. However, the trade-off curve will always be truncated at the fidelity threshold $\frac{n}{m}$. At the point of $\varepsilon = 1-\frac{n}{m}$, demanding a slightly higher fidelity will make the probabilistic distillation impossible, as shown in Fig. 7.3.

7.4 ASSISTED DISTILLATION OF COHERENCE

The setting of *assisted distillation* is based on a scenario in which the distillation of a resource is aided by another party, typically assumed to hold a purifying quantum state, whose assistance is limited to performing local measurements and communicating their results via a classical channel. Just as assisted entanglement distillation [207,208] found a natural application in tasks involving environment-assisted communication and quantum error correction [208–211], the assisted distillation of coherence can find use in any remote quantum information processing protocol where one aims to increase the coherence available to a spatially separated system without direct access to it [176]. Although this setting has recently garnered both theoretical [212,213] as well as experimental [214,215] attention, assisted coherence distillation has not been characterised in practical scenarios with physically relevant restrictions in mind.

In this section, we establish an operational characterisation of assisted coherence distillation in the one-shot setting based on the results derived in Sec. 7.2. Specifically, we evaluate the one-shot rate of assisted coherence distillation exactly, obtaining several equivalent representations of this quantity. We introduce an efficiently computable semidefinite programming bound on the rate of distillation, and show in particular that the bound is tight for all qubit and qutrit systems. A point of note is that the approach presented here does not rely on the entropy-based methods established for the characterisation of similar tasks in the resource theory of entanglement [185,187], leading to a significantly simpler framework and streamlined proofs.

In the protocol of assisted distillation, we consider two parties (Alice and Bob) who share a pure quantum state $|\psi_{AB}\rangle$, and Alice's task is to assist Bob in distilling coherence from his part of the shared system by performing local measurements on her part of the system and communicating the results to Bob. Since each such measurement (which can, without loss of generality, assumed to be rank one [187]) is in a one-to-one correspondence with a particular decomposition of Bob's system $\rho_B \coloneqq \mathrm{Tr}_A |\psi_{AB}\rangle\langle\psi_{AB}|$ [216,217], this effectively means that Alice's role in the protocol is to allow Bob to access any pure-state decomposition of ρ_B. Therefore, the best achievable assisted distillation rate under a class of operations $\mathcal{O}$ on Bob's system can be expressed as

$$C^{(1),\varepsilon}_{A,\mathcal{O}}(\rho_B) \coloneqq \log \max \left\{ m \in \mathbb{N} \mid F_{A,\mathcal{O}}\left(\rho_B, m\right) \geq 1-\varepsilon \right\} \tag{7.81}$$

with the figure of merit being the average fidelity of distillation [187], defined as

$$F_{A,\mathcal{O}}\left(\rho_B, m\right) \coloneqq \max \left\{ \left\langle \sum_i p_i \, \Lambda_i\left(\psi_i\right), \Psi_m \right\rangle \;\middle|\; \rho_B = \sum_i p_i \psi_i, \; \Lambda_i \in \mathcal{O} \; \forall \, i \right\} \tag{7.82}$$

where the maximisation is over all pure-state decompositions of ρ_B. The definition of $F_{A,\mathcal{O}}$ is motivated by the fact that, after Alice announces her measurement

[207] DiVincenzo et al., 1999
[208] Smolin et al., PRA 2005
[209] Gregoratti & Werner, J. Mod. Opt. 2003
[210] Hayden & King, QIC 2005
[211] Winter, Markov Proc. Rel. Fld. 2007
[176] Chitambar et al., PRL 2016
[212] Chitambar & Hsieh, PRL 2016
[213] Streltsov et al., PRX 2017
[214] Wu et al., Optica 2017
[215] Wu et al., PRL 2018
[185] Brandão & Datta, IEEE TIT 2011
[187] Buscemi & Datta, IEEE TIT 2013
[216] Schrödinger, Math. Pr. Cam. Phil. S. 1936
[217] Hughston et al., PLA 1993

result to Bob, he knows exactly which of the states $\{|\psi_i\rangle\}$ he is in possession of, and can apply an appropriate operation Λ_i best suited for the particular state; since the measurement outcome itself behaves probabilistically, $F_{A,\mathcal{O}}$ is then defined to characterise the best fidelity of assisted distillation achievable on average.

Employing results from the theory of entanglement distillation [208], it has been shown that the asymptotic rate of assisted coherence distillation under incoherent operations is given by the regularised coherence of assistance C_a [176], that is,

[208] Smolin et al., PRA 2005

[176] Chitambar et al., PRL 2016

$$C^{\infty}_{A,\mathrm{IO}}(\rho_B) \coloneqq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}C^{(1),\varepsilon}_{A,\mathrm{IO}}(\rho^{\otimes n}) = \lim_{n\to\infty}\frac{1}{n}C_a(\rho_B^{\otimes n}) \tag{7.83}$$

where

$$C_a(\rho_B) \coloneqq \max\left\{\sum_i p_i\, S\left(\Delta(\psi_i)\right) \;\middle|\; \rho_B = \sum_i p_i\psi_i\right\} \tag{7.84}$$

with the optimisation performed over all pure-state decompositions of ρ_B. Perhaps surprisingly, it turns out that the above regularised coherence of assistance admits a closed formula given by the entropy of the diagonal part of the state in consideration:

$$\lim_{n\to\infty}\frac{1}{n}C_a(\rho_B^{\otimes n}) = S\left(\Delta(\rho_B)\right). \tag{7.85}$$

We stress that the coherence of assistance (7.84) acquires an operational meaning in the task of assisted coherence distillation only in this asymptotic setting — although several authors have referred to C_a as the one-shot equivalent of $C^{\infty}_{A,\mathrm{IO}}$, this is not motivated operationally, and the best achievable rate in the non-asymptotic regime has not been characterised thus far.

In the following, we will omit the subscript B and write ρ for Bob's system, working under the assumption that Alice holds a purifying system.

7.4.1 *Fidelity and rate of distillation*

Notice first that the average fidelity of assisted distillation can be equivalently given as

$$\begin{aligned} F_{A,\mathcal{O}}(\rho, m) &= \max\left\{\sum_i p_i \max_{\Lambda_i\in\mathcal{O}}\langle\Lambda_i(\psi_i), \Psi_m\rangle \;\middle|\; \rho = \sum_i p_i\psi_i\right\} \\ &= \max\left\{\sum_i p_i F_{\mathcal{O}}(\psi_i, m) \;\middle|\; \rho = \sum_i p_i\psi_i\right\} \end{aligned} \tag{7.86}$$

since the diagonal unitary operation adjusting the phases of the given pure state $|\psi_i\rangle$ is always a free operation.

Now, we know from Thm. 7.11 that the fidelity of distillation under any of the sets of operations $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}, \mathrm{SIO}, \mathrm{IO}\}$ is the same, and admits an explicit formula as

$$F_{\mathcal{O}}(\psi, m) = \frac{1}{m}\left\||\psi\rangle\right\|^2_{[m]}. \tag{7.87}$$

Interestingly, this immediately implies that the rate of assisted coherence distillation, one-shot as well asymptotic, will be the same under any of the sets $\mathcal{O} \in \{\mathrm{MIO}, \mathrm{DIO}, \mathrm{SIO}, \mathrm{IO}\}$.

Although $\||\psi\rangle\|_{[m]}$ is exactly computable for any pure state, we will now not make use of this exact expression, and instead use its dual representation:

$$\||\psi\rangle\|_{[m]} = \max\left\{ |\langle\psi|\omega\rangle| \;\middle|\; \||\omega\rangle\|_{\ell_\infty} \leq 1,\ \||\omega\rangle\|_{\ell_2} = \sqrt{m} \right\}. \tag{7.88}$$

Notice that now m can be considered as a continuous parameter, and we will hereafter treat it as such. Consider then a family of sets of density matrices defined by

$$\mathcal{M}_m \coloneqq \mathrm{conv}\left\{ |\omega\rangle\langle\omega| \;\middle|\; \||\omega\rangle\|_{\ell_\infty} \leq \frac{1}{\sqrt{m}},\ \||\omega\rangle\|_{\ell_2} = 1 \right\} \tag{7.89}$$

which allows us to equivalently write

$$F_{\mathcal{O}}(\psi, m) = \max_{\omega \in \mathcal{M}_m} \langle \psi, \omega \rangle = \max_{\omega \in \mathcal{M}_m} F(\psi, \omega). \tag{7.90}$$

Thm. 5.26 tells us that, since $\mathcal{M}_m$ is a convex hull of rank-one projectors, we have

$$\begin{aligned} F_{A,\mathcal{O}}(\rho, m) &= \max\left\{ \sum_i p_i \max_{\omega \in \mathcal{M}_m} F(\psi_i, \omega) \;\middle|\; \rho = \sum_i p_i \psi_i \right\} \\ &= \max_{\omega \in \mathcal{M}_m} F(\rho, \omega), \end{aligned} \tag{7.91}$$

that is, to compute the average fidelity of assisted distillation it suffices to maximise the fidelity of the state ρ over the set $\mathcal{M}_m$, and the optimisation over pure-state decompositions of ρ is not necessary.

Consider now the function defined as

$$\vartheta(\omega) \coloneqq \min\left\{ \max_i \||\psi_i\rangle\|^2_{\ell_\infty} \;\middle|\; \omega = \sum_i p_i \psi_i \right\}. \tag{7.92}$$

The crucial property of ϑ is that, for any $m \geq 1$ and any $\omega \in \mathbb{D}$, we have $\omega \in \mathcal{M}_m \iff \vartheta(\omega) \leq \frac{1}{m}$ by definition of $\mathcal{M}_m$. Letting

$$\mathcal{B}_\varepsilon(\rho) \coloneqq \left\{ \omega \in \mathbb{D} \;\middle|\; F(\rho, \omega) \geq 1 - \varepsilon \right\} \tag{7.93}$$

be the ε-ball of ρ in purified distance, we can then use this function to represent

the one-shot assisted distillation rate as

$$\begin{aligned} C^{(1),\varepsilon}_{A,\mathcal{O}}(\rho) &= \log \max\left\{ m \in \mathbb{N} \;\middle|\; F_{A,\mathcal{O}}(\rho, m) \geq 1-\varepsilon \right\} \\ &= \log \max\left\{ m \in \mathbb{N} \mid F(\rho,\omega) \geq 1-\varepsilon,\ \omega \in \mathcal{M}_m \right\} \\ &= \left\lfloor \log \max\left\{ m \;\middle|\; F(\rho,\omega) \geq 1-\varepsilon,\ \omega \in \mathbb{D},\ \vartheta(\omega) \leq \frac{1}{m} \right\} \right\rfloor_{\log} \\ &= \left\lfloor -\log \min_{\omega \in \mathcal{B}_\varepsilon(\rho)} \left\{ k \;\middle|\; \vartheta(\omega) \leq k \right\} \right\rfloor_{\log} \\ &= \left\lfloor -\log \min_{\omega \in \mathcal{B}_\varepsilon(\rho)} \vartheta(\omega) \right\rfloor_{\log}. \end{aligned} \tag{7.94}$$

The difficulty in evaluating ϑ, however, means that one cannot expect $C^{(1),\varepsilon}_{A,\mathcal{O}}(\rho)$ to be computable in general cases, and prompts our investigation of appropriate relaxations.

Note that a quantifier equivalent to ϑ was used to bound the one-shot assisted distillable entanglement by Buscemi and Datta in Ref. [187], derived there using a complementary set of methods.

[187] Buscemi & Datta, IEEE TIT 2013

We remark that [213] considered also more general settings of asymptotic assisted distillation, for example the scenario where Alice is restricted to performing incoherent local operations, as well as the case where Alice and Bob together can perform a larger class of operations called the separable incoherent operations. Interestingly, it was shown that the asymptotic rate of assisted distillation is the same in all of these cases. Our results extend this analysis by showing that even by allowing Bob access to the maximal set of operations MIO one still cannot improve the rate of assisted distillation. We also generalise this equivalence to the non-asymptotic regime.

[213] Streltsov et al., PRX 2017

7.4.2 SDP relaxation

Consider a relaxation of the set $\mathcal{M}_m$ defined as follows:

$$\widetilde{\mathcal{M}}_m \coloneqq \left\{ \omega \in \mathbb{D} \;\middle|\; \|\Delta(\omega)\|_\infty \leq \frac{1}{m} \right\}. \tag{7.95}$$

The inclusion $\mathcal{M}_m \subseteq \widetilde{\mathcal{M}}_m$ is then obvious, since any $|\psi\rangle\langle\psi| \in \mathcal{M}_m$ is contained in $\widetilde{\mathcal{M}}_m$, and hence also all convex combinations of such rank-one terms. Although it might be tempting to conjecture that $\mathcal{M}_m = \widetilde{\mathcal{M}}_m$, this can be shown to be true only in dimension $d \leq 3$, and in general the inclusion can be strict.

Theorem 7.20. *In dimension $d \in \{2,3\}$ it holds that $\mathcal{M}_m = \widetilde{\mathcal{M}}_m$ for all m, but for all $d \geq 4$ there exist m s.t. $\mathcal{M}_m \subsetneq \widetilde{\mathcal{M}}_m$.*

Proof. A particular simplification occurs in the case $m = d$, where we get

$$\widetilde{\mathcal{M}}_d = \left\{ X \geq 0 \;\middle|\; X_{ii} = \frac{1}{d} \;\forall\, i \right\}. \tag{7.96}$$

This set (up to a multiplicative factor of $\frac{1}{d}$) corresponds to the set of so-called correlation matrices. In particular, it is known that the extremal points of the set $\widetilde{\mathcal{M}}_d$ are given by rank-one matrices only in the case of $d \in \{2, 3\}$, and for any $d \geq 4$, there exist extremal points of $\widetilde{\mathcal{M}}_d$ of rank at least 2 [218–220]. Since no such extremal rank-2 matrix can be written as a convex combination of rank-one matrices in $\widetilde{\mathcal{M}}_d$, we have $\widetilde{\mathcal{M}}_d \neq \mathcal{M}_d$ when $d \geq 4$.

[218] Christensen & Vesterstrøm, Math. Ann. 1979

[219] Grone et al., Lin. Alg. Appl. 1990

[220] Li & Tam, SIAM J. Matrix Anal. 1994

Consider now the case of $d \in \{2, 3\}$ and take $\rho \in \widetilde{\mathcal{M}}_m$ for any $m \geq 1$. We will assume without loss of generality that $\Delta(\rho) > 0$, since otherwise one can apply an inconsequential permutation to write $\rho = \rho' \oplus 0$, reducing the problem to the lower-dimensional case. Define $X \coloneqq \Delta(\rho)^{-1/2}\, \rho\, \Delta(\rho)^{-1/2}$ so that $\Delta(X) = \mathbb{1}$. Since X is a correlation matrix, by the result of [218] discussed above it admits a convex pure-state convex decomposition as $X = \sum_i p_i \xi_i'$ with each $\Delta(\xi_i') = \mathbb{1}$. Defining $|\xi_i\rangle \coloneqq \Delta(\rho)^{1/2} |\xi_i'\rangle$ we get $\rho = \sum_i p_i \xi_i$ with $\Delta(\xi_i) = \Delta(\rho)$ for all i, which means in particular that ρ admits a rank-one convex decomposition in $\widetilde{\mathcal{M}}_m$. This implies that $\widetilde{\mathcal{M}}_m \subseteq \mathcal{M}_m$, concluding the proof that $\widetilde{\mathcal{M}}_m = \mathcal{M}_m$ in dimension 2 and 3. ∎

Notice that the proof of the above Theorem also shows a general property of quantum states:

> **Corollary 7.21.** *Every $\rho \in \mathbb{D}$ in dimension $d \leq 3$ admits a convex decomposition into pure states as $\rho = \sum_i p_i \psi_i$ such that $\Delta(\psi_i) = \Delta(\rho)$ for all i.*

This in particular solves an open question raised in [176] concerning the additivity of the coherence of assistance C_a in dimension $d = 3$; we will address this point in more detail in Sec. 7.4.3.

[176] Chitambar et al., PRL 2016

The most important point about relaxing the set $\mathcal{M}_m$ to $\widetilde{\mathcal{M}}_m$ is that $\widetilde{\mathcal{M}}_m$ can be represented by simple linear matrix inequalities, allowing us to reduce many of the intractable optimisation problems involved in computing the one-shot assisted distillable coherence to efficiently computable semidefinite programs (SDP). In particular, since the fidelity function is known to be computable with an SDP [221,222], we can define the semidefinite program

$$\widetilde{F}_{A,\mathcal{O}}(\rho, m) \coloneqq \max_{\omega \in \widetilde{\mathcal{M}}_m} F(\rho, \omega) \tag{7.97}$$

[221] Watrous, Theory Comput. 2009

[222] Watrous, Chic. J. Th. Comp. Sci. 2013

and the following SDP relaxation of the one-shot assisted distillation rate, which

is tight for all states in dimension 2 and 3:

$$\begin{aligned}\widetilde{C}^{(1),\varepsilon}_{A,\mathcal{O}}(\rho) &\coloneqq \log\max\left\{ m \;\middle|\; \widetilde{F}_{A,\mathcal{O}}(\rho, m) \geq 1-\varepsilon \right\} \\ &= \log\max\left\{ m \;\middle|\; F(\rho,\omega) \geq 1-\varepsilon,\ \omega \in \mathbb{D},\ \|\Delta(\omega)\|_\infty \leq \frac{1}{m} \right\} \\ &= -\log \min_{\omega \in \mathcal{B}_\varepsilon(\rho)} \|\Delta(\omega)\|_\infty .\end{aligned} \tag{7.98}$$

In fact, we can show that the fidelity of assisted distillation for qubit and qutrit systems admits an analytical solution. To this end, let $|\delta(\rho)\rangle$ denote the d-dimensional vector obtained from the square roots of the diagonal elements of ρ, i.e. $\delta(\rho) = \mathrm{diag}(\sqrt{\Delta(\rho)})$. We can then notice the following relation:

Proposition 7.22. *$F_{A,\mathcal{O}}(\rho, m) \leq \frac{1}{m}\|\delta(\rho)\|^2_{[m]}$, with equality if $d \in \{2,3\}$.*

Proof. Let $\mathbb{R}^d_+$ denote the set of all vectors $|x\rangle$ such that $x_i \geq 0$ for all i. We first make note of the fact that

$$\left\{ |\delta(\omega)\rangle \;\middle|\; \omega \in \widetilde{\mathcal{M}}_m \right\} = \left\{ |x\rangle \;\middle|\; \||x\rangle\|_{\ell_\infty} \leq \frac{1}{\sqrt{m}},\ \||x\rangle\|_{\ell_2} \leq 1,\ |x\rangle \in \mathbb{R}^d_+ \right\}. \tag{7.99}$$

This can be seen by noting that every matrix $\omega \in \widetilde{\mathcal{M}}_m$ is a normalised density matrix, which implies the conditions that $\delta(\omega) \in \mathbb{R}^d_+$ as well as $\mathrm{Tr}(\omega) = \||\delta(\omega)\rangle\|^2_{\ell_2} = 1$, and each such ω satisfies $\|\Delta(\omega)\|_\infty = \||\delta(\omega)\rangle\|^2_{\ell_\infty} \leq \frac{1}{m}$. We now use the fact that, for any vector $|y\rangle \in \mathbb{R}^d_+$, to compute the m-distillation norm it suffices to optimise over vectors in $\mathbb{R}^d_+$ (see Thm. 7.4), that is

$$\||y\rangle\|_{[m]} = \max\left\{ \langle y|x\rangle \;\middle|\; \||x\rangle\|_{\ell_\infty} \leq 1,\ \||x\rangle\|_{\ell_2} = \sqrt{m},\ |x\rangle \in \mathbb{R}^d_+ \right\}. \tag{7.100}$$

Using the monotonicity of fidelity under quantum channels, we now have

$$\begin{aligned} F_{A,\mathcal{O}}(\rho, m) &\leq \widetilde{F}_{A,\mathcal{O}}(\rho, m) \\ &= \max_{\omega \in \widetilde{\mathcal{M}}_m} F(\rho,\omega) \\ &\leq \max_{\omega \in \widetilde{\mathcal{M}}_m} F(\Delta(\rho), \Delta(\omega)) \\ &= \max_{\omega \in \widetilde{\mathcal{M}}_m} \left\| \sqrt{\Delta(\rho)}\sqrt{\Delta(\omega)} \right\|_1^2 \\ &= \max_{\omega \in \widetilde{\mathcal{M}}_m} \left(\sum_i \delta(\rho)_i\, \delta(\omega)_i \right)^2 \\ &= \max\left\{ \langle \delta(\rho)|x\rangle^2 \;\middle|\; \||x\rangle\|_{\ell_\infty} \leq \frac{1}{\sqrt{m}},\ \||x\rangle\|_{\ell_2} = 1,\ |x\rangle \in \mathbb{R}^d_+ \right\} \\ &= \frac{1}{m}\|\delta(\rho)\|^2_{[m]} . \end{aligned} \tag{7.101}$$

This establishes the first part of the Proposition. To show achievability, note by Corr. 7.21 that any density matrix in $d \in \{2, 3\}$ admits a pure-state decomposition into $\rho = \sum_i p_i \psi_i$ such that $\Delta(\psi_i) = \Delta(\rho)\ \forall i$. Writing ρ in this decomposition, we then have

$$\begin{aligned} F_{A,\mathcal{O}}(\rho, m) &\geq \sum_i p_i F_{\mathcal{O}}(\psi_i, m) \\ &= \sum_i p_i \left\| |\psi_i\rangle \right\|_{[m]}^2 \\ &= \left\| |\psi_i\rangle \right\|_{[m]}^2 \\ &= \left\| \delta(\rho) \right\|_{[m]}^2 . \end{aligned} \tag{7.102}$$

■

Based on numerical evidence, we can conjecture that $\widetilde{F}_{A,\mathcal{O}}(\rho, m)$ is equal to $\frac{1}{m} \left\| \delta(\rho) \right\|_{[m]}^2$ in any dimension, which would give a clear interpretation to the considered quantity $\left\| \delta(\rho) \right\|_{[m]}$.

7.4.3 *Coherence of assistance*

A characterisation of the protocol of assisted coherence distillation in the limit of infinitely many i.i.d. copies was introduced in Ref. [176], where the relation between the operational quantity $C_{A,\mathrm{IO}}^{\infty}$ and the coherence of assistance was explored. Recall that the coherence of assistance, defined in (7.84), satisfies

[176] Chitambar et al., PRL 2016

$$C_a^{\infty}(\rho) \coloneqq \lim_{n\to\infty} \frac{1}{n} C_a(\rho^{\otimes n}) = C_{A,\mathrm{IO}}^{\infty}(\rho) = S(\Delta(\rho)), \tag{7.103}$$

and the results of our work show that one can replace IO with any set of operations SIO, DIO, MIO.

In [176] it was shown that $C_a(\rho) = C_a^{\infty}(\rho)$ for any qubit state ρ, but the problem of additivity of $C_a(\rho)$ for qutrit systems was left as an open question, later explored also in [223]. Below, we answer it in the affirmative [1].

[223] Zhao et al., PRA 2017

Proposition 7.23. *$C_a(\rho) = C_a^{\infty}(\rho)$ for any state ρ with $d \leq 3$, but in $d \geq 4$ there exist states such that $C_a^{\infty} > C_a(\rho)$.*

Proof. Recall that, in general, $C_a^{\infty}(\rho) \geq C_a(\rho)$. By Corollary 7.21, every state ρ in $d \leq 3$ admits a pure-state decomposition into states with diagonal elements equal to those of ρ. Writing ρ in this decomposition, we have

$$C_a(\rho) \geq \sum_i p_i S(\Delta(\rho)) = S(\Delta(\rho)) = C_a^{\infty}(\rho). \tag{7.104}$$

When $d \geq 4$, from Thm. 7.20 we know that there exist states $\rho \in \mathcal{M}_d$ which cannot be written as a convex combination of pure states which all satisfy

[1] The fact that coherence of assistance is not additive for $d > 3$ has been correctly pointed out in Ref. [176], but the proof provided there has a gap.

$\Delta(\psi_i) = \frac{1}{d}\mathbb{1}$, which in particular means that for every decomposition of such a state we have

$$\sum_i p_i S(\Delta(\psi_i)) < \sum_i p_i S(\Delta(\rho)) = S\left(\frac{1}{d}\mathbb{1}\right) \tag{7.105}$$

since the uniform distribution is the unique maximiser of the Shannon entropy. ∎

We remark that Proposition 7.23 disproves a claim in Ref. [74] that C_a is additive for any d.

[74] Uhlmann, Entropy 2010

8 One-shot entanglement manipulation beyond LOCC

Although the choice of LOCC as the free operations in the resource theory of entanglement is very well-motivated operationally, this class of operations is difficult to characterise in practice. Several relaxations of this class have been considered: most notably the class of separable operations [106], separability-preserving operations [224], as well as PPT operations [225]. The purpose of this chapter is to establish a description of the one-shot distillation of entanglement under such extensions of LOCC, thus providing upper bounds on the capabilities of LOCC and new insights about entanglement manipulation in the one-shot setting. The methods generalise our approach to the distillation of coherence in Ch. 7, showing that equivalent results hold also for the resource theory of entanglement. We will express the relevant operational quantities in the convex geometric framework of gauge functions, allowing us to adapt the methods of this thesis straightforwardly. We in particular show that the quantification of one-shot distillable entanglement reduces to the computation of one-shot distillable coherence for the simplified cases of pure states as well as maximally correlated states.

[106] Bennett et al., PRA 1999

[224] Brandão & Plenio, CMP 2010

[225] Rains, PRA 1999

We begin with the establishment of a family of entanglement quantifiers, akin to the approach taken for coherence.

8.1 A FAMILY OF ENTANGLEMENT QUANTIFIERS

We will consider a bipartite system shared between parties A and B, with d_A and d_B denoting the dimensions of the corresponding spaces. Let $d = \min\{d_A, d_B\}$. For a pure state $|\psi\rangle$, let $\xi(|\psi\rangle)$ denote the vector of the Schmidt coefficients of ψ. We will consider the following sets of Hermitian matrices:

$$\begin{aligned} \mathsf{PPT} &= \left\{ X \;\middle|\; \mathrm{Tr}(X) = 1,\ X^{T_B} \geq 0 \right\} \\ \mathsf{PPT}_+ &= \left\{ X \;\middle|\; \mathrm{Tr}(X) = 1,\ X \geq 0,\ X^{T_B} \geq 0 \right\} \\ \mathsf{SEP} &= \mathrm{conv}\left\{ |\psi\rangle\langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{V} \right\} \end{aligned} \tag{8.1}$$

where X^{T_B} is the partial transpose of X and $\mathcal{V}$ is the set of all pure product states. We have discussed the choices of gauges for the sets SEP and PPT_+ in Ch. 6, but we would now like to consider an operational framework mirroring our considerations for quantum coherence. To this end, letting $\mathcal{S}$ denote one of the

above sets, recall that the generalised robustness is given as

$$\begin{aligned} R_{\mathcal{S}}^{\mathbb{D}}(\rho) &= \min\left\{\lambda \geq 0 \;\middle|\; \rho \leq (\lambda+1)X,\ X \in \mathcal{S}\right\}. \\ &= \max\left\{\langle \rho, W\rangle \;\middle|\; -\mathbb{1} \leq W,\ W \in -\mathcal{S}^{\star}\right\} \end{aligned} \tag{8.2}$$

and extend the definition to the family of measures

$$T_{\mathcal{S}}^{m}(\rho) \coloneqq \max\left\{\langle \rho, W\rangle \;\middle|\; -\mathbb{1} \leq W \leq m\mathbb{1},\ W \in -\mathcal{S}^{\star}\right\}. \tag{8.3}$$

We note this class of measures has also been considered in [100]. We include a complete investigation of the measures below.

[100] Brandão, PRA 2005

First of all, by a direct application of Thm. 5.21 we have the following.

Proposition 8.1. *Let $\mathcal{S} \in \{\mathsf{PPT}, \mathsf{PPT}_+, \mathsf{SEP}\}$, and consider the class of CPTP operations $\mathcal{O}$ such that $\rho \in \mathcal{S} \Rightarrow \Lambda(\rho) \in \mathcal{S}\ \forall \Lambda \in \mathcal{O}$. Then, for each $m \geq 1$, $T_{\mathcal{S}}^{m}$ is faithful with respect to the set $\mathcal{S}$ and strongly monotonic under the operations $\mathcal{O}$.*

The above established in particular that all of the measures are strong monotones under LOCC (we will consider the classes of operations in more detail in the next section). Note, however, that T_{PPT}^{m} and $T_{\mathsf{PPT}_+}^{m}$ are not faithful as entanglement measures, as they are zero for all PPT states.

Proposition 8.2. *The measures $T_{\mathcal{S}}^{m}$ can be equivalently expressed as*

$$T_{\mathcal{S}}^{m}(\rho) = \min\left\{m \operatorname{Tr}(\rho - X)_{+} + \operatorname{Tr}(\rho - X)_{-} \;\middle|\; X \in \mathcal{S}^{\star\star}\right\}. \tag{8.4}$$

Proof. Same as Prop. 7.2. ∎

Proposition 8.3. *For any $\mathcal{S} \in \{\mathsf{PPT}, \mathsf{PPT}_+, \mathsf{SEP}\}$ it holds that*

$$T_{\mathcal{S}}^{(d-1)}(\rho) = R_{\mathcal{S}}^{\mathbb{D}}(\rho). \tag{8.5}$$

Proof. Notice that $\mathsf{SEP} \subset \mathsf{PPT}_+ \subset \mathsf{PPT} \Rightarrow \mathsf{PPT}^{\star} \subset \mathsf{PPT}_+^{\star} \subset \mathsf{SEP}^{\star}$. Letting $\lambda_{\min}$ denote the smallest and $\lambda_{\max}$ the largest eigenvalue of a given matrix, we now use the property that $-W \in \mathsf{SEP}^{\star} \Rightarrow \lambda_{\min}(-W) \geq (1-d)\,\lambda_{\max}(-W)$ [137, 5.5] to obtain

[137] Johnston & Kribs, JMP 2010

$$\lambda_{\max}(W) \leq d - 1. \tag{8.6}$$

It follows that the feasible sets for $R_{\mathcal{S}}^{\mathbb{D}}$ and $T_{\mathcal{S}}^{(d-1)}$ are equal and so the problems are equivalent. ∎

From the above two Propositions, we have that the family of measures $T_{\mathcal{S}}^{(m)}$ interpolates between the modified trace distance $T_{\mathcal{S}}^{(1)}(\rho) = \min_{X \in \mathcal{S}^{\star\star}} \|\rho - X\|_1$ (for $m = 1$) and the robustness of $\mathcal{S}$ (for $m = d - 1$).

Recall the definition of the m-distillation norm as

$$\||x\rangle\|_{[m]} = \inf_{|x\rangle=|y\rangle+|z\rangle} \||y\rangle\|_{\ell_1} + \sqrt{m}\,\||z\rangle\|_{\ell_2}\,. \tag{8.7}$$

We then have the following.

Theorem 8.4. *For a pure state $|\psi\rangle$, let $\xi(|\psi\rangle)$ denote the vector of its Schmidt coefficients. We then have*

$$T_{\mathsf{SEP}}^{(m-1)}(\psi) = \|\xi(|\psi\rangle)\|_{[m]}^2 - 1. \tag{8.8}$$

Proof. To begin, mirroring exactly the proof of Thm. 7.4 we obtain that

$$T_{\mathsf{SEP}}^{(m-1)}(\psi) = \left(\min_{|\psi\rangle=|x\rangle+|y\rangle} \Gamma_{\mathcal{V}}(|x\rangle) + \Gamma_{\mathcal{N}_m}(|y\rangle)\right)^2 - 1 \tag{8.9}$$

where $\mathcal{N}_m = \left\{|\psi\rangle \;\middle|\; \||\psi\rangle\|_{\ell_2} = \frac{1}{\sqrt{m}}\right\}$ and $\mathcal{V}$ is the set of all pure product states. Writing any vector $|x\rangle$ in its Schmidt decomposition as $|x\rangle = \sum_i \eta_i\,|ii\rangle$, we have that $\Gamma_{\mathcal{N}_m}(|x\rangle)^2 = m\sum_i \eta_i^2$ and $\Gamma_{\mathcal{V}}(|x\rangle) = \sum_i \eta_i$ (see our discussion in Ch. 6). By optimising over vectors in the Schmidt basis of $|\psi\rangle$ only, the problem reduces to the m-distillation norm of the Schmidt vector, and we thus have $T_{\mathsf{SEP}}^{(m-1)}(\psi) \leq \|\xi(|\psi\rangle)\|_{[m]}^2 - 1$.

To show the opposite inequality, we use gauge duality to write

$$T_{\mathsf{SEP}}^{(m-1)}(\psi) = \max\left\{|\langle\psi|x\rangle|^2 \;\middle|\; \Gamma^\circ_{\mathcal{N}_m}(|x\rangle) \leq 1,\ \Gamma^\circ_{\mathcal{V}}(|x\rangle) \leq 1\right\} - 1 \tag{8.10}$$

where we have, for $|x\rangle = \sum_i \eta_i\,|ii\rangle$, $\Gamma^\circ_{\mathcal{N}_m}(|x\rangle)^2 = \frac{1}{m}\sum_i \eta_i^2$, and $\Gamma^\circ_{\mathcal{V}}(|x\rangle) = \max_i \eta_i$. By optimising over all vectors in the Schmidt basis of $|\psi\rangle$, the result follows. ∎

Proposition 8.5. *For any state we have*

$$R^{\mathbb{D}}_{\mathsf{PPT}}(\rho) = R^{\mathbb{D}}_{\mathsf{PPT}_+}(\rho). \tag{8.11}$$

Proof. Let $W \in -\mathsf{PPT}_+^\star$ be the optimal dual solution for $R^{\mathbb{D}}_{\mathsf{PPT}_+}$, which means it satisfies $-\mathbb{1} \leq W$ and $W = N + Q^{T_B}$ for $N, Q \leq 0$. But then notice that $W' := W - N \in -\mathsf{PPT}^\star$ is also feasible, and we have $\langle\rho, W'\rangle = \langle\rho, W\rangle - \langle\rho, N\rangle \geq \langle\rho, W\rangle$, so in fact it suffices to optimise over $W \in -\mathsf{PPT}^\star$. ∎

8.1.1 *Generalisation to arbitrary sets*

In the operational characterisation of entanglement distillation, it will be necessary to consider also generalisations of the above measures beyond sets of

normalised (unit trace) Hermitian operators. To allow for this, we will now consider more general sets of Hermitian operators $\mathcal{Q}$ and define the quantity

$$G_{\mathcal{Q}}^{(m)}(\rho) \coloneqq \sup\left\{ \langle \rho, W\rangle \;\middle|\; 0 \leq W \leq \mathbb{1},\ W \in \frac{1}{m}\mathcal{Q}^{\circ} \right\}. \tag{8.12}$$

It is straightforward to see that if $\mathcal{Q} \subseteq \mathbb{H}_1$, then $G_{\mathcal{Q}}^{(m)}(\rho)$ is equal to $\frac{1}{m}\left(T_{\mathcal{Q}}^{(m-1)}(\rho)+1\right)$, although this relation does not hold in general.

Proposition 8.6. *For any compact set $\mathcal{Q} \subseteq \mathbb{H}$, we have that*

$$G_{\mathcal{Q}}^{(m)}(\rho) = \inf_{D\in\mathbb{H}} \operatorname{Tr}(\rho - D)_+ + \frac{1}{m}\Gamma_{\mathcal{Q}}(D). \tag{8.13}$$

Proof. Notice from the definition that

$$G_{\mathcal{Q}}^{(m)} = \Gamma_{(-\mathbb{H}_+)\cup\mathbb{D}\cup m\mathcal{Q}}. \tag{8.14}$$

Using the composition rule for gauge functions, we have

$$\begin{aligned} G_{\mathcal{Q}}^{(m)}(\rho) &= (\Gamma_{(-\mathbb{H}_+)} \dot{\vee} \Gamma_{\mathbb{D}} \dot{\vee} \Gamma_{m\mathcal{Q}})(\rho) \\ &= \inf\left\{ \Gamma_{\mathbb{D}}(Y) + \Gamma_{m\mathcal{Q}}(Z) \;\middle|\; \rho = X + Y + Z,\ X \in -\mathbb{H}_+ \right\} \\ &= \inf\left\{ \operatorname{Tr} Y + \frac{1}{m}\Gamma_{\mathcal{Q}}(Z) \;\middle|\; \rho - Z \leq Y,\ Y \in \mathbb{H}_+ \right\} \\ &= \inf \operatorname{Tr}(\rho - Z)_+ + \frac{1}{m}\Gamma_{\mathcal{Q}}(Z). \end{aligned} \tag{8.15}$$

■

In our consideration of entanglement distillation, we will relate $G_{\mathcal{Q}}^{(m)}$ with another quantity that we previously encountered, the hypothesis testing relative entropy

$$D_H^{\varepsilon}(\rho\|X) = -\log\min\left\{ \langle M, X\rangle \;\middle|\; 0 \leq M \leq \mathbb{1},\ 1 - \langle M, \rho\rangle \leq \varepsilon \right\}. \tag{8.16}$$

Let us first note a general correspondence between the hypothesis testing relative entropy and gauge functions, showing that D_H^{ε} minimised over a set of operators gives a suitably "smoothed" polar gauge.

Theorem 8.7. *Let $\mathcal{Q}$ be a compact set of Hermitian operators. Then*

$$\inf_{X\in\operatorname{conv}\mathcal{Q}} D_H^{\varepsilon}(\rho\|X) = -\log \inf_{\substack{\langle\rho,W\rangle\geq 1-\varepsilon \\ 0\leq W\leq\mathbb{1}}} \Gamma_{\mathcal{Q}}^{\circ}(W). \tag{8.17}$$

Proof. We have

$$\begin{aligned}
-\log \inf_{\substack{\langle \rho, W\rangle \geq 1-\varepsilon \\ 0\leq W\leq \mathbb{1}}} \Gamma^{\circ}_{\mathcal{Q}}(W) &= -\log \inf_{\substack{\langle \rho, W\rangle \geq 1-\varepsilon \\ 0\leq W\leq \mathbb{1}}} \sup_{X\in\mathcal{Q}^{\circ\circ}} \langle X, W\rangle \\
&= -\log \sup_{X\in\mathcal{Q}^{\circ\circ}} \inf_{\substack{\langle \rho, W\rangle \geq 1-\varepsilon \\ 0\leq W\leq \mathbb{1}}} \langle X, W\rangle \\
&= -\log \sup_{X\in\operatorname{conv}\mathcal{Q}} \inf_{\substack{\langle \rho, W\rangle \geq 1-\varepsilon \\ 0\leq W\leq \mathbb{1}}} \langle X, W\rangle \qquad (8.18) \\
&= \inf_{X\in\operatorname{conv}\mathcal{Q}} -\log \inf_{\substack{\langle \rho, W\rangle \geq 1-\varepsilon \\ 0\leq W\leq \mathbb{1}}} \langle X, W\rangle \\
&= \inf_{X\in\operatorname{conv}\mathcal{Q}} D^{\varepsilon}_{H}(\rho\|X)
\end{aligned}$$

where the second equality follows by Sion's minimax theorem, since the sets $\left\{W \;\middle|\; \langle \rho, W\rangle \geq 1-\varepsilon,\ 0 \leq W \leq \mathbb{1}\right\}$ and $\mathcal{Q}^{\circ\circ}$ are both convex and compact. We have replaced the optimisation over $\mathcal{Q}^{\circ\circ}$ with an optimisation over $\operatorname{conv}(\mathcal{Q})$ without loss of generality, since the problem has the same optimal value in both cases — either there exists an $X \in \operatorname{conv}(\mathcal{Q})$ such that $D^{\varepsilon}_{H}(\rho\|X) < \infty$, or we have $D^{\varepsilon}_{H}(\rho\|X) = D^{\varepsilon}_{H}(\rho\|0) = \infty$ for all X. ∎

8.2 ONE-SHOT ENTANGLEMENT DISTILLATION

Denoting by Ψ_m the maximally entangled state $|\Psi_m\rangle = \sum_{i=1}^{m} \frac{1}{\sqrt{m}} |ii\rangle$, we consider the task of distilling copies of Ψ_m under a given class of operations $\mathcal{O}$. The fidelity of distillation under $\mathcal{O}$ is defined by

$$F_{\mathcal{O}}(\rho, m) \coloneqq \max_{\Lambda\in\mathcal{O}} \langle \Lambda(\rho), \Psi_m\rangle . \qquad (8.19)$$

The one-shot ε-error distillable entanglement is then defined as the maximum over all distillation rates achievable under the given class of operations with an error tolerance of ε, that is,

$$E^{(1),\varepsilon}_{d,\mathcal{O}}(\rho) \coloneqq \log \max\left\{ m \in \mathbb{N} \;\middle|\; F_{\mathcal{O}}(\rho, m) \geq 1-\varepsilon \right\}. \qquad (8.20)$$

We then have the asymptotic distillable entanglement as

$$E^{\infty}_{d,\mathcal{O}}(\rho) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} E^{(1),\varepsilon}_{d,\mathcal{O}}(\rho^{\otimes n}). \qquad (8.21)$$

We can now make explicit the connection between the quantifiers discussed earlier and entanglement distillation. Specifically, if one can show that the fidelity of distillation under a given class of operations is given by $G^{(m)}_{\mathcal{Q}}$ for some set $\mathcal{Q}$, then the optimal rate of distillation can be computed as the hypothesis testing entropy minimised over conv $\mathcal{Q}$. Formally, we have the following.

Theorem 8.8. *Let $\mathcal{O}$ be a class of CPTP operations, and $\mathcal{Q}$ a compact set of Hermitian operators. If a given state ρ satisfies*

$$F_{\mathcal{O}}(\rho, m) = G_{\mathcal{Q}}^{(m)}(\rho), \tag{8.22}$$

then

$$E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) = \left\lfloor \min_{X \in \operatorname{conv}\mathcal{Q}} D_H^{\varepsilon}(\rho \| X) \right\rfloor_{\log}. \tag{8.23}$$

Proof. By assumption, we have

$$\begin{aligned} E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) &= \log \max \left\{ m \in \mathbb{N} \;\middle|\; \langle \rho, W \rangle \geq 1-\varepsilon,\ 0 \leq W \leq \mathbb{1},\ W \in \frac{1}{m}\mathcal{Q}^{\circ} \right\} \\ &= \left\lfloor -\log \min \left\{ k \in \mathbb{R} \;\middle|\; \langle \rho, W \rangle \geq 1-\varepsilon,\ 0 \leq W \leq \mathbb{1},\ W \in k\mathcal{Q}^{\circ} \right\} \right\rfloor_{\log} \\ &= \left\lfloor -\log \min_{\substack{\langle \rho, W \rangle \geq 1-\varepsilon \\ 0 \leq W \leq \mathbb{1}}} \Gamma_{\mathcal{Q}^{\circ}}(W) \right\rfloor_{\log} \\ &= \left\lfloor \min_{X \in \operatorname{conv}\mathcal{Q}} D_H^{\varepsilon}(\rho \| X) \right\rfloor_{\log} \end{aligned} \tag{8.24}$$

where the last equality follows from Thm. 8.7. ∎

Although we leave open the question of when exactly a given class of operations leads to a fidelity of distillation of the form given by $G_{\mathcal{Q}}^{(m)}$, we will see below that this is a very common phenomenon among different classes of operations relevant to the resource theory of entanglement.

8.2.1 PPT and PPT-preserving operations

The class of PPT operations, due to Rains [156,225], is defined to consist of all CPTP maps $\Lambda : AB \to A'B'$ whose Choi matrix $J_\Lambda \in \mathcal{L}(ABA'B')$ satisfies $J_\Lambda^{T_{BB'}} \geq 0$. In the literature, PPT operations have sometimes been referred to as "PPT-preserving operations" (see e.g. [226,227]), motivated by the fact that they preserve the positivity of the partial transpose of any Hermitian operator. This name has often caused confusion, as it could suggest that the only requirement placed on this class of operations is such that if $\sigma \in \mathsf{PPT}_+$ then $\Lambda(\sigma) \in \mathsf{PPT}_+$. It is not difficult to see that the definitions are, in fact, inequivalent, and this difference was already pointed out in [228]. An accurate way of referring to the class of PPT operations could be *completely* PPT-preserving [229], since the condition $J_\Lambda^{T_{BB'}} \geq 0$ ensures the preservation of positivity when the map acts on a part of a larger system, akin to completely positive maps.

For clarity, let us begin with the precise definitions.

[156] Rains, IEEE TIT 2001
[225] Rains, PRA 1999
[226] Eggeling et al., PRL 2001
[227] Audenaert et al., PRL 2003
[228] Horodecki, QIC 2001
[229] Matthews & Winter, PRA 2008

Definition 8.9. A CPTP map $\Lambda : AB \to A'B'$ is called PPT if either of the following equivalent conditions is satisfied [225]:

[225] Rains, PRA 1999

(i) The Choi matrix J_Λ is PPT with respect to the partition $AA'|BB'$, i.e. $J_\Lambda^{T_{BB'}} \geq 0$.

(ii) The map $T_{B'} \circ \Lambda \circ T_B$ is completely positive.

We will use PPT to denote the set of all such maps.

A CPTP map $\Lambda : AB \to A'B'$ is called PPT_+-preserving if

$$\sigma \in \mathsf{PPT}_+ \Rightarrow \Lambda(\sigma) \in \mathsf{PPT}_+. \tag{8.25}$$

We will use P-PPT_+ to denote the set of all such maps.

A CPTP map $\Lambda : AB \to A'B'$ is called PPT-preserving if either of the following equivalent conditions is satisfied:

(i) $X \in \mathsf{PPT} \Rightarrow \Lambda(X) \in \mathsf{PPT}$.

(ii) The map $T_{B'} \circ \Lambda \circ T_B$ is positive.

We will use P-PPT to denote the set of all such maps.

We now characterise the operational capabilities of the different sets of operations. Note that the fidelity of distillation under the class PPT has previously been obtained by Rains [156], and an explicit expression for the rate of distillation in terms of D_H^ε appeared more recently in [184]. To our knowledge, the other classes have not been considered before in this context.

[156] Rains, IEEE TIT 2001

[184] Fang et al., preprint 2017

Theorem 8.10. *The one-shot distillable entanglement under the classes of operations* PPT, P-PPT, *and* P-PPT_+ *is given by*

$$\begin{aligned} E_{d,\text{P-PPT}_+}^{(1),\varepsilon}(\rho) &= \left\lfloor \min_{\sigma \in \mathsf{PPT}_+} D_H^\varepsilon(\rho\|\sigma) \right\rfloor_{\log} \\ E_{d,\mathrm{PPT}}^{(1),\varepsilon}(\rho) = E_{d,\text{P-PPT}}^{(1),\varepsilon}(\rho) &= \left\lfloor \min_{X \in \mathsf{PPT}'} D_H^\varepsilon(\rho\|X) \right\rfloor_{\log} \end{aligned} \tag{8.26}$$

where $\mathsf{PPT}' = \mathrm{conv}(\mathsf{PPT} \cup -\mathsf{PPT}) = \left\{ X \;\middle|\; \left\|X^{T_B}\right\|_1 \leq 1 \right\}$.

Proof. Since Ψ_m is invariant under any unitary of the form $U \otimes U^*$, it is in particular invariant under the twirling $\mathcal{T}(\cdot) \coloneqq \int (U \otimes U^*) \cdot (U \otimes U^*)^\dagger \, \mathrm{d}U$, where the integration is performed with respect to the Haar measure of the unitary group. We can then without loss of generality consider only trace-preserving

operations of the form $\Lambda = \mathcal{T} \circ \Lambda$, giving

$$\Lambda(Z) = \langle Z, X\rangle \Psi_m + \frac{\langle Z, \mathbb{1} - X\rangle}{m^2 - 1} (\mathbb{1} - \Psi_m) \tag{8.27}$$

as this is the most general form of an operator invariant under twirling [230]. Note in particular that

[230] Horodecki & Horodecki, PRA 1999

$$\begin{aligned} \Lambda(Z) \geq 0 &\iff \langle Z, X\rangle \geq 0 \\ \Lambda(Z)^{T_B} \geq 0 &\iff -\frac{1}{m} \leq \langle Z, X\rangle \leq \frac{1}{m}. \end{aligned} \tag{8.28}$$

Noting in addition that the complete positivity of Λ imposes $0 \leq X \leq \mathbb{1}$, we can constrain the map Λ such that $\Lambda(\sigma) \in \mathsf{PPT}_+$ for any $\sigma \in \mathsf{PPT}_+$ to get

$$F_{d,\text{P-PPT}_+}(\rho, m) = \max\left\{ \langle \rho, X\rangle \;\middle|\; 0 \leq X \leq \mathbb{1},\ \langle \sigma, X\rangle \leq \frac{1}{m}\ \forall \sigma \in \mathsf{PPT}_+ \right\}, \tag{8.29}$$

and similarly, by imposing that $S \in \mathsf{PPT} \Rightarrow \Lambda(S) \in \mathsf{PPT}$ we have

$$\begin{aligned} F_{d,\text{P-PPT}}(\rho, m) &= \max\left\{ \langle \rho, X\rangle \;\middle|\; 0 \leq X \leq \mathbb{1},\ |\langle S, X\rangle| \leq \frac{1}{m}\ \forall S \in \mathsf{PPT} \right\} \\ &= \max\left\{ \langle \rho, X\rangle \;\middle|\; 0 \leq X \leq \mathbb{1},\ \langle S, X\rangle \leq \frac{1}{m}\ \forall S \in \mathsf{PPT}' \right\}. \end{aligned} \tag{8.30}$$

Noting that the Choi matrix of the map (8.27) is given by $J_\Lambda = X_{AB} \otimes \Psi_{m_{A'B'}} + \frac{1}{m^2-1} (\mathbb{1} - X)_{AB} \otimes (\mathbb{1} - \Psi_m)_{A'B'}$, an explicit computation yields

$$J_\Lambda^{T_{BB'}} \geq 0 \iff -\frac{1}{m}\mathbb{1} \leq X^{T_B} \leq \frac{1}{m}\mathbb{1}, \tag{8.31}$$

which is precisely the condition $\langle S, X\rangle \leq \frac{1}{m}\ \forall S \in \mathsf{PPT}'$, yielding the equality between $F_{d,\text{P-PPT}}(\rho, m)$ and $F_{d,\text{PPT}}(\rho, m)$.

Since the distillation fidelities $F_{d,\text{P-PPT}_+}$ and $F_{d,\text{P-PPT}}$ are precisely of the form G_{PPT_+} and $G_{\mathsf{PPT}'}$, respectively, the result follows by Thm. 8.8. ∎

The Theorem establishes an interesting operational equivalence between the set of operations PPT and P-PPT, while explicitly showing a difference with the channels P-PPT$_+$. In addition, we recall an argument in [231] which showed a gap between PPT and P-PPT$_+$ operations by proving that the negativity (a known monotone under PPT [92]) is not a monotone under P-PPT$_+$. This no longer applies to P-PPT — by definition, the negativity corresponds to the robustness $R^{\mathsf{PPT}}_{\mathsf{PPT}}$ and it follows straightforwardly from Thm. 5.21 that this is a strong monotone under P-PPT. Further monotones under P-PPT$_+$ can be easily established as $R^{\mathsf{PPT}_+}_{\mathsf{PPT}_+}$ or $R^{\mathbb{D}}_{\mathsf{PPT}_+}$, where we recall from Prop. 8.5 that $R^{\mathbb{D}}_{\mathsf{PPT}_+}(\rho) = R^{\mathbb{D}}_{\mathsf{PPT}}(\rho)$. A monotone which lower bounds all of these can additionally be obtained as $R^{\mathbb{D}}_{\mathsf{PPT}'}$, and it has recently found use in providing bounds for distillable of entanglement [157] as well as quantum channel capacities [232,233].

[231] Chitambar et al., preprint 2017

[92] Vidal & Werner, PRA 2002

[157] Wang & Duan, PRA 2016

[232] Wang et al., preprint 2017

[233] Berta & Wilde, NJP 2018

Although it is not easy to characterise the asymptotic rates of distillation under PPT and P-PPT maps, we have the following characterisation of distillable entanglement under P-PPT$_+$, thus establishing a limit on the asymptotic performance of PPT and P-PPT.

Corollary 8.11. *The asymptotic distillable entanglement under* P-PPT$_+$ *is given by the regularised relative entropy with respect to the set* PPT$_+$*,*

$$E_{d,\text{P-PPT}_+}^{\infty}(\rho) = \lim_{n\to\infty} \min_{\sigma\in\text{PPT}_+} \frac{1}{n} D(\rho^{\otimes n}\|\sigma). \tag{8.32}$$

Proof. Follows directly from the generalised quantum Stein's lemma [234]. ■

[234] Brandão & Plenio, CMP 2010

We remark that, although $E_{d,\text{PPT}}^{\infty}(\rho)$ is not known in general, it has been shown that there exist states such that $E_{d,\text{PPT}}^{\infty}(\rho) < E_{d,\text{P-PPT}_+}^{\infty}(\rho)$ [196].

[196] Wang & Duan, PRL 2017

One can then wonder if the operational equivalence between PPT and P-PPT extends beyond the task of entanglement distillation. We will now show that, indeed, it does, and the one-shot entanglement *cost* is also equal under the two sets of operations. Precisely, we define

$$E_{c,\mathcal{O}}^{(1),\varepsilon}(\rho) \coloneqq \log \min_{\Lambda\in\mathcal{O}} \left\{ m\in\mathbb{N} \mid F(\rho, \Lambda(\Psi_m)) \geq 1-\varepsilon \right\} \tag{8.33}$$

and we obtain the following result.

Theorem 8.12. *The one-shot entanglement cost under the classes of operations* PPT*,* P-PPT*, and* P-PPT$_+$ *is given by*

$$\begin{aligned} E_{c,\text{P-PPT}_+}^{(1),\varepsilon}(\rho) &= \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \log\left(R_{\text{PPT}_+}^{\text{PPT}_+}(\widetilde{\rho})+1\right) \right\rceil_{\log} \\ E_{c,\text{PPT}}^{(1),\varepsilon}(\rho) &= E_{c,\text{P-PPT}}^{(1),\varepsilon}(\rho) = \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \Upsilon_{\text{PPT}}(\widetilde{\rho}) \right\rceil_{\log} \end{aligned} \tag{8.34}$$

where

$$\begin{aligned} \Upsilon_{\text{PPT}}(\rho) &\coloneqq \log \min_{\sigma\in\text{PPT}_+} \left\{ \lambda\in\mathbb{R}_+ \;\middle|\; \frac{\rho+(\lambda-1)\sigma}{\lambda} \in \text{PPT},\ \frac{(\lambda+1)\sigma-\rho}{\lambda} \in \text{PPT} \right\} \\ &= \log \min_{\omega\in\mathbb{D}} \left\{ \lambda \;\middle|\; -(\lambda-1)\omega^{T_B} \leq \rho^{T_B} \leq (\lambda+1)\omega^{T_B} \right\}. \end{aligned} \tag{8.35}$$

We defer the proof to the Appendix. The crucial point of this result, together with Thm. 8.10, is that the classes PPT and P-PPT are equivalent in all entanglement manipulation protocols, leading to the same rates of distillation and dilution already in the one-shot setting. This is not a priori obvious, as the definitions allow for a gap between the two operations to exist.

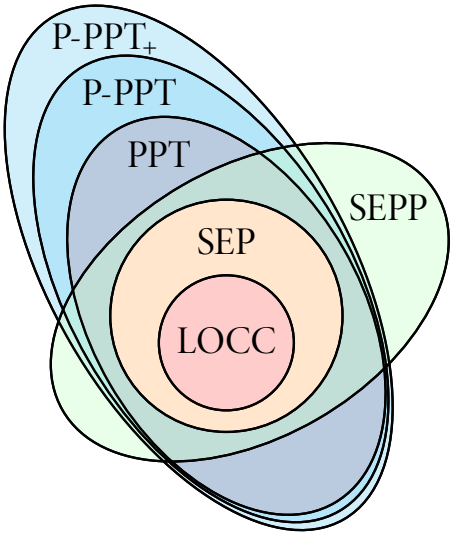


Figure 8.1: The hierarchy of operations considered in this chapter. Each of the considered classes of CPTP maps constitutes a relaxation of the class LOCC.

8.2.2 *Pure-state distillation and separability-preserving operations*

One of the first relaxations of LOCC in the literature was the class of separable operations (SEP) [52,106], corresponding to all quantum channels $\Lambda : AB \to A'B'$ whose Choi matrix is separable across the bipartition $AA'|BB'$. Alternatively, this can be viewed as the class of resource-free operations in the resource theory of entanglement (cf. Sec. 5.3), that is, all operations which admit a Kraus decomposition into separable operators. It is obvious that this is a strictly smaller class than PPT, and it has been shown to be strictly larger than LOCC [106].

[52] Vedral et al., PRL 1997

[106] Bennett et al., PRA 1999

The class of separability-preserving operations SEPP is defined as all CPTP maps Λ such that $\sigma \in \mathsf{SEP} \Rightarrow \Lambda(\sigma) \in \mathsf{SEP}$, that is, as the maximal class of free operations in the resource theory of entanglement. Notice that, this class does not *completely* preserve separability, in the sense that it could generate entanglement if applied to a part of a larger system; if such complete preservation is imposed, we instead obtain the class of separable operations again. The inclusions between the different classes of operations are shown in Fig. 8.1.

We will additionally consider a subset of LOCC operations in which only one party can send classical information to the other party, denoted as one-way LOCC (1-LOCC).

The fidelity of distillation under SEPP was first derived in [224], and can be characterised as follows.

[224] Brandão & Plenio, CMP 2010

Lemma 8.13. $F_{\mathrm{SEPP}}(\rho, m) = G^{(m)}_{\mathsf{SEP}}(\rho) = \dfrac{1}{m}\left(T^{(m-1)}_{\mathsf{SEP}}(\rho) + 1\right)$.

Proof. Follows in exactly the same way as the proof of Thm. 8.10, since isotropic states of the form

$$\langle Z, X\rangle \Psi_m + \frac{\langle Z, \mathbb{1} - X\rangle}{m^2 - 1} (\mathbb{1} - \Psi_m) \tag{8.36}$$

are separable if and only if they are PPT [230]. ■

[230] Horodecki & Horodecki, PRA 1999

By the inclusion $\mathsf{SEP} \subset \mathsf{PPT}_+ \subset \mathsf{PPT}'$, we then have that

$$\begin{aligned} F_{\text{SEPP}}(\rho, m) &\geq F_{\text{P-PPT}_+}(\rho, m) \\ &\geq F_{\text{PPT}}(\rho, m) \end{aligned} \tag{8.37}$$

thus establishing a hierarchy of rates of distillation between the operations SEPP, P-PPT$_+$, and PPT — notice that this does not follow from their definition, as there is no inclusion between the sets SEPP and PPT or between SEPP and P-PPT$_+$.

From Theorem 8.4, we then immediately get the fidelity of distillation under SEPP of a pure state $|\psi\rangle$ as $\frac{1}{m} \|\xi(|\psi\rangle)\|^2_{[m]}$, where we recall that $\xi(|\psi\rangle)$ is the vector of the Schmidt coefficients. This can then be related to the optimal fidelity of pure-state distillation under LOCC and one-way LOCC, which can be obtained from a result by Vidal et al. [235]. We will now rederive the exact expression for the fidelity of pure-state distillation under LOCC in terms of the m-distillation norm, and in particular establish an operational equivalence between all relevant sets of operations in the distillation of entanglement from pure states.

[235] Vidal et al., PRA 2000

Theorem 8.14. *For any pure state $|\psi\rangle$, any integer $m \geq 1$, and any set of operations $\mathcal{O} \in \{1\text{-LOCC}, \text{LOCC}, \text{PPT}, \text{P-PPT}, \text{P-PPT}_+, \text{SEPP}\}$, the fidelity of distillation is given by*

$$F_{\mathcal{O}}(\psi, m) = \frac{1}{m} \|\xi(|\psi\rangle)\|^2_{[m]} \,. \tag{8.38}$$

Proof. It is well-known that the deterministic transformation from $|\psi\rangle$ to another pure state $|\eta\rangle$ is possible with (both one-way and two-way) LOCC if and only if the Schmidt vector $\xi(|\psi\rangle) \coloneqq (\alpha_1, \ldots, \alpha_d)^T$ is majorised by the Schmidt vector $\xi(|\eta\rangle) \coloneqq (\beta_1, \ldots, \beta_d)^T$ [198], that is,

[198] Nielsen, PRL 1999

$$\sum_{i=1}^{k} \alpha_i^2 \leq \sum_{i=1}^{k} \beta_i^2 \quad \forall k \in \{1, \ldots, d\} \tag{8.39}$$

where we have assumed without loss of generality that the Schmidt coefficients are given in non-increasing order. Following the proof of Thm. 7.11, we can then see that the Schmidt coefficients of $|\psi\rangle$ are majorised by the Schmidt coefficients of the state

$$|\eta\rangle = \left(\alpha_1, \ldots, \alpha_{m-k^\star}, \underbrace{\frac{L_2}{\sqrt{k^\star}}, \ldots, \frac{L_2}{\sqrt{k^\star}}}_{k^\star \text{ times}}\right)^T \tag{8.40}$$

expressed in the Schmidt basis of $|\Psi_m\rangle$, where $k^\star = \arg\min_{1\leq k\leq m} \frac{1}{k} \|\alpha_{m-k+1:d}\|_{\ell_2}^2$ and $L_2 = \|\alpha_{m-k^\star+1:d}\|_{\ell_2}$. This then gives

$$\begin{aligned} F_{\text{1-LOCC}}(\psi, m) &\geq |\langle\Psi_m|\eta\rangle|^2 \\ &= \frac{1}{m}\left(\|\alpha_{1:m-k^\star}\|_{\ell_1} + \sqrt{k^\star}\,\|\alpha_{m-k^\star+1:d}\|_{\ell_2}\right)^2 \\ &= \frac{1}{m}\,\|\xi(|\psi\rangle)\|_{[m]}^2 \,. \end{aligned} \tag{8.41}$$

On the other hand, since 1-LOCC $\subset$ SEPP, we have

$$F_{\text{LOCC}}(\psi, m) \leq F_{\text{SEPP}}(\psi_m) = \frac{1}{m}\left(T_{\text{SEP}}^{(m-1)}(\psi) + 1\right) = \frac{1}{m}\,\|\xi(|\psi\rangle)\|_{[m]}^2 \tag{8.42}$$

by Thm. 8.4, which concludes the proof. ∎

This leads to an interesting characterisation of the m-distillation norm in two different ways. Writing $|\psi\rangle \xrightarrow{\text{LOCC}} |\eta\rangle$ to denote that the pure state transformation is possible with LOCC, we have

$$\begin{aligned} \|\xi(|\psi\rangle)\|_{[m]} &= \sqrt{m}\,\max\left\{ |\langle\Psi_m|\eta\rangle| \;\middle|\; |\psi\rangle \xrightarrow{\text{LOCC}} |\eta\rangle \right\} \\ &= \sqrt{m}\,\max\left\{ |\langle\psi|\eta\rangle| \;\middle|\; |\eta\rangle \xrightarrow{\text{LOCC}} |\Psi_m\rangle \right\}, \end{aligned} \tag{8.43}$$

where the maximisation is over normalised state vectors $|\eta\rangle$.

As a straightforward corollary of the above result, we can characterise the quantifiers $G_{\mathcal{S}}^{(m)}$ for several different sets.

Corollary 8.15. *For any $1 \leq m \leq d$, any pure state $|\psi\rangle$, and any $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+, \mathsf{PPT}'\}$ we have*

$$G_{\mathcal{S}}^{(m)}(|\psi\rangle\langle\psi|) = \frac{1}{m}\,\|\xi(|\psi\rangle)\|_{[m]}^2 \tag{8.44}$$

8.2.3 *Maximally correlated states*

Let us consider a bipartite system with $d = d_A = d_B$. A *maximally correlated state* is any state of the form $\rho_{\text{mc}} = \sum_{i,j} a_{ij}\,|ii\rangle\langle jj|$ for some local orthonormal bases $\{|i\rangle\}$. The name for this class of states comes from the fact is that the two parties are guaranteed to obtain the same measurement results for any measurement in their local basis $\{|i\rangle\}$. Notice that any maximally correlated state has a corresponding single-party state $\rho = \sum_{i,j} a_{ij}\,|i\rangle\langle j|$ with the same coefficients in the local basis.

It has been conjectured in several works that the resource theory of coherence is equivalent to the resource theory of entanglement restricted to maximally correlated states [55,56]. Although this conjecture is still unsolved in full generality, we know that many operational quantifiers such as the entanglement of formation, relative entropy of entanglement (and other Rényi entropy–based entanglement measures), and asymptotic distillable entanglement can be evaluated on maximally correlated states by computing the corresponding coherence quantifiers, typically significantly simpler to evaluate and satisfying useful properties such as additivity [55,123]. Furthermore, an operational equivalence between SIO or IO transformations acting on ρ and LOCC operations acting on ρ_{mc} has been suggested, although so far this conjecture has been proved only in specific cases [55,56].

[55] Winter & Yang, PRL 2016

[56] Chitambar & Gour, PRL 2016

[123] Zhu et al., JPA 2017

To obtain a result allowing us to quantify the one-shot distillable entanglement of maximally correlated states, we will fix a choice of basis $\{|i\rangle\}$ for the single-party state ρ, and use $\mathcal{I}$ to denote the set of all incoherent states in this basis. Furthermore, we define the subset of separable states $\mathsf{SEP}_{\mathrm{mc}} \coloneqq \mathrm{conv}\{|ii\rangle\langle ii|\}$ where $\{|ii\rangle\}_{i=1}^{d}$ is the basis of the corresponding maximally correlated state ρ_{mc}.

Theorem 8.16. *For any maximally correlated state and any $m \geq 1$ it holds that*

$$F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m) = F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m) = G_{\mathcal{I}}^{(m)}(\rho). \tag{8.45}$$

Proof. Using Thm. 8.6, we have

$$\begin{aligned} F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m) &\leq F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m) \\ &= \min_{X \in \mathsf{SEP}^{\star\star}} \mathrm{Tr}(\rho_{\mathrm{mc}} - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \\ &\leq \min_{X \in \mathsf{SEP}_{\mathrm{mc}}^{\star\star}} \mathrm{Tr}(\rho_{\mathrm{mc}} - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \\ &= \min_{X \in \mathcal{I}^{\star\star}} \mathrm{Tr}(\rho - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \\ &= G_{\mathcal{I}}^{(m)}(\rho). \end{aligned} \tag{8.46}$$

On the other hand, let $W^\star = \sum_{i,j} W_{ij} |i\rangle\langle j|$ be the optimal solution to the dual problem of

$$G_{\mathcal{I}}^{(m)}(\rho) = \max\left\{ \langle \rho, W\rangle \;\middle|\; 0 \leq W \leq \mathbb{1},\ W \in \frac{1}{m}\mathcal{I}^\circ \right\}. \tag{8.47}$$

Notice that $W^\star \in \frac{1}{m}\mathcal{I}^\circ$ is equivalent to $\Delta(W) \leq \frac{1}{m}\mathbb{1}$, that is, $\max_i W_{ii} \leq \frac{1}{m}$. Consider then the matrix $W_{\mathrm{mc}} \coloneqq \sum_{i,j} W_{ij} |ii\rangle\langle jj|$, defined in the basis of ρ_{mc}. The eigenvalues of $W_{\mathrm{mc}}^{T_B}$ can be straightforwardly verified to be $\{W_{ii},\ \pm|W_{ij}|\}_{i,j=1}^{d}$. Positivity of $W^\star$ imposes that

$$\max_{i,j} |W_{ij}| \leq \max_i W_{ii} \tag{8.48}$$

from which it follows that

$$\Gamma^{\circ}_{\mathsf{PPT}'}(W_{\mathrm{mc}}) = \left\| W_{\mathrm{mc}}^{T_B} \right\|_{\infty} = \max_i W_{ii} \leq \frac{1}{m}. \tag{8.49}$$

This means that $W_{\mathrm{mc}} \in \frac{1}{m}\mathsf{PPT}'^{\circ}$, and so

$$\begin{aligned} F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m) &= G^{(m)}_{\mathsf{PPT}'}(\rho_{\mathrm{mc}}) \\ &= \max\left\{ \langle \rho_{\mathrm{mc}}, W\rangle \;\middle|\; 0 \leq W \leq \mathbb{1},\ W \in \frac{1}{m}\mathsf{PPT}'^{\circ} \right\} \\ &\geq \langle \rho_{\mathrm{mc}}, W_{\mathrm{mc}}\rangle \\ &= G^{(m)}_{\mathcal{I}}(\rho). \end{aligned} \tag{8.50}$$

■

Notice that

$$G^{(m)}_{\mathcal{I}}(\rho) = \frac{1}{m}\left(T^{(m-1)}_{\mathcal{I}}(\rho) + 1 \right) \tag{8.51}$$

with $T^{(m)}_{\mathcal{I}}$ being the coherence measures defined in Ch. 7.

Interestingly, in contrast to many other results which show an exact equality between operational quantities in the resource theory of coherence and the resource theory of entanglement of maximally correlated states, our result above shows a slight discrepancy between the two resource theories: in particular, it is not difficult to find numerical examples of states such that

$$F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m) = G^{(m)}_{\mathcal{I}}(\rho) > G^{(m)}_{\mathcal{J}}(\rho) = F_{\mathrm{MIO}}(\rho, m), \tag{8.52}$$

and so the distillable entanglement of a maximally correlated state under the maximal set of operations in the resource theory of entanglement (SEPP) can be strictly larger than the distillable entanglement the corresponding single-partite state under the maximal set of operations in the resource theory of coherence (MIO). This shows in particular that, in the distillation of entanglement from maximally correlated states under SEPP, it is not sufficient to consider operations whose output remains in the maximally correlated subspace — indeed, if this were the case, any such operation could always be mapped to a corresponding MIO operation, and the fidelities $F_{\mathrm{MIO}}(\rho, m)$ and $F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m)$ would be equal.

Nevertheless, the result reduces the much more complicated optimisation problems in the quantification of one-shot distillable entanglement to the simple task of evaluating coherence measures.

8.2.4 *Isotropic states*

Consider $d = d_A = d_B$ and define the isotropic states as

$$\rho_f = f\Psi_d + \frac{1-f}{d^2-1}\left(\mathbb{1} - \Psi_d\right) \tag{8.53}$$

with $0 \leq f \leq 1$. We then have the following result.

Proposition 8.17. *For any isotropic state ρ_f and any $\mathcal{O} \in \{\mathrm{PPT}, \text{P-PPT}_+, \mathrm{SEPP}\}$, it holds that*

$$F_{\mathcal{O}}(\rho_f, m) = \begin{cases} \frac{1}{m} & f \leq \frac{1}{d} \\ \frac{df-1}{d-1} + \frac{d(1-f)}{m(d-1)} & f \geq \frac{1}{d}. \end{cases} \tag{8.54}$$

Proof. This is a slight extension of the result in [156], and the proof is based thereon. [156] Rains, IEEE TIT 2001

Take $1 \leq m \leq d$. If $f \leq \frac{1}{d}$, then $\rho_f \in \mathsf{SEP}$, and we have $F_{\mathrm{PPT}}(\rho_f, m) = F_{\mathrm{SEPP}}(\rho_f, m) = \frac{1}{m}$. If $f \geq \frac{1}{d}$, then on one hand we can take

$$W = \frac{k-1}{(d-1)k} d\Psi_d + \frac{d-k}{(d-1)k} \mathbb{1} \tag{8.55}$$

which satisfies

$$\begin{aligned} \left\| W^{T_B} \right\|_\infty &\leq \frac{k-1}{(d-1)k} \left\| (d\Psi_d)^{T_B} \right\|_\infty + \frac{d-k}{(d-1)k} \left\| \mathbb{1} \right\|_\infty \\ &= \frac{1}{k} \end{aligned} \tag{8.56}$$

since $(d\Psi_d)^{T_B}$ is the swap operator, whose eigenvalues are easily checked to be ± 1. This gives

$$\begin{aligned} F_{\mathrm{PPT}}(\rho, m) &\geq \langle \rho_f, W \rangle \\ &= \frac{df-1}{d-1} + \frac{d(1-f)}{m(d-1)}. \end{aligned} \tag{8.57}$$

On the other hand, take

$$X = \frac{d(1-f)}{d^2-1} \left(\Psi_d + \frac{1}{d} \mathbb{1} \right). \tag{8.58}$$

It is known that $\Psi_d + \frac{1}{d} \in \mathsf{SEP}^{**}$ [94], which gives [94] Vidal & Tarrach, PRA 1999

$$\begin{aligned} F_{\mathrm{SEPP}}(\rho_f, m) &\leq \mathrm{Tr}(\rho_f - X)_+ + \frac{1}{m} \mathrm{Tr}\, X \\ &= \frac{df-1}{d-1} + \frac{d(1-f)}{m(d-1)}. \end{aligned} \tag{8.59}$$

■

8.2.5 *Beyond deterministic distillation*

The investigation of probabilistic entanglement transformations and distillation under different classes of operations has been considered in several works

[38,154,200,202,236,237]. The setting can be defined analogously to the characterisation of coherence distillation in Sec. 7.3. We do not consider this task explicitly here, although a particularly interesting open question in this regard is to characterise the capabilities of separability-preserving operations in probabilistic distillation.

Similarly, the formalism of assisted distillation of entanglement has been considered before [187]. Here, in direct analogy with our results for coherence in Sec. 7.4, we obtain several new results as direct corollaries of our investigation in this Chapter. Most notably, it follows immediately that one-shot assisted distillable entanglement is exactly the same under any class of operations ranging from one-way LOCC to P-PPT$_+$ to SEPP, and some of the proofs in [187] can be simplified by employing our definition of the m-distillation norm. The results follow in exactly the same way as 7.4, so we do not consider them again.

[38] Bennett et al., PRA 1996

[154] Vidal, PRL 1999

[200] Ishizaka & Plenio, PRA 2005

[202] Lo & Popescu, PRA 2001

[236] Jonathan & Plenio, PRL 1999

[237] Rozpędek et al., PRA 2018

[187] Buscemi & Datta, IEEE TIT 2013

Conversion of nonclassicality into entanglement

9

We have so far investigated two signature features of quantum theory: coherence and entanglement. Coherence embodies the *nonclassicality* of quantum mechanics, that is, the possibility for a quantum system to exist in a superposition of "classical" states. The latter are usually determined based on physical considerations: in preceding chapters, we looked specifically at the resource theory of coherence, where the classical states form an orthonormal basis, while in continuous-variable systems, they are often identified with the Glauber-Sudarshan coherent states [238,239].

Superposition underlies other nonclassical phenomena such as quantum correlations among parts of a quantum system. In particular, entanglement itself stems from the superposition principle in conjunction with the tensor product structure associated to composite systems. Despite the common origin, entanglement and superposition are formalised according to different resource theories: the former being tied to the paradigm of spatially separated laboratories which can only implement LOCC operations for free, while the second specified by the inability to create superpositions of the classical states for free. Consequently, these two resources, like two currencies, enjoy different uses in quantum technologies. It thus becomes particularly relevant to investigate the connection between these two types of resource beyond a merely conceptual standpoint, and to devise operational schemes that allow the dynamical transformation of one into the other.

Several works have analyzed this problem. In quantum optics, nonclassicality gets mapped into entanglement by a beam splitter [240–244], while, in the discrete-variable scenario, it is the controlled NOT (CNOT) gate [245,246] that plays a similar role. The quantitative interplay between the degree of nonclassicality and the bipartite entanglement obtained from it has been investigated as well [2,17–19,61,243,247]. These studies have advanced our understanding of nonclassicality as a resource in systems of arbitrary dimension [15,18,19,248–253].

In this chapter, we will investigate the conversion of nonclassicality, which generalises the previously considered notion of quantum coherence, into multipartite entanglement. We will first show that there always exists a state-independent unitary mapping, realised by operations which alone cannot create nonclassicality, such that the presence of k-level nonclassicality in the state of a single d-level system is necessary and sufficient to create $k+1$-partite entanglement between the system and k ancillas. To exemplify such a conversion procedure, we specialise to quantum coherence as an instance of nonclassicality, and in-

[238] Glauber, Phys. Rev. 1963

[239] Sudarshan, PRL 1963

[240] Kim et al., PRA 2002

[241] Wang, PRA 2002

[242] Wolf et al., PRL 2003

[243] Asbóth et al., PRL 2005

[244] Ivan et al., Phys Rev A 2011

[245] Streltsov et al., PRL 2011

[246] Piani et al., PRL 2011

[2] Sperling & Vogel, Phys. Scr. 2015

[17] Vogel & Sperling, PRA 2014

[18] Killoran et al., PRL 2016

[19] Theurer et al., PRL 2017

[61] Streltsov et al., PRL 2015

[247] Ge et al., PRA 2015

[15] Baumgratz et al., PRL 2014

[248] Brandão & Plenio, Nat. Phys 2008

[249] Zhang et al., PRA 2016

[250] Streltsov et al., PRL 2016

[251] Xu, PRA 2016

[252] Tan et al., PRL 2017

[253] Mukhopadhyay et al., preprint 2017

troduce an explicit physical protocol which directly converts k-level coherence into $k+1$-body multipartite entanglement. The protocol entangles a single qudit with up to d qubits by a sequential application of generalised CNOT gates (free operations in the resource theory of coherence). The protocol can be further extended via the decoupling of the qudit system by LOCC (free operations in the resource theory of entanglement), to provide a mapping of k-coherence into multipartite entanglement of the ancillary qubits alone. Finally, we show that the initial amount of k-coherence places a quantitative restriction on the amount of entanglement that can be converted from it — in particular, the fidelity-based geometric measure of $k+1$-partite entanglement at the output of the protocol is exactly equal to the fidelity-based geometric measure of k-coherence in the input state of the d-level system. The chapter is based on results published in [BR5].

[BR5] Regula et al., NJP 2018

9.1 RELATIONS BETWEEN NONCLASSICALITY AND ENTANGLEMENT

Nonclassicality is a notion that depends on the preassigned set of states that are deemed "classical". Choosing a finite set of states $\{|\chi_i\rangle\}$ which spans the whole Hilbert space $\mathcal{H}$ to constitute the pure classical states, as dictated by the physics of the problem under investigation, one asks whether a mixed state ρ can be represented as a convex combination of classical states only. Nonclassicality therefore naturally fits into the resource theoretic formalism considered throughout this thesis, where the set of free states $\mathcal{C}$ is formed by the convex hull of $\{|\chi_i\rangle\}$.

For finite-dimensional systems, the notion of nonclassicality is often understood as quantum coherence. However, following [18,19], we note that the approach considered here is more general, since one does not require the states $\{|\chi_i\rangle\}$ to be orthogonal. This provides a common framework applicable e.g. to classical sets formed by SU(N) Gilmore-Perelomov coherent states [254,255] in discrete-variable systems and Glauber-Sudarshan coherent states [238,239] in continuous-variable systems, where the latter are not orthogonal yet such that any finite subset thereof is linearly independent [17].

[18] Killoran et al., PRL 2016

[19] Theurer et al., PRL 2017

[254] Gilmore, Annals of Physics 1972

[255] Perelomov, Commun.Math. Phys. 1972

[238] Glauber, Phys. Rev. 1963

[239] Sudarshan, PRL 1963

[17] Vogel & Sperling, PRA 2014

[2] Sperling & Vogel, Phys. Scr. 2015

The framework adopted here leads to a natural measure of the level of nonclassicality of a state. For a pure state, one can indeed define the *nonclassical rank* (NR) [2,18] as

$$\mathrm{NR}\left(|\psi\rangle\right) = \min\left\{ r \in \mathbb{N} \;\middle|\; |\psi\rangle = \sum_{i=1}^{r} c_i |\chi_i\rangle\,,\ |\chi_i\rangle \in \mathcal{C} \right\}, \tag{9.1}$$

with nonzero complex coefficients c_i. This clearly resembles the definition of the Schmidt rank $\mathrm{SR}(|\psi\rangle)$ of bipartite entangled states, and it can be extended to mixed states in the same way as the the Schmidt rank is extended to the Schmidt

number SN: we thus define the *nonclassical number* (NN) of a mixed state ρ as

$$\mathrm{NN}(\rho) = \min_{\{p_i,|\psi_i\rangle\}} \max_i \mathrm{NR}(|\psi_i\rangle), \tag{9.2}$$

where the minimisation is performed over all pure-state convex decompositions of ρ into $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. In other words, in every such decomposition at least one state has nonclassical rank $\mathrm{NR}(|\psi_i\rangle) \geq \mathrm{NN}(\rho)$, and there exists a decomposition where all pure states have nonclassical rank $\mathrm{NR}(|\psi_i\rangle) \leq \mathrm{NN}(\rho)$.

Killoran et al. [18] showed that there always exists an isometry, consisting of adding an ancilla and applying a global unitary, which maps each pure state of nonclassical rank k into a bipartite entangled pure state of Schmidt rank k. In fact, as we show below, this result can be straightforwardly extended to the general case of mixed states:

[18] Killoran et al., PRL 2016

Proposition 9.1. *Let $\mathcal{H}$ be a d-dimensional Hilbert space, $\mathbb{D}$ the corresponding set of density operators, and $\mathcal{H}_{anc} \cong \mathcal{H}$ the Hilbert space of an ancillary system. Then if the classical pure states $\{|\chi_i\rangle\}_{i=1}^{d}$ form a linearly independent set spanning $\mathcal{H}$, there exists an isometry $W : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_{anc}$ such that for any state $\rho \in \mathbb{D}$ we have $\mathrm{NN}(\rho) = \mathrm{SN}(W\rho W^\dagger)$.*

Proof. To begin, let us note the fact that the set of all possible pure states belonging to pure-state decompositions of $\rho \in \mathbb{D}$ is given precisely by the set of pure states in the support of ρ [216,217]. By the result of [18], we have that there exists a unitary U such that $\mathrm{NR}(|\psi\rangle) = \mathrm{SR}(U(|\psi\rangle \otimes |\psi_{\mathrm{anc}}\rangle))\ \forall\, |\psi\rangle \in \mathcal{H}$ where $|\psi_{\mathrm{anc}}\rangle \in \mathcal{H}_{\mathrm{anc}}$ is a fixed reference state for the ancilla system. The isometry W is given by attaching the ancilla state $|\psi_{\mathrm{anc}}\rangle$ composed with the action of the unitary U. Following [243,256], we note that there is a one-to-one correspondence between the pure-state decompositions of ρ and the decompositions of $\rho' = W\rho W^\dagger$, given exactly by the action of W. Notice in particular that W can be inverted on its image, and that any $|\psi'\rangle$ in the support of ρ' has as preimage $W^\dagger|\psi'\rangle$ in the support of ρ, which means that by the properties of W one has $\mathrm{NR}(W^\dagger|\psi'\rangle) = \mathrm{SR}(|\psi'\rangle)$. Assuming ρ has $\mathrm{NN}(\rho) = k$, then it is possible to find a pure-state decomposition of it which only contains states with nonclassicality rank less or equal to k. The pure states in one such decomposition will then be transformed by the action of W into entangled states of Schmidt rank at most k, which will form a pure-state decomposition of ρ'. This proves that $\mathrm{SN}(\rho') \leq \mathrm{NN}(\rho)$. On the other hand, suppose $\mathrm{SN}(\rho') = l$; then it is possible to find a pure-state decomposition of ρ' such that it only contains states with Schmidt rank less or equal to l. Under the action of $W^\dagger$, such a decomposition gives rise to a pure-state decomposition of ρ whose elements have nonclassicality rank at most l. This proves that $\mathrm{NN}(\rho) \leq \mathrm{SN}(\rho')$. ■

[216] Schrödinger, Math. Pr. Cam. Phil. S. 1936

[217] Hughston et al., PLA 1993

[243] Asbóth et al., PRL 2005

[256] Piani & Adesso, PRA 2012

The simplest example of a relation of this kind can be obtained in the resource theory of coherence, where the faithful conversion of coherent states into

maximally correlated states of the form $\rho_{\mathrm{mc}} = \sum_{i,j} \rho_{ij} |ii\rangle\langle jj|$ can be performed with the isometry $W|i\rangle = |ii\rangle$. As a corollary of this relation, the Schmidt number of a maximally correlated state can be computed by considering the nonclassical number of the single-party state $\rho = \rho_{i,j} |i\rangle\langle j|$.

We will show that an analogous faithful conversion of multilevel nonclassicality into genuine multipartite entanglement is always possible. As before, we define a pure state $|\psi\rangle$ to be k-producible if it can be written as $|\psi\rangle = |\psi_1\rangle \otimes \ldots \otimes |\psi_m\rangle$ with each $|\psi_j\rangle$ pertaining to at most k parties, and a mixed state ρ to be k-producible if it can be written as a convex combination of k-producible pure states. We call a state ρ genuinely k-partite entangled if it is k-producible but not $k-1$-producible; equivalently, under such conditions we say that ρ has *entanglement depth* $\mathrm{D_E}(\rho) = k$ [257]. A 1-producible state ρ has $\mathrm{D_E}(\rho) = 1$ and is fully separable.

[257] Sørensen & Mølmer, PRL 2001

Theorem 9.2. *Let $\mathcal{H}$ be a d-dimensional Hilbert space, and $\mathcal{H}_{anc}$ the Hilbert space of an ancillary system. Then if the classical pure states $\{|\chi_i\rangle\}_{i=1}^{d} \in \mathcal{H}$ form a linearly independent set spanning $\mathcal{H}$, there exists an isometry $U : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_{anc}^{\otimes d}$ such that for any state $\rho \in \mathbb{D}$ with nonclassical number $\mathrm{NN}(\rho) = k$, $U\rho U^\dagger$ is genuinely $k+1$-partite entangled iff ρ is nonclassical ($2 \leq k \leq d$) and $U\rho U^\dagger$ is fully separable iff ρ is classical ($k = 1$).*

Proof. We adapt the methods of Ref. [18] to show the existence of this mapping. Let us consider the case of pure states first. Define

[18] Killoran et al., PRL 2016

$$|c_i\rangle = |\chi_i\rangle \otimes |\psi_{\mathrm{anc}}\rangle \in \mathcal{H} \otimes \mathcal{H}_{\mathrm{anc}}^{\otimes d} \tag{9.3}$$

with $|\psi_{\mathrm{anc}}\rangle \in \mathcal{H}_{\mathrm{anc}}^{\otimes d}$ a fixed (fully unentangled) reference state of the ancilla systems. Define $\{|b_i\rangle\}_{i=1}^{d} \in \mathcal{H}_{\mathrm{anc}}^{\otimes d}$ as

$$|b_i\rangle \coloneqq |0\rangle^{\otimes i-1} \otimes |\lambda\rangle \otimes |0\rangle^{\otimes d-i} \tag{9.4}$$

with $|\lambda\rangle \coloneqq \sqrt{\lambda}\,|0\rangle + \sqrt{1-\lambda}\,|1\rangle$, where $|0\rangle, |1\rangle \in \mathcal{H}_{\mathrm{anc}}$ are orthonormal, and $0 \leq \lambda \leq 1$. The Gram matrix of a set of states $\{|\phi_i\rangle\}$ is defined as $[G^{(\phi)}]_{ij} = \langle\phi_i|\phi_j\rangle$, and has full rank iff the $|\phi_i\rangle's$ are linearly independent [258]. Define a μ-dependent matrix $B(\mu)$ by

[258] Jozsa & Schlienz, PRA 2000

$$[B(\mu)]_{ij} = \begin{cases} 1 & i = j \\ \mu & i \neq j \end{cases}. \tag{9.5}$$

We have $G^{(b)} = B(\lambda)$. Define $M(\epsilon) \coloneqq G^{(c)} \circ B(1+\epsilon)$ where $\circ$ is the Schur (entrywise) product. Since $\lim_{\epsilon\to 0} M(\epsilon) = G^{(c)} > 0$ and $\mathrm{diag}(M(\epsilon)) = (1, 1, \ldots, 1)$, it follows that, for sufficiently small $\epsilon > 0$, $M(\epsilon)$ is the Gram matrix $G^{(a)}$ of a set of linearly independent states $\{|a_i\rangle\}_{i=1}^{d}$. We then have $G^{(c)} = G^{(a)} \circ B\left(\frac{1}{1+\epsilon}\right) = G^{(a)} \circ G^{(b)}$, for $\lambda = (1+\epsilon)^{-1}$, which means that the sets of states $\{|c_i\rangle\}$ and

$\{|a_i\rangle \otimes |b_i\rangle\}$ have equal Gram matrices, and so there exists a unitary U such that $U|c_i\rangle = |a_i\rangle \otimes |b_i\rangle \ \forall\, i$ [258,259]. The isometry V is defined by the composition of attaching the ancilla state $|\psi_{\text{anc}}\rangle$ followed by the action of U.

[258] Jozsa & Schlienz, PRA 2000

[259] Chefles et al., IJQI 2004

Now consider a general pure qudit state $|\psi\rangle = \sum_{i=1}^{d} \psi_i |\chi_i\rangle$. Then,

$$\begin{aligned} |\psi'\rangle = V|\psi\rangle &= \sum_{i=1}^{d} \psi_i |a_i\rangle |b_i\rangle \\ &= \sum_{i=1}^{d} \psi_i |a_i\rangle |0\rangle^{\otimes i-1} |\lambda\rangle |0\rangle^{\otimes d-i} . \end{aligned} \tag{9.6}$$

It is convenient to use the fact that the entanglement depth of $|\psi'\rangle$ is not affected by a local filter $S \otimes L^{\otimes d}$, with S a qudit operator such that $S|a_i\rangle = |i\rangle$, and L a qubit operator such that $L|0\rangle = |0\rangle$, $L|\lambda\rangle = |1\rangle$. Thus, we can study the entanglement depth of the state

$$|\widetilde{\psi}'\rangle \propto \sum_{i=1}^{d} \psi_i |i\rangle |0\rangle^{\otimes i-1} |1\rangle |0\rangle^{\otimes d-i} = \sum_{i=1}^{d} \psi_i |i\rangle |\underline{2^{d-i}}\rangle , \tag{9.7}$$

where $|\underline{2^{d-i}}\rangle$ is the string of qubits corresponding to 2^{d-i} in binary (padded with zeros from the left as needed), e.g. $|\underline{2^3}\rangle = |00\cdots01000\rangle$ since $2^3 = 1000_2$. It is evident that $|\widetilde{\psi}'\rangle$ is fully product iff there is only one term in the superposition, that is $\mathrm{NR}(|\psi\rangle) = 1$ iff $\mathrm{D_E}(|\psi'\rangle) = 1$. In the following we will consider $\mathrm{NR}(|\psi\rangle) \geq 2$, and in this case we will prove that $\mathrm{D_E}(|\psi'\rangle) = \mathrm{NR}(|\psi\rangle) + 1$.

To show that $\mathrm{NR}(|\psi\rangle) = k \geq 2$ implies $\mathrm{D_E}(|\widetilde{\psi}'\rangle) = k+1$, assume w.l.o.g. that the k nonzero coefficients ψ_i are the first ones. Then

$$|\widetilde{\psi}'\rangle \propto \Big(\sum_{i=1}^{k} \psi_i |i\rangle |0\rangle^{\otimes i-1} |1\rangle |0\rangle^{\otimes k-i} \Big) |0\rangle^{\otimes d-k} , \tag{9.8}$$

and the claim follows by showing that $\sum_{i=1}^{k} \psi_i |i\rangle |0\rangle^{\otimes i-1} |1\rangle |0\rangle^{\otimes k-i}$ cannot be factorised in any non-trivial way. This holds, as the reduced state of the k ancillary qubits is proportional to $\sum_{i=1}^{k} |\psi_i|^2 |0\rangle\langle 0|^{\otimes i-1} \otimes |1\rangle\langle 1| \otimes |0\rangle\langle 0|^{\otimes k-i}$, so the marginal state of any subset of these k qubits is evidently mixed. On the other hand, to prove that $\mathrm{D_E}(|\psi'\rangle) = \mathrm{D_E}(V|\psi\rangle) = k+1 \geq 3$ implies $\mathrm{NR}(|\psi\rangle) = k$, notice that the isometry is invertible on its image. Thus, $|\psi\rangle = V^\dagger |\psi'\rangle$; since we have just proven that $\mathrm{D_E}(V|\psi\rangle) = \mathrm{NR}(|\psi\rangle) + 1$, we arrive at the claim for pure states.

The mixed-state case again begins by noting that there is a one-to-one correspondence between the pure state decompositions $\{p_i, |\psi_i\rangle\}$ of ρ and $\{p_i, |\psi_i'\rangle\}$ of $\rho' = V\rho V^\dagger$, with each input-output pair of states respecting the relation just discussed: either $\mathrm{NR}(|\psi_i\rangle) = \mathrm{D_E}(|\psi_i'\rangle) = 1$ or $\mathrm{NR}(|\psi_i\rangle) + 1 = \mathrm{D_E}(|\psi_i'\rangle)$. Thus, with the exception of the first (trivial) case, we have $\mathrm{NR}(\rho) + 1 = \mathrm{D_E}(\rho')$. Indeed, the pure-state mapping with the above properties, together with the definitions of nonclassical number and entanglement depth, entail that $\mathrm{NN}(\rho) = m$ implies $\mathrm{D_E}(\rho') \leq m+1$, and, in turn, $\mathrm{D_E}(\rho') = l$ implies $\mathrm{NN}(\rho') \leq l-1$. These relations can only be satisfied for $l = m+1$. ■

Theorem 9.2 shows that there always exists an isometry which faithfully converts the k-level nonclassicality of a quantum system into multipartite entanglement with k other ancillary systems. We note that the specifics of the mappings are not fixed by the theorem, and one could always devise other ways to convert the nonclassicality into entanglement. In particular, the mappings presented in the proofs only use two levels of each ancillary system, resulting in entanglement akin to that of W states [260]. One may consider other kinds of operations which create qualitatively different multipartite entanglement — for instance, one can instead attach a number d of $d+1$-dimensional ancilla systems and choose $\{|b_j\rangle\}_{j=1}^{d}$ such that $|b_j\rangle = |\lambda_j\rangle^{\otimes d}$ and $|\lambda_j\rangle = \sqrt{\lambda^{1/d}}\,|0\rangle + \sqrt{1-\lambda^{1/d}}\,|j\rangle$ where $\{|j\rangle\}$ is now an orthonormal basis for $\mathcal{H}_{\text{anc}}$. Following a similar argument to Thm. 9.2, this will then introduce a generalised GHZ-type entanglement between the qudits, entangling as many levels of the systems as the nonclassical rank of the original state. However, the choice of a W-type mapping in the theorems makes the conversion quite appealing in practice, as it only requires qubit ancillas, and, as we show below, enables one to create entanglement by a sequential application of two-body gates on the nonclassical system and each ancilla.

[260] Dür et al., PRA 2000

9.2 COHERENCE CONVERSION PROTOCOL

We will now specialise to the framework of quantum coherence. Here, the classical states $\{|i\rangle\}_{i=1}^{d}$ are taken to form a fixed orthonormal basis for $\mathcal{H}$. As before (cf. Sec. 6.1), we can then define a hierarchy of coherence levels by considering the coherence rank $\mathrm{CR}(|\psi\rangle)$ and the coherence number

$$\mathrm{CN}(\rho) \coloneqq \min_{\{p_i,|\psi_i\rangle\}} \max_i \mathrm{CR}\left(|\psi_i\rangle\right). \tag{9.9}$$

We will refer to states with coherence number at most k as *k-coherent states*. Clearly, 1-coherence corresponds to classicality, k-coherence for any $k \geq 2$ stands as a fine graining of the usual notion of coherence, and d-coherence is the maximal coherence level of a d-level system.

The k-coherence of a single qudit can be converted into multipartite entanglement in different physical ways. To show this, we design a protocol to convert k-coherence into $k+1$-partite entanglement between the qudit and k qubit ancillas (following Thm. 9.2), realisable by a sequential application of CNOT gates (see Fig. 9.1). We then provide a natural mapping of k-coherence into k-body entanglement, which can be accomplished by a second step which disentangles the qudit system — either by unitary transformations as in Fig. 9.1(a), or by one-way LOCC as in Fig. 9.1(b). The latter scheme reflects an operational scenario in which input agents are constrained to the resource theory of k-coherence, having at disposal only incoherent ancillas and incoherent operations as used in the first step, while output agents are constrained to the resource theory of entanglement, being bound to use LOCC as in the second step.

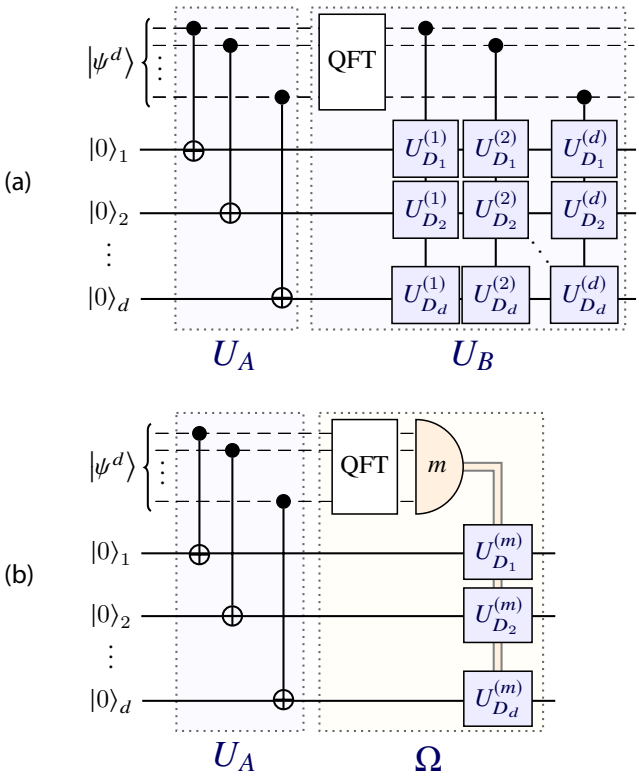


Figure 9.1: Schemes of two protocols to convert k-coherence into multipartite entanglement. Both protocols begin with the global unitary operation U_A which sequentially entangles each level of the qudit system in the state $|\psi^d\rangle$ with a corresponding ancillary qubit by generalised CNOT gates, resulting in a $k+1$-partite entangled state. One can then decouple the qudit system either (a) by a unitary transformation U_B, consisting of a Fourier transform and a disentangling unitary U_D, or (b) via a one-way LOCC operation Ω. Both protocols result in genuine k-partite entanglement between the ancillary qubits.

We illustrate the scheme for pure states, noting that it extends straightforwardly to mixed states. Let $|\Psi\rangle = |\psi^d\rangle \otimes |0\rangle^{\otimes d}$ be the state of the composite system consisting of the qudit initialised in $|\psi^d\rangle$ and d ancilla qubits in a reference pure state $|0\rangle^{\otimes d}$. Consider a unitary activation operation U_A which consists of a sequence of generalised CNOT gates $(\mathbb{1}_d - |i\rangle\langle i|) \otimes \mathbb{1}_2 + |i\rangle\langle i| \otimes \sigma_x$, with σ_x the Pauli x matrix, between the qudit and the i-th ancillary qubit. Explicitly, the sequence realises the unitary

$$U_A = \sum_{i=1}^{d} |i\rangle\langle i| \otimes \mathbb{1}_2^{\otimes i-1} \otimes \sigma_x \otimes \mathbb{1}_2^{\otimes d-i}, \tag{9.10}$$

which transforms the state $|\Psi\rangle = \sum_{i=1}^{d} c_i |i\rangle |0\rangle^{\otimes d}$ into

$$|\Psi'\rangle = U_A |\Psi\rangle = \sum_{i=1}^{d} c_i |i\rangle |\underline{2^{d-i}}\rangle . \tag{9.11}$$

To complete the protocol by mapping into k-partite entanglement among the qubit ancillas only, we now give two alternative approaches. Both methods begin by performing a quantum Fourier transform (QFT) $|j\rangle \to \frac{1}{\sqrt{d}} \sum_m e^{2\pi i jk/d} |m\rangle$ on the qudit only. Then, in the first approach (cf. Fig. 9.1(a)), we can apply a unitary

$$U_D = \sum_{j,m=1}^{d} e^{-2\pi i jm/d} |m\rangle \langle m| \otimes |\underline{2^{d-j}}\rangle \langle \underline{2^{d-j}}| \tag{9.12}$$

to effectively decouple the qudit and the ancilla qubits. This can be understood as the sequential application of d^2 controlled local operations $(\mathbb{1}_d - |m\rangle\langle m|) \otimes \mathbb{1}_2 + |m\rangle\langle m| \otimes U_{D_j}^{(m)}$, with control on the qudit and

$$U_{D_j}^{(m)} = |0\rangle \langle 0| + e^{-2\pi i jm/d} |1\rangle\langle 1| \tag{9.13}$$

acting on the j-th ancillary qubit. After the action of the QFT and U_D, which jointly define the global unitary U_B, the output will be the product state $U_B |\Psi'\rangle = |\Phi^+\rangle |\Psi''\rangle$, where $|\Phi^+\rangle = \sum_i \frac{1}{\sqrt{d}} |i\rangle$ is the maximally coherent state of the qudit, and $|\Psi''\rangle = \sum_{i=1}^{d} c_i |\underline{2^{d-i}}\rangle$ is a k-partite entangled state of the qubit ancillas.

An alternative approach (cf. Fig. 9.1(b)), which might lend itself to a more efficient implementation as it does not require global interactions, is to realise the decoupling of the qudit by an operation Ω consisting of one-way LOCC (see e.g. Ref. [256]). After performing the QFT, one can measure the qudit in the $\{|m\rangle\}$ basis and, depending on the measurement result m, apply the local unitary $U_D^{(m)} = \sum_{j=1}^{d} e^{-2\pi i jm/d} |\underline{2^{d-j}}\rangle \langle \underline{2^{d-j}}| = \bigotimes_j U_{D_j}^{(m)}$ to the remaining d qubits. We then obtain the final state $\Omega(|\Psi'\rangle) = |\Psi''\rangle$, which is exactly the same as the state of the qubits after the unitary transformation U_B from the previous approach.

[256] Piani & Adesso, PRA 2012

We can formalise the properties of the protocol as follows, casting the result in terms of mixed states in general.

Theorem 9.3. *Given the above conversion protocol, consisting of the activation unitary U_A and the decoupling operation (either via U_B or Ω), with $\rho' = U_A(\rho \otimes |0\rangle\langle 0|^{\otimes d})U_A^\dagger$ and $\rho'' = \Omega(\rho')$, the following statements are equivalent for $2 \le k \le d$:*

(i) $\mathrm{CN}(\rho) = k$;

(ii) $\mathrm{D_E}(\rho') = k + 1$;

(iii) $\mathrm{D_E}(\rho'') = k$.

Proof. The equivalence between (i) and (ii) is a direct application of Thm. 9.2. To prove that (i) implies (iii), we first consider pure states, and assume w.l.o.g. that the k nonzero coefficients are the first ones. Then the last $d-k$ qubits in $|\Psi''\rangle$ are in the initial product state $|0\rangle^{\otimes d-k}$, and we need to prove that the first k qubits are genuinely multipartite entangled. This is the case since any non-trivial subset of such qubits is mixed. That (iii) implies (i) can then be proven by observing, as in the proof of Thm. 9.2, that the isometry from $|\psi^d\rangle$ to $|\Psi''\rangle$ can be inverted, and using the just proven fact that (i) implies (ii). The extension to mixed states follows the exact same steps as in Thm. 9.2. ∎

9.2.1 *Quantitative relations*

We will first consider general distance-based resource quantifiers, as discussed in Sec. 5.4. Adapting a notation similar to the previous sections, we will use $\mathcal{I}^{(k-1)}$ to denote the set of states which are at most $k-1$-coherent, and similarly $\mathcal{M}^{(k-1)}$ to denote the set of $k-1$-producible states, i.e. states which are at most $k-1$-partite entangled.

Definition 9.4. Given a quasi-metric $D(\rho,\sigma)$ contractive under CPTP maps, we define the distance-based measure of k-partite entanglement as

$$E_D^{(k)}(\rho) = \inf_{\varsigma\in\mathcal{M}^{(k-1)}} D(\rho,\varsigma) \tag{9.14}$$

and the distance-based measure of k-coherence as

$$C_D^{(k)}(\rho) = \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(\rho,\sigma). \tag{9.15}$$

The distance-based quantifiers of k-coherence and k-partite entanglement have useful properties which will allow us to relate the two resource quantitatively. In particular, within the distance-based framework, we prove the following relation between the degree of coherence of a state ρ and the multipartite entanglement of the output states $\rho' = U_A(\rho\otimes|0\rangle\langle 0|^{\otimes d})U_A^\dagger$ and $\rho'' = \Omega(\rho')$ obtained from the conversion protocol of Thm. 9.3:

Proposition 9.5. *Let D be any distance contractive under CPTP maps. Then*

$$C_D^{(k)}(\rho) \geq E_D^{(k+1)}(\rho'),$$
$$C_D^{(k)}(\rho) \geq E_D^{(k)}(\rho'').$$

Proof. Let $\mathcal{B}^{(k)}$ denote the subset of $\mathcal{M}^{(k)}$ spanned by states of the form $|\Psi'\rangle = \sum_{i=1}^d c_i\,|i\rangle\,|\underline{2^{d-i}}\rangle$ as obtained from the first step of the conversion protocol in

Thm. 9.3. Similarly, let $\mathcal{R}^{(k-1)}$ denote the subset of $\mathcal{M}^{(k-1)}$ spanned by $|\Psi''\rangle = \sum_{i=1}^{d} c_i |\underline{2^{d-i}}\rangle$ as obtained from the second step of the same protocol.

Let us consider ρ' first. We have that

$$\begin{aligned} C_D^{(k)}(\rho) &= \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(\rho,\sigma) \\ &= \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(U_A\rho\otimes|0\rangle\langle 0|^{\otimes d}U_A^\dagger, U_A\sigma\otimes|0\rangle\langle 0|^{\otimes d}U_A^\dagger) \\ &= \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(\rho', U_A\sigma\otimes|0\rangle\langle 0|^{\otimes d}U_A^\dagger) \\ &= \inf_{\delta\in\mathcal{B}^{(k)}} D(\rho',\delta) \\ &\geq \inf_{\varsigma\in\mathcal{M}^{(k)}} D(\rho',\varsigma) \\ &= E_D^{(k+1)}(\rho'), \end{aligned} \tag{9.16}$$

where we have used Thm. 9.3, as well as the facts that $D(\rho\otimes\tau,\sigma\otimes\tau) = D(\rho,\sigma)$ and $D(U\rho U^\dagger, U\sigma U^\dagger) = D(\rho,\sigma)$ for any contractive distance D. Equality clearly holds when there exists $\delta \in \mathcal{B}^{(k)}$ such that

$$\inf_{\varsigma\in\mathcal{M}^{(k)}} D(\rho',\varsigma) = D(\rho',\delta). \tag{9.17}$$

An analogous argument holds for ρ'', One can either follow the steps above with the unitary transformation U_B, or note the contractivity of the distance D under the LOCC operation Ω and obtain:

$$\begin{aligned} C_D^{(k)}(\rho) &= \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(\rho,\sigma) \\ &= \inf_{\sigma\in\mathcal{I}^{(k-1)}} D(\rho', U\sigma\otimes|0\rangle\langle 0|^{\otimes d}U^\dagger) \\ &\geq \inf_{\sigma\in\mathcal{I}^{(k-1)}} D\left(\Omega(\rho'), \Omega\left(U\sigma\otimes|0\rangle\langle 0|^{\otimes d}U^\dagger\right)\right) \\ &= \inf_{\delta\in\mathcal{R}^{(k-1)}} D(\rho'',\delta) \\ &\geq \inf_{\varsigma\in\mathcal{M}^{(k-1)}} D(\rho'',\varsigma) \\ &= E_D^{(k)}(\rho''). \end{aligned} \tag{9.18}$$

■

The amount of k-coherence present in the initial state thus places quantitative constraints on the multipartite entanglement one can obtain from it.

Of particular interest will be the choice $D(\rho,\sigma) = 1 - F(\rho,\sigma)$, that is, the geometric measures as considered in Sec. 5.4.2). We will use the notation $E_G^{(k)}$ and $C_G^{(k)}$ for these quantities. In particular, recall from Thm. 5.26 that

$$E_G^{(k)}(\rho) = 1 - \max\left\{ \sum_i p_i E_G^{(k)}(|\psi_i\rangle\langle\psi_i|) \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}. \tag{9.19}$$

Remarkably, under the geometric quantifiers, the k-coherence of any state and the converted $k+1$-partite entanglement are in fact *equal*. The result relies on the following lemma, which shows that for any pure state $|\Psi'\rangle$ obtained from the conversion protocol, it suffices to optimise the distance-based quantifier of $k+1$-partite entanglement over the set of k-producible output states of the protocol, instead of the whole set of k-producible states.

Lemma 9.6. *Given a state of the form*

$$|\psi\rangle = \sum_{i=1}^{d} c_i \, |i\rangle \, |\underline{2^{d-i}}\rangle \tag{9.20}$$

where we can take $|c_1| \geq |c_2| \geq \ldots \geq |c_d|$ without loss of generality, the closest k-producible state with respect to the fidelity-based geometric measure of entanglement can be chosen as a state $|\psi^c\rangle \in \mathcal{B}^{(k)}$.
In other words, there exists $|\psi^c\rangle \in \mathcal{B}^{(k)}$ such that

$$\max_{|\varsigma\rangle \in \mathcal{M}^{(k)}} F(|\psi\rangle, |\varsigma\rangle) = \max_{|\sigma\rangle \in \mathcal{B}^{(k)}} F(|\psi\rangle, |\sigma\rangle) = F(|\psi\rangle, |\psi^c\rangle). \tag{9.21}$$

The proof is deferred to the Appendix.

The above Lemma finally allows us to prove the equality between the geometric quantifiers for k-coherence and $k+1$-partite entanglement in the first step of the conversion protocol.

Theorem 9.7. *Given ρ and the transformed state ρ' as described in the conversion protocol, we have*

$$C_G^{(k)}(\rho) = E_G^{(k+1)}(\rho'). \tag{9.22}$$

Proof. The proof follows the methods of Refs. [120] and [61].

[120] Streltsov et al., NJP 2010

[61] Streltsov et al., PRL 2015

This result amounts to showing that there exists a state $\chi \in \mathcal{B}^{(k)}$ which is the closest k-producible state to ρ' with respect to the distance given by $D(\rho, \sigma) = 1 - F(\rho, \sigma)$. In other words, we need to find $\chi \in \mathcal{B}^{(k)}$ s.t. $E_G^{(k+1)}(\rho') = 1 - F(\rho', \chi)$.

First, let $\{p_i, |\phi_i\rangle\}$ be the optimal convex decomposition of ρ' which realises the infimum in the convex roof extension of $E_G^{(k+1)}$, that is,

$$E_G^{(k+1)}(\rho') = \sum_i p_i \, E_G^{(k+1)}(|\phi_i\rangle). \tag{9.23}$$

Define $\{|\xi_i\rangle\} \in \mathcal{M}^{(k)}$ to be the closest k-producible states to each of the states $\{|\phi_i\rangle\}$ with respect to the fidelity-based distance, that is,

$$E_G^{(k+1)}(|\phi_i\rangle) = 1 - F(|\xi_i\rangle, |\phi_i\rangle) \quad \forall \, i. \tag{9.24}$$

Now, notice that the support of ρ' is spanned by states of the form

$$|\psi'\rangle = \sum_i p_i \, |i\rangle \, |\underline{2^{d-i}}\rangle \; \in \mathcal{B}^{(k)}, \tag{9.25}$$

which means that any pure state in a convex decomposition of ρ' can be expressed as a complex linear combination of such $|\psi'\rangle$. In particular, each of the states $\{|\phi_i\rangle\}$ can be written as

$$|\phi_i\rangle = \sum_j r_j \, |j\rangle \, |\underline{2^{d-j}}\rangle \, . \tag{9.26}$$

By Lemma 9.6, the closest k-producible state to each $|\phi_i\rangle$ is a state $|\xi_i\rangle \in \mathcal{B}^{(k)}$, and so we have

$$|\xi_i\rangle = \sum_{j=1}^{k-1} e_j \, |j\rangle \, |\underline{2^{d-j}}\rangle \, . \tag{9.27}$$

Now, define

$$\chi = \sum_j q_j \, |\xi_j\rangle \, \langle \xi_j| \tag{9.28}$$

with coefficients

$$q_j = p_j \frac{1 - E_G^{(k+1)} \left(|\phi_j\rangle \right)}{1 - E_G^{(k+1)} \left(\rho' \right)}. \tag{9.29}$$

Since each $|\xi_i\rangle$ is in $\mathcal{B}^{(k)}$, we get that $\chi \in \mathcal{B}^{(k)}$. Recalling that $\sqrt{F}$ satisfies the so-called strong joint concavity property [25], defined as [25] Nielsen & Chuang, 2011

$$\sqrt{F\left(\sum_i a_i \, |\alpha_i\rangle \, \langle \alpha_i| \, , \, \sum_i b_i \, |\beta_i\rangle \, \langle \beta_i| \right)} \geq \sum_i \sqrt{a_i \, b_i \, F\left(|\alpha_i\rangle \, , |\beta_i\rangle \right)}, \tag{9.30}$$

we get

$$\begin{aligned} \sqrt{F(\rho', \chi)} &\geq \sum_j \sqrt{p_j \, q_j \, F\left(|\phi_j\rangle \, , |\xi_j\rangle \right)} \\ &= \sum_j \sqrt{p_j^2 \frac{F\left(|\phi_j\rangle \, , |\xi_j\rangle \right)^2}{1 - E_G^{(k+1)} \left(\rho' \right)}} \\ &= \sqrt{1 - E_G^{(k+1)} \left(\rho' \right)} \end{aligned} \tag{9.31}$$

which follows by Eqs. (9.24) and (9.23). This gives

$$E_G^{(k+1)} \left(\rho' \right) \geq 1 - F(\rho', \chi) \tag{9.32}$$

and since $\chi \in \mathcal{M}^{(k)}$ is a convex combination of k-producible pure states, by definition we also have

$$E_G^{(k+1)} \left(\rho' \right) \leq 1 - F(\rho', \chi) \tag{9.33}$$

and so χ is the closest k-producible state to ρ'. But since $\chi \in \mathcal{B}^{(k)}$ by construction, Prop. 9.5 gives

$$C_G^{(k)}(\rho) = E_G^{(k+1)}(\rho') \tag{9.34}$$

as required. ■

The above result has implications for the quantification of $k+1$-partite entanglement, since for any state of the form ρ' as obtained from the conversion protocol in Thm. 9.3, the entanglement can be quantified by considering the quantification of k-coherence instead. It has been recently shown that optimisation over sets of k-coherent states can be expressed as semidefinite programs [104], and together with the fact that the computation of the fidelity function can be cast as an SDP as well [221,222], we have that the geometric measure of $k+1$-partite entanglement of any state ρ' can be quantified as an SDP. This reduces the very complex problem of k-partite entanglement to a much simpler optimisation with semidefinite constraints, and includes among others the class of maximally correlated states considered previously.

[104] Ringbauer et al., preprint 2017

[221] Watrous, Theory Comput. 2009

[222] Watrous, Chic. J. Th. Comp. Sci. 2013

10 Polynomial measures of entanglement and monogamy

The final chapter of the thesis deals with the problem of quantifying a class of multipartite entanglement measures based on the so-called *polynomial invariants*. Although the original measure of this kind — the concurrence [261] — was introduced in order to characterise the operationally-motivated entanglement of formation, the theory of polynomial invariants and their applications to entanglement quantification has since diverged from the strictly operational considerations and has been investigated in purely mathematical terms. Many of such measures are relatively simple to quantify for pure states by construction, but their computation for mixed states again relies on the concept of the convex roof, which is notoriously difficult to evaluate in general cases. We begin by showing that under certain conditions this problem becomes trivial — precisely, we prove that polynomial entanglement measures of degree 2 are independent of the choice of pure-state decomposition of a mixed state when the latter has only one pure unentangled state in its support. This allows for the analytical evaluation of convex roof extended entanglement measures in classes of rank-two states obeying such condition. As an explicit application of this approach, we investigate possible generalisations of the Coffman-Kundu-Wootters monogamy inequality from three to four qubits, accounting for multipartite entanglement in addition to the bipartite terms. We show that the most natural extension of the inequality does not hold in general, and our results allow us to characterise exactly the violations of this inequality.

This chapter is based on results published in [BR1,BR2,BR12].

[261] Hill & Wootters, PRL 1997

[BR1] Regula & Adesso, PRL 2016

[BR2] Regula et al., PRA 2016

[BR12] Regula et al., PRL 2014

10.1 A GEOMETRIC APPROACH TO ENTANGLEMENT QUANTIFICATION

In the characterisation of properties of multipartite entanglement, the formalism of LOCC is often extended to the so-called stochastic LOCC (SLOCC), representing probabilistic LOCC transformations. This class admits a simplified mathematical structure: given two pure state $|\psi\rangle$, $|\phi\rangle$ consisting of n systems of dimension d, the transformation $|\psi\rangle \to |\phi\rangle$ is possible with SLOCC if and only if there exist local invertible linear operators $L_1, \ldots, L_n \in \mathbb{C}^{d\times d}$ such that $|\phi\rangle = L_1 \otimes \cdots \otimes L_n |\psi\rangle$ [51,260]. Disregarding normalisation, we can take each L_i to have determinant one, and thus belong to the special linear group $\mathrm{SL}(d, \mathbb{C})$.

[51] Eltschka & Siewert, JPA 2014

[260] Dür et al., PRA 2000

A polynomial invariant is then a polynomial function in the coefficients of a pure state $|\psi\rangle$ which is invariant under SLOCC [51]. For a system of n d-

dimensional system, any such a polynomial invariant P of homogeneous degree h can be written as

$$P_h\left(c\,L\,|\psi\rangle\right) = c^h\,P_h\left(|\psi\rangle\right), \tag{10.1}$$

for any $c > 0$ and an invertible linear operator $L \in \mathrm{SL}(d,\mathbb{C})^{\otimes n}$ representing a SLOCC transformation. Then, one can take an appropriate power p of the absolute value of any polynomial invariant P to construct an entanglement measure on pure states,

$$E_h^p(|\psi\rangle) \coloneqq |P_h(|\psi\rangle)|^p. \tag{10.2}$$

For n-qubit states ($d = 2$), the above expression defines a valid entanglement monotone provided $hp \leq 4$ [262,263], and in a slight abuse of terminology we will refer to all such quantities as polynomial measures of entanglement. The concept of polynomial invariants can be used to obtain entanglement measures for different types of entanglement in any number of qubits [264,265] and qudits [266]. Two particularly common monotones obtained in this way are the *concurrence* for two qubits [50,261] and the *three-tangle* for three qubits [267]. Notably, the convex roof extension of the concurrence can be quantified exactly for any system of two qubits [50], although analytical solutions for the convex roof of the three-tangle have been found only in very special cases [268–275] and the first insights into the solutions of the even more complicated case of the convex roof of four-qubit polynomial measures have been obtained only recently [276].

One can relate the properties of entanglement measures based on polynomial invariants with the geometric representation of quantum states on a hypersphere [268,277]. Here, we will limit ourselves to the case of density matrices ρ of rank 2 and the familiar Bloch sphere. Any such state can always be written in its spectral decomposition into the two eigenvectors $\{|\phi_0\rangle, |\phi_1\rangle\}$ corresponding to the two non-zero eigenvalues $\{\lambda_0, \lambda_1\}$ as $\rho = \lambda_0\,|\phi_0\rangle\langle\phi_0| + \lambda_1\,|\phi_1\rangle\langle\phi_1|$, with $\lambda_0 + \lambda_1 = 1$. We can then visualise the support of ρ as the Bloch sphere with the two eigenvectors as poles, and any state $|\omega\rangle$ on the surface of the sphere as corresponding (up to normalisation) to a linear combination of the two eigenvectors, $|\omega\rangle = |\phi_0\rangle + \omega\,|\phi_1\rangle$ for some $\omega \in \mathbb{C}$, with $|\phi_1\rangle$ itself obtained in the limit $\omega \to \infty$. This, in fact, can be understood as the stereographic projection of the Bloch sphere, establishing a bijection between the sphere and the extended complex plane $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ (see Fig. 10.1). To avoid ambiguities, we will use the notation $\boldsymbol{\omega}$ to denote the Bloch vector $\boldsymbol{\omega} \in \mathbb{R}^3$ corresponding to the given state $|\omega\rangle \in \mathrm{span}(\rho)$, with $\boldsymbol{\phi_0} = (0,0,1)^T$ and $\boldsymbol{\phi_1} = \boldsymbol{\infty} = (0,0,-1)^T$.

Let E_h^1 denote a polynomial entanglement measure based on a polynomial invariant P_h of degree h as in Eq. (10.2), with $p = 1$ for simplicity. Let $\{|\phi_0\rangle, |\phi_1\rangle\}$ be the two eigenvectors of a rank-2 state ρ corresponding to its non-zero eigenvalues. Then, the entanglement E_h^1 of any state $|\omega\rangle$ in the support of ρ can be expressed as the absolute value of a polynomial in one complex variable ω:

$$E_h^1\left(\,|\phi_0\rangle + \omega\,|\phi_1\rangle\,\right) = N\prod_{i=1}^{h}|\omega - z_i|, \tag{10.3}$$

[262] Verstraete et al., PRA 2003
[263] Eltschka et al., PRA 2012
[264] Osterloh & Siewert, PRA 2005
[265] Djokovic & Osterloh, JMP 2009
[266] Gour & Wallach, PRL 2013
[50] Wootters, PRL 1998
[261] Hill & Wootters, PRL 1997
[267] Coffman et al., PRA 2000
[268] Lohmayer et al., PRL 2006
[269] Eltschka et al., NJP 2008
[270] Jung et al., PRA 2009
[271] Jung et al., PRA 2009
[272] He et al., Commun. Theor. Phys. 2011
[273] Siewert & Eltschka, PRL 2012
[274] Viehmann et al., Appl. Phys. B 2012
[275] Eltschka & Siewert, PRL 2012
[276] Jung & Park, Quantum Inf Process 2015
[277] Osterloh et al., PRA 2008

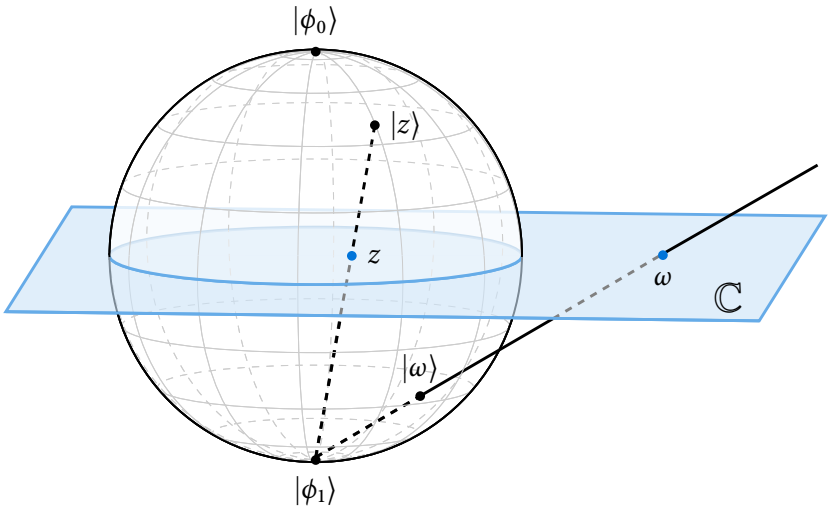


Figure 10.1: The Bloch sphere as the stereographic projection onto the complex plane (shaded blue plane). Every pure state $|\omega\rangle$ with a Bloch vector on the sphere corresponds to the point $\omega \in \mathbb{C}$ which lies at the intersection of the plane with the line containing the Bloch vectors of $|\omega\rangle$ and the projection point $|\phi_1\rangle$. States in the upper hemisphere are mapped to the region of the plane inside the sphere, and states in the lower hemisphere are mapped outside of it. The projection point itself, corresponding to the state $|\phi_1\rangle$, is mapped to a point at infinity.

where N is a normalisation constant, and $\{z_1, \ldots, z_h\}$ are the roots of the polynomial E_h^1 in $\hat{\mathbb{C}}$, defined by

$$E_h^1\Big(|\phi_0\rangle + z_i\,|\phi_1\rangle\Big) = 0. \tag{10.4}$$

The convex hull of the points on the Bloch sphere corresponding to these coefficients $\{z_i\}$ defines the *zero polytope* [268,277]. This concept provides a useful representation of the set of all separable states within the considered Bloch sphere, since the entanglement as measured by the convex roof extension of E_h^1 has to vanish for any convex combination of pure separable states (i.e., for any mixed state inside the zero polytope), but will be strictly positive outside of their convex hull. Various methods for constructing bounds to polynomial entanglement measures in mixed states rely on finding states within the zero polytope and using them to form suitable convex combinations with states outside of it [114,278,279].

A crucial observation is that this expression is in fact proportional to the Euclidean distances between the Bloch vectors of the qubits. We make this precise in the following lemma.

Lemma 10.1. *Any polynomial entanglement measure can be computed for*

[268] Lohmayer et al., PRL 2006

[277] Osterloh et al., PRA 2008

[114] Lewenstein & Sanpera, PRL 1998

[278] Acín et al., PRL 2001

[279] Rodriques et al., PRA 2014

a pure state $|\omega\rangle = |\phi_0\rangle + \omega\,|\phi_1\rangle$ in the support of a rank-2 state ρ as

$$E_h^1\left(|\phi_0\rangle + \omega\,|\phi_1\rangle\right) = N_\rho \prod_i \|\boldsymbol{\omega} - \boldsymbol{z}_i\| \tag{10.5}$$

where $\boldsymbol{\omega}$ and $\boldsymbol{z}_i$ are the Bloch vectors corresponding to $|\omega\rangle$ and each $|z_i\rangle$, respectively, and N_ρ is a normalisation constant, which can depend on the given state ρ.

Proof. We begin by explicitly normalising the state $|\omega\rangle$ and additionally dividing the expression (10.3) by the normalisation factors of all the states $\{|z_i\rangle\}$, to obtain

$$\frac{1}{\prod_i \sqrt{1+|z_i|^2}} E_h^1\left(\frac{|\phi_0\rangle + \omega\,|\phi_1\rangle}{\sqrt{1+|\omega|^2}}\right) = \frac{N \prod_i |\omega - z_i|}{\sqrt{1+|\omega|^2}^{\,h} \prod_i \sqrt{1+|z_i|^2}}. \tag{10.6}$$

Let us now recall the definition of the chordal distance in the stereographic projection, that is, the Euclidean distance between points $\boldsymbol{x}_k$ on the sphere S^2 corresponding to their stereographic projection $\xi_k \in \hat{\mathbb{C}}$ [280]:

[280] Carathéodory, 1954

$$\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = \frac{2\,|\xi_1 - \xi_2|}{\sqrt{1+|\xi_1|^2}\sqrt{1+|\xi_2|^2}}, \tag{10.7}$$

with

$$\|\boldsymbol{x}_1 - \boldsymbol{\infty}\| = \frac{2}{\sqrt{1+|\xi_1|^2}}. \tag{10.8}$$

As we discussed, the Bloch sphere is precisely an (inverse) stereographic projection of $\hat{\mathbb{C}}$ onto a sphere S^2, and so we identify the right-hand side of Eq. (10.6) with a product of distances between pure-state Bloch vectors defined by their respective complex coefficients. This leads to the simple relation

$$\begin{aligned} E_h^1\left(\frac{|\phi_0\rangle + \omega\,|\phi_1\rangle}{\sqrt{1+|\omega|^2}}\right) &= \frac{N}{2^h} \prod_i \sqrt{1+|z_i|^2} \prod_i \|\boldsymbol{\omega} - \boldsymbol{z}_i\| \\ &= N_\rho \prod_i \|\boldsymbol{\omega} - \boldsymbol{z}_i\| \end{aligned} \tag{10.9}$$

where we defined $N_\rho \coloneqq \frac{N}{2^h} \prod_i \sqrt{1+|z_i|^2}$. ∎

We remark that the constant N_ρ only depends on the support of ρ, and as such it will be the same for any two states sharing the same support.

10.1.1 *One-root states*

We begin by noting that an entanglement measure of degree 2 can be obtained simply by choosing $E_h^{2/h}$. Any such quantity is then a valid entanglement monotone for both pure and mixed states [262,263], and the particular choice of degree

[262] Verstraete et al., PRA 2003
[263] Eltschka et al., PRA 2012

2 was found to be particularly useful as it enjoys some simplified scaling properties [274]. We will hereafter make this choice of the degree of the measure.

[274] Viehmann et al., Appl. Phys. B 2012

We consider now a particular case of a state ρ such that there is only one pure product state in its support, that is, the equation $E_h^{2/h}(|\phi_0\rangle + z_i|\phi_1\rangle) = 0$ has only one unique root z with multiplicity h. We will refer to such states as *one-root* states. For any such state, we can then use Lemma 10.1 to express the value of the polynomial entanglement measure as

$$E_h^{2/h}(\rho) = N_\rho \min\left\{ \sum_i p_i \|\psi_i - \boldsymbol{z}\|^2 \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}. \tag{10.10}$$

Geometrically, any state ρ can be identified with a Bloch vector $\boldsymbol{g}$ in the Bloch ball, and any pure-state decomposition of such a state gives a set of Bloch vectors $\{\boldsymbol{a_i}\}$ on the sphere and weights $\{\alpha_i\}$ such that $\sum_i \alpha_i \boldsymbol{a_i} = \boldsymbol{g}$. We will now use this correspondence to establish a surprising property: the value of the convex roof optimisation in (10.10) is in fact *independent* of the chosen pure state decomposition, and no optimisation is necessary. The result is based on the following geometric argument.

Theorem 10.2. *Consider an n-sphere with radius R and center located at $\mathbf{o}$. We will indicate by $\{\lambda_i, \mathbf{p}_i\}$ a finite set of points $\mathbf{p}_i \in \mathbb{R}^{n+1}$ on the sphere with corresponding weights λ_i, normalised so that $\sum_i \lambda_i = 1$. Let us choose a particular point $\mathbf{z}$ on the sphere and denote by $\{\alpha_i, \mathbf{a}_i\}$ a set such that all points $\{\mathbf{a}_i\}$ are equidistant from $\mathbf{z}$, that is $\|\mathbf{z} - \mathbf{a}_i\| = \|\mathbf{z} - \mathbf{a}_k\| \;\forall\, i, k$, according to the standard Euclidean distance. Let $\mathbf{g}$ denote the barycenter of the family of points, that is, $\mathbf{g} = \sum_i \alpha_i \mathbf{a}_i$. Then, for any other set of points $\{\beta_j, \mathbf{b}_j\}$ which lie on the same sphere and share the same barycenter $\mathbf{g}$, the following holds:*

$$\textstyle\sum_j \beta_j \|\mathbf{z} - \mathbf{b}_j\|^2 = \|\mathbf{z} - \mathbf{a}_l\|^2, \qquad \forall\, l. \tag{10.11}$$

Proof. Apollonius' formula [281] says that for any set of points $\{\alpha_i, \mathbf{a}_i\}$ with barycenter $\mathbf{g}$ and any chosen point $\mathbf{z} \in \mathbb{R}^{n+1}$ we can write

[281] Berger, 1987

$$\sum_i \alpha_i \|\mathbf{z} - \mathbf{a}_i\|^2 = \|\mathbf{z} - \mathbf{g}\|^2 + \sum_i \alpha_i \|\mathbf{g} - \mathbf{a}_i\|^2, \tag{10.12}$$

but since we chose $\{\alpha_i, \mathbf{a}_i\}$ to be equidistant from $\mathbf{z}$, we get

$$\|\mathbf{z} - \mathbf{g}\|^2 = \|\mathbf{z} - \mathbf{a}_l\|^2 - \sum_i \alpha_i \|\mathbf{g} - \mathbf{a}_i\|^2, \tag{10.13}$$

for any chosen $\mathbf{a}_l$. Applying the same formula to any other set of points $\{\beta_j, \mathbf{b}_j\}$ with the same barycenter $\mathbf{g}$ gives:

$$\begin{aligned} \sum_j \beta_j \|\mathbf{z} - \mathbf{b}_j\|^2 &= \|\mathbf{z} - \mathbf{g}\|^2 + \sum_j \beta_j \|\mathbf{g} - \mathbf{b}_j\|^2 \\ &= \|\mathbf{z} - \mathbf{a}_l\|^2 - \sum_i \alpha_i \|\mathbf{g} - \mathbf{a}_i\|^2 + \sum_j \beta_j \|\mathbf{g} - \mathbf{b}_j\|^2, \end{aligned} \tag{10.14}$$

by Eq. (10.13). Since the points $\{\alpha_i, \mathbf{a}_i\}$ lie on the n-sphere, we get

$$R^2 = \|\mathbf{a}_i - \mathbf{o}\|^2 = \|\mathbf{a}_i\|^2 - 2\langle \mathbf{a}_i, \mathbf{o}\rangle + \|\mathbf{o}\|^2 \tag{10.15}$$

for each $\mathbf{a}_i$, where $\langle\cdot,\cdot\rangle$ denotes the standard Euclidean inner product. We can average over this expression with the weights $\{\alpha_i\}$ to obtain

$$R^2 = \Big(\sum_i \alpha_i \|\mathbf{a}_i\|^2\Big) - 2\left\langle \sum_i \alpha_i \mathbf{a}_i,\, \mathbf{o}\right\rangle + \|\mathbf{o}\|^2, \tag{10.16}$$

that is, $\sum_i \alpha_i \|\mathbf{a}_i\|^2 = R^2 + 2\langle \mathbf{g}, \mathbf{o}\rangle - \|\mathbf{o}\|^2$. Since the points $\{\beta_j, \mathbf{b}_j\}$ lie on the same n-sphere and share the same barycenter $\mathbf{g}$, we can conclude that $\sum_i \alpha_i \|\mathbf{a}_i\|^2 = \sum_j \beta_j \|\mathbf{b}_j\|^2$, and therefore $\sum_i \alpha_i \|\mathbf{g} - \mathbf{a}_i\|^2 = \sum_j \beta_j \|\mathbf{g} - \mathbf{b}_j\|^2$. Using this in Eq. (10.14) gives the final result announced in Eq. (10.11). ■

The evaluation of $E(\rho)$ is thus made easy and can be carried out exactly in any decomposition. That is, the measure $E_h^{2/h}$ is an affine function on the whole Bloch sphere,

$$E_h^{2/h}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i E_h^{2/h}(|\psi_i\rangle). \tag{10.17}$$

It is particularly instructive to consider a decomposition of ρ such that all pure states $\{|\psi_i\rangle\}$ with Bloch vectors $\{\mathbf{a}_i\}$ (in the notation of Theorem 10.2) lie on the secant plane equidistant from the root point $\mathbf{z}$, see Fig. 10.2. The value of $E(\rho)$ then corresponds to the squared distance from any of these points $\mathbf{a}_i$ to $\mathbf{z}$, according to Lemma 10.1. But since this can be done for any other state with a Bloch vector lying in the same plane, the equidistant plane is in fact a plane of constant entanglement.

Let us now introduce the radial state ρ_c whose vector $\mathbf{c}$ is at the centre of the secant plane, i.e. at the intersection between the plane and the Bloch sphere diameter joining $\mathbf{z}$ with the antipodal point $\mathbf{z}'$. The latter point corresponds to the pure state $|z'\rangle$ with maximal entanglement E on the sphere. As ρ and ρ_c are on the same secant plane, one has $E_h^{2/h}(\rho) = E_h^{2/h}(\rho_c)$. The latter can be evaluated by exploiting the affinity of $E_h^{2/h}$, Eq. (10.17), and taking the decomposition $\rho_c = \frac{1}{2}\|\mathbf{c} - \mathbf{z}'\|\,|z\rangle\langle z| + \frac{1}{2}\|\mathbf{c} - \mathbf{z}\|\,|z'\rangle\langle z'|$. Since $E_h^{2/h}(|z\rangle) = 0$, we finally get, for any one-root state ρ in the Bloch sphere, the closed formula

$$\begin{aligned} E_h^{2/h}(\rho) = E_h^{2/h}(\rho_c) &= \frac{1}{2}\|\mathbf{c} - \mathbf{z}\|\; E_h^{2/h}(|z'\rangle) \\ &= \frac{1}{2}\|\rho_c - |z\rangle\langle z|\|_1\; E_h^{2/h}(|z'\rangle), \end{aligned} \tag{10.18}$$

where we have exploited the fact that the trace distance between 2-dimensional density matrices is equal to the Euclidean distance between their Bloch vectors [53]. With some simple geometry, this can be computed as

[53] Bengtsson & Życzkowski, 2007

$$E(\rho) = \frac{1}{2}\big|1 - \langle z|\rho|z\rangle + \langle z'|\rho|z'\rangle\big|\; E(|z'\rangle). \tag{10.19}$$

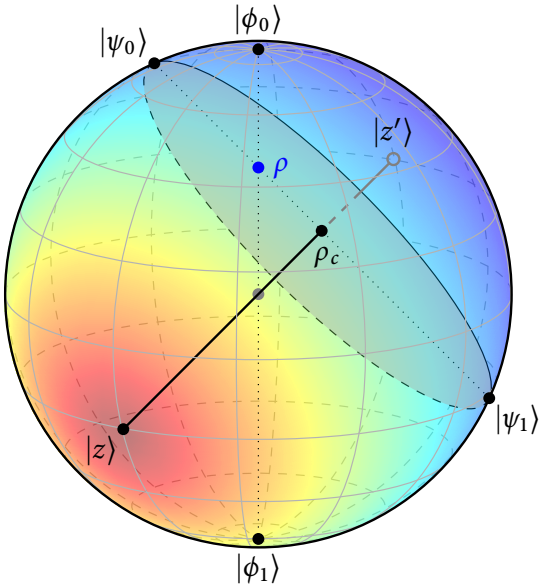


Figure 10.2: The entanglement of a rank-2 one-root state ρ, with detail of an equidistant decomposition $\rho = p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|$, visualised on the Bloch sphere with the polar states $\{|\phi_0\rangle, |\phi_1\rangle\}$ being the eigenvectors of ρ. The flat shaded surface is the plane of constant entanglement for ρ and the radial state ρ_c. The coloring on the sphere corresponds to the value of entanglement, which amounts to the distance from the root state $|z\rangle$, with red meaning vanishing entanglement, and blue meaning maximum entanglement, reached on the antipodal state $|z'\rangle$.

Remark 10.3. Similar methods can be employed to show that, for any rank-2 density matrix, the convex roof of polynomial measures of any degree can be quantified exactly when the measure has two unique roots with equal multiplicity [BR3]. This provides a generalisation of the results discussed herein, but these cases are not relevant to the applications we wish to discuss, so we have decided not to include them for the sake of brevity.

10.1.2 *Examples*

The result can be readily applied to any suitable measure. For two qubits, a polynomial measure of degree 2 is the concurrence C [261], for which our result can be verified to coincide with the known general solutions from [50,261]. A more complicated case is that of three qubits, where the representative choice of a tripartite entanglement measure is the so-called three-tangle τ, and its square root defines a valid quantifier of degree 2 [267,274]. Although this quantity found a fundamental application in describing the monogamy relations of entanglement [267], at present no closed solution exists in general for its evaluation on mixed states, beyond a few special cases [268–275]. Using our results, one can construct examples of states which have not been characterised before [BR1].

[BR3] Regula & Adesso, PRA 2016
[261] Hill & Wootters, PRL 1997
[50] Wootters, PRL 1998
[267] Coffman et al., PRA 2000
[274] Viehmann et al., Appl. Phys. B 2012
[268] Lohmayer et al., PRL 2006
[269] Eltschka et al., NJP 2008
[270] Jung et al., PRA 2009
[271] Jung et al., PRA 2009
[272] He et al., Commun. Theor. Phys. 2011
[273] Siewert & Eltschka, PRL 2012
[275] Eltschka & Siewert, PRL 2012
[BR1] Regula & Adesso, PRL 2016

Beyond specific examples, one can wonder whether a more systematic characterisation of one-root three-qubit states is possible, so as to gauge the relevant range of applicability of our exact results. The answer is affirmative. Notice that every rank-2 three-qubit state ρ can be purified to a four-qubit state $|\Psi\rangle \in \mathbb{C}^{16}$, and conversely the set of marginals obtained by tracing out one qubit from arbitrary four-qubit pure states $|\Psi\rangle$ completely characterises the set of rank-2 three-qubit states ρ. We can then aim to identify the one-root three-qubit states in terms of their four-qubit purifications. To this end, we recall that while pure four-qubit states have an infinite number of SLOCC-inequivalent classes [260], that is, subsets of states which cannot be transformed into one another by performing SLOCC operations, from the point of view of classifying their entanglement properties they can in fact be conveniently grouped into 9 classes [282,283]. Each of these classes then forms a subset $\Upsilon_\mu \subset \mathbb{C}^{16}$ (for $\mu = 1, \ldots, 9$) represented by a generating family $|G_\mu^{a,b,c,d}\rangle$ (dependent on at most four continuous complex parameters a, b, c, d), such that all the states $|\Psi_\mu\rangle \in \Upsilon_\mu$ belonging to the μ^{th} class are constructed as $|\Psi_\mu\rangle = L\,|G_\mu\rangle \,/\, \|L\,|G_\mu\rangle\,\|$, where $L \in \mathrm{SL}(2,\mathbb{C})^{\otimes 4}$ is a SLOCC operation. The union of all the nine classes $\bigcup_{\mu=1}^{9} \Upsilon_\mu$ covers the Hilbert space of *all* four-qubit pure states, up to permutations of the qubits.

[260] Dür et al., PRA 2000

[282] Verstraete et al., PRA 2002

[283] Chterental & Djokovic, 2007

We now make a useful observation. When two pure states are SLOCC-equivalent, the supports of their corresponding reduced subsystems are spanned by SLOCC-equivalent bases, which means that all states in the reduced supports are related by an invertible linear transformation [284,285]. Since for any two SLOCC-equivalent states the polynomial entanglement measure E_h^p either vanishes on both or is strictly nonzero on both [274], we have that SLOCC operations preserve the number of product states in the support of the reduced subsystems. In other words, the number of roots in the zero polytope for the marginals of four-qubit states is a SLOCC invariant.

[284] Chen & Chen, PRA 2006

[285] Chen & Chen, PRA 2006

[274] Viehmann et al., Appl. Phys. B 2012

It then suffices to check the three-qubit marginals of the generators $|G_\mu\rangle$ to look for the one-root property. This can be done analytically after some straightforward algebra, and we find as a result that three of the nine classes of four-qubit pure states have three-qubit marginals which can enjoy the one-root property, thus establishing whole classes of three-qubit one-root states. We stress that our results entail that for all these rank-2 mixed states ρ we can exactly calculate the convex roof extended entanglement measure $\sqrt{\tau}$ thanks to Theorem 10.2.

Explicitly, the classes whose marginals are generally one-root are: class 5 (tracing out qubit 2 or qubit 4 only), and classes 7 and 8 (tracing out qubit 2, 3, or 4 only). The corresponding sets of three-qubit one-root states are given

therefore by

$$\rho^{L}_{\mu,k} = \mathrm{Tr}_k\left[\frac{L\,|G^a_\mu\rangle\langle G^a_\mu|\,L^\dagger}{\mathrm{Tr}(L\,|G^a_\mu\rangle\langle G^a_\mu|\,L^\dagger)}\right], \quad \forall\ L \in \mathrm{SL}(2,\mathbb{C})^{\otimes 4}, \\ \text{with}\ \ k \in \begin{cases} \{2,4\}, & \text{if } \mu = 5; \\ \{2,3,4\}, & \text{if } \mu = 7 \text{ or } 8. \end{cases} \tag{10.20}$$

For completeness, we report the relevant (unnormalised) generators: $|G^a_5\rangle = a(|0000\rangle + |0101\rangle + |1010\rangle + |1111\rangle) + i\,|0001\rangle + |0110\rangle - i\,|1011\rangle$, $|G_7\rangle = |0000\rangle + |0101\rangle + |1000\rangle + |1110\rangle$, $|G_8\rangle = |0000\rangle + |1011\rangle + |1101\rangle + |1110\rangle$, where $a \in \mathbb{C}$ with $\mathrm{Re}(a) \geq 0$. More classes of four-qubit states with one-root marginals can be obtained as special cases of the states in other SLOCC classes; of particular importance will be the class $|G^{a,b,c}_2\rangle$, which has one-root marginals for certain values of the parameters a, b, c. This will be addressed in more detail in the next section.

The square root of three-tangle for all the states in Eq. (10.20) is given exactly by Eq. (10.18); if one prefers, it can also be evaluated numerically in any convex decomposition (e.g. the spectral one), with no optimisation required. We finally note that many entanglement bounds for convex roof extended measures will be tight on the one-root states because of their simplified properties. For instance, bounds such as the best separable approximation for two qubits [114], the best W approximation for three qubits [278], and the generalised best zero-E approximation [279] are based on finding a convex decomposition for an arbitrary state ρ in terms of states with vanishing entanglement and at most one state with non-vanishing entanglement. However, for one-root states, such a decomposition is possible only in one way: that is, into a pair formed by the root $|z\rangle$ and some other state $|\omega\rangle$; hence the entanglement of ρ is trivially given by $E(|\omega\rangle)$ with the corresponding weight.

[114] Lewenstein & Sanpera, PRL 1998

[278] Acín et al., PRL 2001

[279] Rodriques et al., PRA 2014

The main issue with our characterisation of the SLOCC classes discussed above is that any generic four-qubit pure state in fact belongs to the class $|G_1\rangle$ [282], which does not enjoy the one-root property on its marginals. This in particular means that this property is only encountered on zero-measure subsets of the set of all rank-2 three-qubit density matrices, and therefore does not describe a generic behaviour of such systems. One can then ask: are one-root states useful in the characterisation of multipartite entanglement at all? We will now show that they indeed are.

[282] Verstraete et al., PRA 2002

10.2 MONOGAMY OF ENTANGLEMENT

One of the fundamental traits of multipartite quantum entanglement is its *monogamy* — an inherent constraint on how it can be shared among multiple parties. The concept was first formalised for a system of three qubits in the seminal work

of Coffman, Kundu, and Wootters (CKW) [267], showing that the entanglement of a qubit with another pair of qubits (quantified with the *tangle* τ) places a bound on the total amount of pairwise entanglement between the considered qubit and each qubit from the pair. This relation, referred to as the CKW monogamy inequality, can be expressed for a pure state $|\phi\rangle$ of three qubits as

$$\tau^{(1)}_{1|23}\left(|\phi\rangle\right) \geq \tau^{(2)}_{1|2}\left(|\phi\rangle\right) + \tau^{(2)}_{1|3}\left(|\phi\rangle\right) \tag{10.21}$$

where $\tau^{(1)}_{i|jk}$ represents the *one-tangle*, also known as the *linear entropy*, quantifying the bipartite entanglement between a chosen qubit i and the rest of the system in the state $|\phi\rangle$:

$$\tau^{(1)}_{i|jk}\left(|\phi\rangle\right) = 4\det\left(\mathrm{Tr}_{jk}\,|\phi\rangle\langle\phi|\right) \tag{10.22}$$

and $\tau^{(2)}_{i|j}$ denotes the *two-tangle*, which measures the pairwise entanglement for the reduced subsystem of qubits i and j:

$$\tau^{(2)}_{i|j}\left(|\phi\rangle\right) = C\left(\mathrm{Tr}_{k}\,|\phi\rangle\langle\phi|\right)^2 \tag{10.23}$$

where C is the concurrence.

Remarkably, the residual term in this inequality is invariant under SLOCC as well as under permutations of the parties ijk, and in fact quantifies the genuine three-partite entanglement in the system [260]. It is is now known as the *three-tangle*:

$$\tau^{(3)}_{1|2|3}\left(|\phi\rangle\right) = \tau^{(1)}_{1|23}\left(|\phi\rangle\right) - \tau^{(2)}_{1|2}\left(|\phi\rangle\right) - \tau^{(2)}_{1|3}\left(|\phi\rangle\right) \tag{10.24}$$

and it admits a straightforward closed-form expression as a polynomial of homogeneous degree 4 in the coefficients of the state $|\phi\rangle$ [262,264,267].

Even though it is tailored for particular choices of the entanglement measure, the CKW monogamy inequality is often regarded as a quantitative constraint capturing a characteristic property of entanglement itself, which distinguishes it from other, weaker forms of nonclassical correlations [286,287]. The relation was later extended by Osborne and Verstraete [288] to describe the distribution of pairwise entanglement in a system of n qubits, following the conjecture by CKW [267], and takes the form

$$\tau^{(1)}_{1|\,2\cdots n}\left(|\phi\rangle\right) \geq \sum_{j=2}^{n} \tau^{(2)}_{1|j}\left(|\phi\rangle\right). \tag{10.25}$$

It is, however, obvious that this relation does not fully describe the entanglement of an n-partite state, since on the right-hand side it only contains the entanglement of pairs of particles and disregards genuine multipartite correlations. Indeed, many attempts have been made to construct a generalised monogamy inequality for n qubits [BR12,289–292]. To date, no extension in the spirit of inequality (10.25) was found to provide a constraint on entanglement sharing which could satisfactorily take into account all the forms of multipartite entanglement in the system.

[267] Coffman et al., PRA 2000

[260] Dür et al., PRA 2000

[262] Verstraete et al., PRA 2003

[264] Osterloh & Siewert, PRA 2005

[286] Streltsov et al., PRL 2012

[287] Lancien et al., PRL 2016

[288] Osborne & Verstraete, PRL 2006

[BR12] Regula et al., PRL 2014

[289] Eltschka et al., PRA 2009

[290] Gour & Wallach, JMP 2010

[291] Cornelio, PRA 2013

[292] Eltschka & Siewert, PRL 2015

10.2.1 Strong monogamy inequalities

From eqs. (10.21) and (10.25), one might expect a natural extension of the CKW inequality to hold, where the terms corresponding to different kinds of entanglement of up to $(n-1)$ subsystems are considered. However, there are several questions which make such a generalisation non-trivial; most importantly, how exactly to define and quantify the different types of multipartite entanglement. Let us consider the simplest generalisation of the monogamy inequality for arbitrary pure states $|\phi\rangle$ of four qubits, which could take the form (referred to as "strong monogamy" [BR12,293]):

[BR12] Regula et al., PRL 2014

[293] Adesso & Illuminati, PRL 2007

$$\tau^{(1)}_{1|234} \overset{?}{\geq} \tau^{(2)}_{1|2} + \tau^{(2)}_{1|3} + \tau^{(2)}_{1|4} + \tau^{(3)}_{1|2|3} + \tau^{(3)}_{1|2|4} + \tau^{(3)}_{1|3|4}, \tag{10.26}$$

with $|\phi\rangle$ omitted for simplicity. Here, the square root of the three-tangle $\tau^{(3)}$ can be extended to mixed states by taking the convex roof as

$$\tau^{(3)}_{i|j|k}(\rho) = \left[\min\left\{\sum_n p_n \sqrt{\tau^{(3)}_{i|j|k}(|\psi_n\rangle)} \;\middle|\; \rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|\right\}\right]^2 \tag{10.27}$$

Interestingly, this generalised form of monogamy has been shown to hold in a range of systems of not just four, but n qubits, including mixtures of W states and generalised GHZ states [BR12] as well as generalised W states, which saturate the inequality [267,269,294]. It has also been investigated using different entanglement measures, including the square of convex-roof extended negativity [148] and the standard negativity [295]. Additionally, a relation equivalent to (10.26) has been shown to hold for all permutationally-invariant continuous-variable Gaussian states [293]. All of these results suggest that it might be possible to establish a similar relation for qubits.

[267] Coffman et al., PRA 2000

[269] Eltschka et al., NJP 2008

[294] Kim, PRA 2014

[148] Choi & Kim, PRA 2015

[295] Karmakar et al., PRA 2016

However, and rather surprisingly, this natural extension turns out not to hold in general, even for four qubits.

10.2.2 Violations of strong monogamy

Let us now denote by $\tau^{(4)}_{1|2|3|4}$ the residual term of the inequality, that is,

$$\tau^{(4)}_{1|2|3|4} \coloneqq \tau^{(1)}_{1|234} - \tau^{(2)}_{1|2} - \tau^{(2)}_{1|3} - \tau^{(2)}_{1|4} - \tau^{(3)}_{1|2|3} - \tau^{(3)}_{1|2|4} - \tau^{(3)}_{1|3|4}, \tag{10.28}$$

which makes Eq. (10.26) equivalent to $\tau^{(4)}_{1|2|3|4} \overset{?}{\geq} 0$. To investigate the violations of strong monogamy, we first performed extensive numerical tests in search of states whose residual $\tau^{(4)}_{1|2|3|4}$ is negative. As described in Sec. 10.1.2, any pure four-qubit state can be written as

$$|\phi\rangle = \left(A_1 \otimes A_2 \otimes A_3 \otimes A_4\right) |G_i^{a,b,c,d}\rangle\,, \tag{10.29}$$

where A_k are invertible linear operators, $|G_i^{a,b,c,d}\rangle$ is one of the nine generating states corresponding to one of the nine SLOCC classes of four-qubit entanglement, and a, b, c, d are complex parameters. A natural method to optimise over the set of pure four-qubit states is then to randomise the parameters a, b, c, d as well as the SLOCC transformations A_k applied to the generating states of each class.

Noting that the three-qubit reduced subsystems of each pure four-qubit state are mixed states of rank 2, we employed the so-called zero-E approximation [279] to provide a tight upper bound on the convex roof of the mixed-state three-tangle. This means that our numerical tests in fact bound the residual $\tau^{(4)}_{1|2|3|4}$ from below, and while violations are a priori inconclusive, all positive results confirm the inequality. It is not possible to strictly falsify the strong monogamy inequality unless we have an exact result for the convex roof.

[279] Rodriques et al., PRA 2014

A test of 10^6 randomly generated pure states in each SLOCC class, with all parameters a,b,c,d and invertible linear operators A_k randomised, has returned no violations of the inequality. This provides strong evidence that the natural extension of monogamy given by Eq. (10.26) holds for a *generic* choice of parameters defining the SLOCC classes. However, the situation is more subtle for non-generic states. Specifically, in class 2, generated by

$$\begin{aligned} |G_2^{a,b,c}\rangle = \frac{a+b}{2}\big(\,|0000\rangle + |1111\rangle\,\big) + \frac{a-b}{2}\big(\,|0011\rangle + |1100\rangle\,\big) \\ + c\big(\,|0101\rangle + |1010\rangle\,\big) + |0110\rangle\,, \end{aligned} \tag{10.30}$$

we can find degenerate subclasses with $a = c$ or $b = c$ which do, in fact, lead to violations of the strong monogamy inequality. A further numerical test of random invertible linear operators applied to 5×10^6 states in each of these degenerate subclasses revealed 24,626 violations in the former and 24,654 in the latter, with the most violating state having a residual value of $\tau^{(4)}_{1|2|3|4} \approx -0.097$, which is far beyond any possible error due to numerical precision. As anticipated, violations cannot be conclusively confirmed until one establishes that the adopted upper bound for the mixed-state three-tangles is in fact tight. Fortunately, this is possible in the considered case.

To see this, we note the peculiarity of the two degenerate subclasses of SLOCC class 2. Although the marginals of $|G_2^{a,b,c}\rangle$ are not one-root states in general, it can be straightforwardly verified that taking $a = c$ or $b = c$ does in fact always give four-qubit state whose marginals are all one-root, that is, they only have a single unentangled pure state in their support. By Thm. 10.2, we know that not only can the value of the three-tangle be evaluated exactly for such states, it is in fact the same in every convex decomposition. Therefore, the values of the residual $\tau^{(4)}_{1|2|3|4}$ calculated numerically are in fact the *exact* values, and any violation of the inequality found in these subclasses is a true violation.

In order to present an explicit example, we would like to characterise a family of violations analytically. An example family of states with particularly large vio-

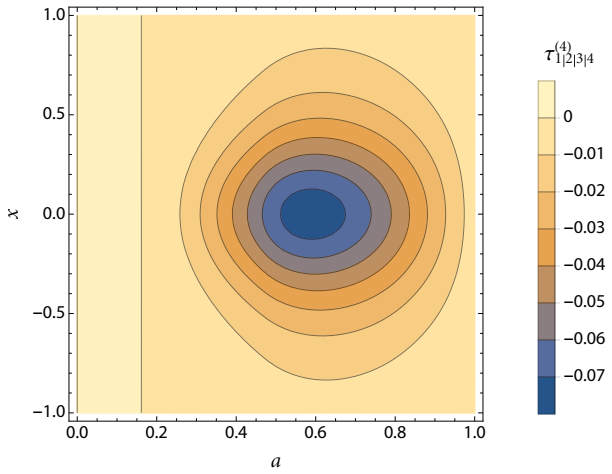


Figure 10.3: The exact residual $\tau^{(4)}_{1|2|3|4}$ as defined in Eq. (10.28) of the state $\rho_{a,x}$ in (10.32) with respect to the parameters a and x.

lations of the inequality can be constructed by beginning with the state $|G_2^{a,b,c}\rangle$ with $b = c$ and additionally imposing $b = c = ia$ for a parameter $a \geq 0$. We can then choose to parametrise an invertible linear operator A_x of determinant 1 in a form which leads to an analytically tractable violating state:

$$A_x = \begin{pmatrix} (1+i)\left(x+\frac{1}{3}\right) & \left(\frac{1}{3}-i\right)+x \\ i\left(x-\frac{2}{3}\right) & \frac{1+i}{6}\left(3x-2-3i\right) \end{pmatrix} \tag{10.31}$$

with a real parameter x. The simplest case to consider is to apply the transformation to the first qubit only, giving the state (up to normalisation)

$$\rho_{a,x} = \left(A_x \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1}\right) |G_2^{a,ia,ia}\rangle\langle G_2^{a,ia,ia}| \left(A_x \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1}\right)^\dagger \tag{10.32}$$

with an explicit dependence on the parameters a and x. We stress that $\rho_{a,x}$ belongs to the degenerate subclass of SLOCC class 2 for all a and x, so we know that we can reach the convex roof of the three-tangles $\tau^{(3)}_{i|j|k}(\rho_{a,x})$ in any decomposition, and therefore the residual $\tau^{(4)}_{1|2|3|4}$ of $\rho_{a,x}$ can be quantified exactly by simply computing the spectral decomposition of the state. While the explicit analytical expression for the residual is lengthy and omitted here, we plot it in Figure 10.3 for different values of the parameters.

As can be seen from the figure, the strong monogamy inequality is violated for a large subset of values of a and x. One can easily notice that the SLOCC transformation with $x = 0$ leads to the biggest violations. To obtain the minimum with regards to a, we analyze the marginal one-, two-, and three-tangles of the subsystems, and find that the value of $a = a_0 = \frac{5}{6\sqrt{2}}$ leads in fact to the

minimum value of the residual $\tau^{(4)}_{1|2|3|4} \approx -0.076$ within the considered family of states.

We have thus explicitly demonstrated a significant violation of the strong monogamy inequality. This shows that the natural extension of the CKW monogamy inequality in Eqs. (10.26) does not hold in general. A natural follow-up question is then whether there exists a different form of a monogamy relation for which violations do not occur. The fact that the natural extension of the CKW inequality does not hold suggests that the three-partite terms in the inequality should not be given the same weight as the bi- and single-partite terms. This can be achieved e.g. by modifying the definition of the three-tangle to

$$\tau^{(3)}_{i|j|k}(\rho) = \left[\min\left\{\sum_n p_n \sqrt{\tau^{(3)}_{i|j|k}(|\psi_n\rangle)} \;\middle|\; \rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|\right\}\right]^{\mu}. \tag{10.33}$$

or to

$$\tau^{(3)}_{i|j|k}(\rho) = \left[\min\left\{\sum_n p_n \left(\tau^{(3)}_{i|j|k}(|\psi_n\rangle)\right)^{1/\mu} \;\middle|\; \rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|\right\}\right]^{\mu} \tag{10.34}$$

for some power $\mu > 2$. The first approach has the advantage that it preserves the degree of the measure inside the convex roof, and we thus recover many useful properties such as the one-root property. On the other hand, the second approach preserves the overall degree of the inequality, which could be considered to be a more desirable behaviour. Although one can verify numerically that both of these approaches are able to remove the violations of the inequality encountered before, we were not able to prove the validity of either of the approaches in the general case, and thus leave a possible extension of the strong monogamy inequality as an open question.

Bibliography

[BR1] B. Regula and G. Adesso, "Entanglement Quantification Made Easy: Polynomial Measures Invariant under Convex Decomposition", Physical Review Letters **116**, 070504 (2016) (cit. on pp. 141, 147).

[BR2] B. Regula, A. Osterloh, and G. Adesso, "Strong monogamy inequalities for four qubits", Physical Review A **93**, 052338 (2016) (cit. on p. 141).

[BR3] B. Regula and G. Adesso, "Geometric approach to entanglement quantification with polynomial measures", Physical Review A **94**, 022324 (2016) (cit. on p. 147).

[BR4] B. Regula, "Convex geometry of quantum resource quantification", Journal of Physics A: Mathematical and Theoretical **51**, 045303 (2018) (cit. on pp. 42, 66).

[BR5] B. Regula, M. Piani, M. Cianciaruso, T. R. Bromley, A. Streltsov, and G. Adesso, "Converting multilevel nonclassicality into genuine multipartite entanglement", New Journal of Physics **20**, 033012 (2018) (cit. on pp. 66, 68, 69, 129).

[BR6] B. Regula, K. Fang, X. Wang, and G. Adesso, "One-Shot Coherence Distillation", Physical Review Letters **121**, 010401 (2018) (cit. on p. 82).

[BR7] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso, "Probabilistic distillation of quantum coherence", arXiv:1804.09500 (2018) (cit. on p. 82).

[BR8] B. Regula, L. Lami, and A. Streltsov, "Non-asymptotic assisted distillation of quantum coherence", arXiv:1807.04705 (2018) (cit. on p. 82).

[BR9] L. Lami, B. Regula, X. Wang, R. Nichols, A. Winter, and G. Adesso, "Gaussian quantum resource theories", arXiv:1801.05450 (2018).

[BR10] T. R. Bromley, M. Cianciaruso, S. Vourekas, B. Regula, and G. Adesso, "Accessible bounds for general quantum resources", Journal of Physics A: Mathematical and Theoretical **51**, 325303 (2018).

[BR11] N. Johnston, C.-K. Li, S. Plosker, Y.-T. Poon, and B. Regula, "Robustness of k-coherence and k-entanglement", arXiv:1806.00653 (2018) (cit. on pp. 53, 70, 74).

[BR12] B. Regula, S. Di Martino, S. Lee, and G. Adesso, "Strong Monogamy Conjecture for Multiqubit Entanglement: The Four-Qubit Case", Physical Review Letters **113**, 110501 (2014) (cit. on pp. 141, 150, 151).

[1] M. Horodecki and J. Oppenheim, "(Quantumness in the context of) Resource theories", International Journal of Modern Physics B **27**, 1345019 (2012) (cit. on p. 7).

[2] J. Sperling and W. Vogel, "Convex ordering and quantification of quantumness", Physica Scripta **90**, 074024 (2015) (cit. on pp. 7, 128, 129).

[3] F. G. S. L. Brandão and G. Gour, "Reversible framework for quantum resource theories", Phys. Rev. Lett. **115**, 070503 (2015) (cit. on pp. 7, 41, 90).

[4] L. del Rio, L. Kraemer, and R. Renner, "Resource theories of knowledge", arXiv:1511.08818 (2015) (cit. on p. 7).

[5] B. Coecke, T. Fritz, and R. W. Spekkens, "A mathematical theory of resources", Information and Computation **250**, 59–86 (2016) (cit. on p. 7).

[6] G. Gour, "Quantum resource theories in the single-shot regime", Physical Review A **95**, 062314 (2017) (cit. on pp. 7, 41, 54, 81).

[7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement", Reviews of Modern Physics **81**, 865–942 (2009) (cit. on pp. 7, 15, 19, 41, 76).

[8] G. Gour and R. W. Spekkens, "The resource theory of quantum reference frames: manipulations and monotones", New Journal of Physics **10**, 033023 (2008) (cit. on pp. 7, 20).

[9] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, "Quantum Coherence, Time-Translation Symmetry, and Thermodynamics", Physical Review X **5**, 021001 (2015) (cit. on p. 7).

[10] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, "Resource Theory of Quantum States Out of Thermal Equilibrium", Physical Review Letters **111**, 250404 (2013) (cit. on p. 7).

[11] M. Horodecki, P. Horodecki, and J. Oppenheim, "Reversible transformations from pure to mixed states and the unique measure of information", Physical Review A **67**, 062104 (2003) (cit. on p. 7).

[12] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, "Operational Framework for Nonlocality", Physical Review Letters **109**, 070401 (2012) (cit. on p. 7).

[13] J. I. de Vicente, "On nonlocality as a resource theory and nonlocality measures", Journal of Physics A: Mathematical and Theoretical **47**, 424017 (2014) (cit. on p. 7).

[14] J. Aberg, "Quantifying Superposition", arXiv:quant-ph/0612146 (2006) (cit. on pp. 7, 20–22, 88).

[15] T. Baumgratz, M. Cramer, and M. B. Plenio, "Quantifying Coherence", Physical Review Letters **113**, 140401 (2014) (cit. on pp. 7, 20–22, 41, 55, 70, 82, 83, 128).

[16] A. Streltsov, G. Adesso, and M. B. Plenio, "Quantum coherence as a resource", Reviews of Modern Physics **89**, 041003 (2017) (cit. on pp. 7, 20–22, 41, 83).

[17] W. Vogel and J. Sperling, "Unified quantification of nonclassicality and entanglement", Physical Review A **89**, 052302 (2014) (cit. on pp. 7, 128, 129).

[18] N. Killoran, F. E. S. Steinhoff, and M. B. Plenio, "Converting Nonclassicality into Entanglement", Physical Review Letters **116**, 080402 (2016) (cit. on pp. 7, 66, 128–131).

[19] T. Theurer, N. Killoran, D. Egloff, and M. Plenio, "Resource Theory of Superposition", Physical Review Letters **119**, 230401 (2017) (cit. on pp. 7, 41, 66, 128, 129).

[20] S. Das, C. Mukhopadhyay, S. S. Roy, S. Bhattacharya, A. S. De, and U. Sen, "Wave-particle duality employing quantum coherence in superposition with non-distinguishable pointers", arXiv:1705.04343 (2017) (cit. on p. 7).

[21] R. Gallego and L. Aolita, "Resource Theory of Steering", Physical Review X **5**, 041008 (2015) (cit. on p. 7).

[22] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, "Quantifying Contextuality", Physical Review Letters **112**, 120401 (2014) (cit. on p. 7).

[23] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation", New Journal of Physics **16**, 013009 (2014) (cit. on pp. 7, 77).

[24] E. Wakakuwa, "Operational Resource Theory of Non-Markovianity", arXiv:1709.07248 (2017) (cit. on p. 7).

[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th (Cambridge University Press, New York, NY, USA, 2011) (cit. on pp. 11, 83, 139).

[26] M. M. Wilde, *Quantum information theory*, 2nd ed. (Cambridge University Press, 2017) (cit. on pp. 11, 83).

[27] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018) (cit. on pp. 11, 35).

[28] A. Uhlmann, "The "transition probability" in the state space of a *-algebra", Reports on Mathematical Physics **9**, 273–279 (1976) (cit. on pp. 13, 63).

[29] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels", Physical Review Letters **70**, 1895–1899 (1993) (cit. on p. 15).

[30] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states", Physical Review Letters **69**, 2881–2884 (1992) (cit. on p. 15).

[31] L. Gurvits, "Classical Deterministic Complexity of Edmonds' Problem and Quantum Entanglement", in Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing, STOC '03 (2003), pp. 10–19 (cit. on p. 16).

[32] A. Peres, "Separability Criterion for Density Matrices", Physical Review Letters **77**, 1413–1415 (1996) (cit. on p. 16).

[33] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions", Physics Letters A **223**, 1–8 (1996) (cit. on pp. 16, 59).

[34] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition", Physics Letters A **232**, 333–339 (1997) (cit. on pp. 16, 75).

[35] C. E. Shannon, "A Mathematical Theory of Communication", Bell System Technical Journal **27**, 379–423 (1948) (cit. on p. 17).

[36] W. van Dam and P. Hayden, "Universal entanglement transformations without communication", Physical Review A **67**, 060302 (2003) (cit. on p. 18).

[37] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels", Physical Review Letters **76**, 722–725 (1996) (cit. on pp. 18, 81).

[38] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, "Concentrating partial entanglement by local operations", Physical Review A **53**, 2046–2052 (1996) (cit. on pp. 18, 81, 102, 127).

[39] E. M. Rains, "Rigorous treatment of distillable entanglement", Physical Review A **60**, 173–178 (1999) (cit. on pp. 18, 81).

[40] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction", Physical Review A **54**, 3824–3851 (1996) (cit. on p. 19).

[41] P. M. Hayden, M. Horodecki, and B. M. Terhal, "The asymptotic entanglement cost of preparing a quantum state", Journal of Physics A: Mathematical and General **34**, 6891 (2001) (cit. on p. 19).

[42] P. W. Shor, "Equivalence of Additivity Questions in Quantum Information Theory", Communications in Mathematical Physics **246**, 473–473 (2004) (cit. on p. 19).

[43] P. Hayden and A. Winter, "Counterexamples to the maximal p-norm multiplicativity conjecture for all p > 1", Communications in Mathematical Physics **284**, 263–280 (2008) (cit. on p. 19).

[44] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs", Nature Physics **5**, 255–257 (2009) (cit. on p. 19).

[45] M. Horodecki, P. Horodecki, and R. Horodecki, "Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?", Physical Review Letters **80**, 5239–5242 (1998) (cit. on p. 19).

[46] P. Horodecki and R. Horodecki, "Distillation and bound entanglement", Quantum Information and Computation **1**, 45–75 (2001) (cit. on p. 19).

[47] L. Clarisse, "Entanglement Distillation; A Discourse on Bound Entanglement in Quantum Information Theory" (University of York, Dec. 10, 2006), arXiv:quant-ph/0612072 (cit. on p. 19).

[48] A. Uhlmann, "Entropy and Optimal Decompositions of States Relative to a Maximal Commutative Subalgebra", Open Syst. Inf. Dyn. **5**, 209–228 (1998) (cit. on pp. 19, 36).

[49] G. Vidal, "Entanglement monotones", Journal of Modern Optics **47**, 355–376 (2000) (cit. on p. 19).

[50] W. K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits", Physical Review Letters **80**, 2245–2248 (1998) (cit. on pp. 19, 142, 147).

[51] C. Eltschka and J. Siewert, "Quantifying entanglement resources", Journal of Physics A: Mathematical and Theoretical **47**, 424005 (2014) (cit. on pp. 19, 141).

[52] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying Entanglement", Physical Review Letters **78**, 2275–2279 (1997) (cit. on pp. 20, 55, 61, 65, 121).

[53] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Dec. 6, 2007), 479 pp. (cit. on pp. 20, 61, 65, 146).

[54] F. Levi and F. Mintert, "A quantitative theory of coherent delocalization", New Journal of Physics **16**, 033007 (2014) (cit. on p. 20).

[55] A. Winter and D. Yang, "Operational resource theory of coherence", Physical Review Letters **116**, 120404 (2016) (cit. on pp. 20, 21, 55, 67, 81, 87, 90, 93, 94, 124).

[56] E. Chitambar and G. Gour, "Critical examination of incoherent operations and a physically consistent resource theory of quantum coherence", Physical Review Letters **117**, 030401 (2016) (cit. on pp. 20, 21, 55, 94, 98, 101, 124).

[57] I. Marvian and R. W. Spekkens, "How to quantify coherence: Distinguishing speakable and unspeakable notions", Physical Review A **94**, 052324 (2016) (cit. on pp. 20, 21, 55, 89).

[58] J. I. de Vicente and A. Streltsov, "Genuine quantum coherence", Journal of Physics A: Mathematical and Theoretical **50**, 045301 (2017) (cit. on pp. 20, 55).

[59] B. Yadin, J. Ma, D. Girolami, M. Gu, and V. Vedral, "Quantum Processes Which Do Not Use Coherence", Physical Review X **6**, 041028 (2016) (cit. on p. 21).

[60] E. Chitambar and G. Gour, "Comparison of incoherent operations and measures of coherence", Physical Review A **94**, 052336 (2016) (cit. on pp. 21, 82, 89, 90, 93).

[61] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, "Measuring Quantum Coherence with Entanglement", Physical Review Letters **115**, 020403 (2015) (cit. on pp. 22, 67, 128, 138).

[62] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, "Robustness of coherence: an operational and observable measure of quantum coherence", Physical Review Letters **116**, 150502 (2016) (cit. on pp. 22, 67, 81, 83).

[63] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970) (cit. on pp. 24, 27, 28, 32, 37, 48).

[64] J.-B. Hiriart-Urrutyand C. Lemaréchal, *Convex analysis and minimization algorithms I: fundamentals* (Springer-Verlag, Berlin, 1993) (cit. on pp. 24, 27, 35).

[65] J.-B. Hiriart-Urrutyand C. Lemaréchal, *Convex analysis and minimization algorithms II: advanced theory and bundle methods* (Springer-Verlag, Berlin, 1993) (cit. on p. 24).

[66] F. Bach, "Convex relaxations of structured matrix factorizations", arXiv:1309.3117 (2013) (cit. on pp. 27, 45).

[67] M. Friedlander, I. Macedo, and T. Pong, "Gauge optimization and duality", SIAM Journal on Optimization **24**, 1999–2022 (2014) (cit. on p. 27).

[68] R. M. Freund, "Dual gauge programs, with applications to quadratic programming and the minimum-norm problem", Mathematical Programming **38**, 47–67 (1987) (cit. on p. 28).

[69] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The Convex Geometry of Linear Inverse Problems", Foundations of Computational Mathematics **12**, 805–849 (2012) (cit. on p. 30).

[70] N. Bourbaki, *Topological Vector Spaces* (Springer-Verlag, Berlin, 2003) (cit. on p. 32).

[71] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004) (cit. on pp. 32, 35, 48).

[72] L. Vandenberghe and S. Boyd, "Semidefinite programming", SIAM Review **38**, 49–95 (1996) (cit. on p. 35).

[73] M. Sion, "On general minimax theorems", Pacific Journal of Mathematics **8**, 171–176 (1958) (cit. on p. 35).

[74] A. Uhlmann, "Roofs and Convexity", Entropy **12**, 1799–1832 (2010) (cit. on pp. 36, 111).

[75] A. Lima, "On Continuous Convex Functions and Split Faces", Proceedings of the London Mathematical Society **s3-25**, 27–40 (1972) (cit. on p. 36).

[76] H. J. M. Peters and P. P. Wakker, "Convex functions on non-convex domains", Economics Letters **22**, 251–255 (1986) (cit. on p. 36).

[77] M. Tawarmalani and N. V. Sahinidis, "Convex extensions and envelopes of lower semi-continuous functions", Mathematical Programming **93**, 247–263 (2002) (cit. on p. 36).

[78] O. Bucicovschi and J. Lebl, "On the continuity and regularity of convex extensions", Journal of Convex Analysis **20**, 1113 (2013) (cit. on p. 36).

[79] M. Yan, "Extension of Convex Function", Journal of Convex Analysis **21**, 965 (2014) (cit. on p. 36).

[80] I. Shcherbatyi and B. Andres, "Convexification of Learning from Constraints", in Pattern Recognition: 38th German Conference, GCPR 2016, Hannover, Germany, September 12-15, 2016, Proceedings, edited by B. Rosenhahn and B. Andres (2016), pp. 79–90 (cit. on p. 36).

[81] M. B. Plenio and S. Virmani, "An Introduction to Entanglement Measures", Quantum Inf. Comput. **7**, 1–51 (2007) (cit. on pp. 41, 49, 75).

[82] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, "Robustness of asymmetry and coherence of quantum states", Physical Review A **93**, 042107 (2016) (cit. on pp. 41, 52, 67, 83).

[83] M. Howard and E. Campbell, "Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing", Physical Review Letters **118**, 090501 (2017) (cit. on pp. 41, 77–79).

[84] A. Hartkämper and H. Neumann, eds., *Foundations of Quantum Mechanics and Ordered Linear Spaces* (Springer, 1974), 355 pp. (cit. on pp. 43, 48, 68).

[85] E. Richard, G. Obozinski, and J.-P. Vert, "Tight Convex Relaxations for Sparse Matrix Factorization", in Advances in Neural Information Processing Systems (2014), pp. 3284–3292 (cit. on pp. 44, 70).

[86] E. Richard, G. Obozinski, and J.-P. Vert, "Tight convex relaxations for sparse matrix factorization (Technical report)", arXiv:1407.5158 (2014) (cit. on pp. 44, 70).

[87] G. Jameson, *Summing and Nuclear Norms in Banach Space Theory* (Cambridge University Press, 1987) (cit. on pp. 45, 73).

[88] W. Matthews, S. Wehner, and A. Winter, "Distinguishability of Quantum States Under Restricted Families of Measurements with an Application to Quantum Data Hiding", Communications in Mathematical Physics **291**, 813–843 (2009) (cit. on p. 48).

[89] D. Reeb, M. J. Kastoryano, and M. M. Wolf, "Hilbert's projective metric in quantum information theory", Journal of Mathematical Physics **52**, 082201 (2011) (cit. on pp. 48, 50).

[90] A. Jenčová, "Base norms and discrimination of generalized quantum channels", Journal of Mathematical Physics **55**, 022201 (2014) (cit. on p. 48).

[91] L. Lami, C. Palazuelos, and A. Winter, "Ultimate Data Hiding in Quantum Mechanics and Beyond", Communications in Mathematical Physics, 1–48 (2018) (cit. on p. 48).

[92] G. Vidal and R. F. Werner, "Computable measure of entanglement", Physical Review A **65**, 032314 (2002) (cit. on pp. 49, 72, 75, 77, 119).

[93] O. Rudolph, "Further Results on the Cross Norm Criterion for Separability", Quantum Information Processing **4**, 219–239 (2005) (cit. on pp. 49, 73, 76).

[94] G. Vidal and R. Tarrach, "Robustness of entanglement", Physical Review A **59**, 141–155 (1999) (cit. on pp. 49, 50, 53, 60, 72, 74, 126).

[95] A. C. Thompson, "On certain contraction mappings in a partially ordered vector space.", Proceedings of the American Mathematical Society **14**, 438–443 (1963) (cit. on p. 50).

[96] Ş. Cobzaş and M.-D. Rus, "Normal Cones and Thompson Metric", in *Topics in Mathematical Analysis and Applications*, Springer Optimization and Its Applications (Springer, Cham, 2014), pp. 209–258 (cit. on p. 50).

[97] D. Hilbert, "Über die gerade Linie als kürzeste Verbindung zweier Punkte", Mathematische Annalen **46**, 91–96 (1895) (cit. on p. 50).

[98] P. J. Bushell, "Hilbert's metric and positive contraction mappings in a Banach space", Archive for Rational Mechanics and Analysis **52**, 330–338 (1973) (cit. on p. 50).

[99] N. Datta, "Min- and Max-Relative Entropies and a New Entanglement Monotone", IEEE Transactions on Information Theory **55**, 2816–2826 (2009) (cit. on pp. 50, 65, 92).

[100] F. G. S. L. Brandão, "Quantifying entanglement with witness operators", Physical Review A **72**, 022310 (2005) (cit. on pp. 51, 59, 83, 113).

[101] M. Steiner, "Generalized robustness of entanglement", Physical Review A **67**, 054305 (2003) (cit. on pp. 52, 72, 77).

[102] L. Clarisse, "On the schmidt robustness of pure states", Journal of Physics A: Mathematical and General **39**, 4239 (2006) (cit. on pp. 52, 73).

[103] D. Cavalcanti, F. G. S. L. Brandão, and M. O. Terra Cunha, "Are all maximally entangled states pure?", Physical Review A **72**, 040303 (2005) (cit. on pp. 52, 76).

[104] M. Ringbauer, T. R. Bromley, M. Cianciaruso, S. Lau, G. Adesso, A. G. White, A. Fedrizzi, and M. Piani, "Quasi-device-independent witnessing of genuine multilevel quantum coherence", arXiv:1707 . 05282 (2017) (cit. on pp. 52, 66, 68, 70, 140).

[105] N. Johnston and D. W. Kribs, "Duality of Entanglement Norms", Houston Journal of Mathematics, 831–847 (2015) (cit. on pp. 52, 69, 73, 74).

[106] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, "Quantum nonlocality without entanglement", Physical Review A **59**, 1070–1091 (1999) (cit. on pp. 55, 112, 121).

[107] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, "Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)", Communications in Mathematical Physics **328**, 303–326 (2014) (cit. on pp. 55, 72).

[108] K. Bu and C. Xiong, "A note on cohering power and de-cohering power", Quantum Information and Computation **13**, 1206 (2017) (cit. on pp. 58, 101).

[109] B. M. Terhal, "Detecting quantum entanglement", Theoretical Computer Science, Natural Computing **287**, 313–335 (2002) (cit. on p. 59).

[110] O. Gühne and G. Tóth, "Entanglement detection", Physics Reports **474**, 1–75 (2009) (cit. on p. 59).

[111] O. Gühne, M. Reimpell, and R. F. Werner, "Estimating Entanglement Measures in Experiments", Physical Review Letters **98**, 110502 (2007) (cit. on p. 59).

[112] J. Eisert, F. G. S. L. Brandão, and K. M. R. Audenaert, "Quantitative entanglement witnesses", New Journal of Physics **9**, 46 (2007) (cit. on p. 59).

[113] F. G. S. L. Brandão and R. O. Vianna, "Witnessed entanglement", International Journal of Quantum Information **04**, 331–340 (2006) (cit. on p. 60).

[114] M. Lewenstein and A. Sanpera, "Separability and Entanglement of Composite Quantum Systems", Physical Review Letters **80**, 2261–2264 (1998) (cit. on pp. 60, 143, 149).

[115] J. Eisert, K. Audenaert, and M. B. Plenio, "Remarks on entanglement measures and non-local state distinguishability", Journal of Physics A: Mathematical and General **36**, 5605 (2003) (cit. on p. 62).

[116] X.-D. Yu, D.-J. Zhang, G. F. Xu, and D. M. Tong, "Alternative framework for quantifying coherence", Physical Review A **94**, 060302 (2016) (cit. on pp. 62, 81, 83).

[117] L.-F. Qiao, J. Gao, A. Streltsov, S. Rana, R.-J. Ren, Z.-Q. Jiao, C.-Q. Hu, X.-Y. Xu, C.-Y. Wang, H. Tang, A.-L. Yang, Z.-H. Ma, M. Lewenstein, and X.-M. Jin, "Activation of entanglement from quantum coherence and superposition", arXiv:1710.04447 (2017) (cit. on p. 62).

[118] A. Shimony, "Degree of Entanglementa", Annals of the New York Academy of Sciences **755**, 675–679 (1995) (cit. on pp. 63, 72, 186).

[119] T.-C. Wei and P. M. Goldbart, "Geometric measure of entanglement and applications to bipartite and multipartite quantum states", Physical Review A **68**, 042307 (2003) (cit. on pp. 63, 72).

[120] A. Streltsov, H. Kampermann, and D. Bruß, "Linking a distance measure of entanglement to its convex roof", New Journal of Physics **12**, 123004 (2010) (cit. on pp. 63, 138).

[121] R. Jozsa, "Fidelity for Mixed Quantum States", Journal of Modern Optics **41**, 2315–2323 (1994) (cit. on p. 63).

[122] D. Cavalcanti, "Connecting the generalized robustness and the geometric measure of entanglement", Physical Review A **73**, 044302 (2006) (cit. on p. 64).

[123] H. Zhu, M. Hayashi, and L. Chen, "Coherence and entanglement measures based on Rényi relative entropies", Journal of Physics A: Mathematical and Theoretical **50**, 475303 (2017) (cit. on pp. 65, 71, 124).

[124] S. Rana, P. Parashar, A. Winter, and M. Lewenstein, "Logarithmic coherence: operational interpretation of ℓ_1-norm coherence", Physical Review A **96**, 052336 (2017) (cit. on p. 65).

[125] F. Dupuis, L. Krämer, P. Faist, J. M. Renes, and R. Renner, "Generalized entropies", in *XVIIth International Congress on Mathematical Physics*, 0 vols. (WORLD SCIENTIFIC, Dec. 20, 2012), pp. 134–153 (cit. on p. 65).

[126] S. Rana, P. Parashar, and M. Lewenstein, "Trace-distance measure of coherence", Physical Review A **93**, 012110 (2016) (cit. on pp. 65, 70, 82).

[127] S. Chin, "Coherence number as a discrete quantum resource", Physical Review A **96**, 042336 (2017) (cit. on pp. 66, 70).

[128] S. Chin, "Generalized coherence concurrence and path distinguishability", Journal of Physics A: Mathematical and Theoretical **50**, 475302 (2017) (cit. on p. 66).

[129] X. Yuan, H. Zhou, Z. Cao, and X. Ma, "Intrinsic randomness as a measure of quantum coherence", Physical Review A **92**, 022124 (2015) (cit. on p. 67).

[130] X. Qi, T. Gao, and F. Yan, "Measuring coherence with entanglement concurrence", Journal of Physics A: Mathematical and Theoretical **50**, 285301 (2017) (cit. on p. 67).

[131] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Oct. 22, 2012), 663 pp. (cit. on p. 68).

[132] A. Argyriou, R. Foygel, and N. Srebro, "Sparse prediction with the k-support norm", in Advances in Neural Information Processing Systems (2012), pp. 1457–1465 (cit. on p. 68).

[133] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Structured Sparsity through Convex Optimization", Statistical Science **27**, 450–468 (2012) (cit. on p. 69).

[134] G. S. Mudholkar and M. Freimer, "A structure theorem for the polars of unitarily invariant norms", Proceedings of the American Mathematical Society **95**, 331–337 (1985) (cit. on p. 69).

[135] B. M. Terhal and P. Horodecki, "Schmidt number for density matrices", Physical Review A **61**, 040301 (2000) (cit. on p. 71).

[136] D. Chruściński and A. Kossakowski, "Spectral Conditions for Positive Maps", Communications in Mathematical Physics **290**, 1051–1064 (2009) (cit. on p. 71).

[137] N. Johnston and D. W. Kribs, "A family of norms with applications in quantum information theory", Journal of Mathematical Physics **51**, 082202 (2010) (cit. on pp. 71–74, 113).

[138] G. Sentís, C. Eltschka, O. Gühne, M. Huber, and J. Siewert, "Quantifying Entanglement of Maximal Dimension in Bipartite Mixed States", Physical Review Letters **117**, 190502 (2016) (cit. on p. 71).

[139] J. Chen, S. Grogan, N. Johnston, C.-K. Li, and S. Plosker, "Quantifying the coherence of pure quantum states", Physical Review A **94**, 042313 (2016) (cit. on p. 71).

[140] H. Zhu, Z. Ma, Z. Cao, S.-M. Fei, and V. Vedral, "Operational one-to-one mapping between coherence and entanglement measures", Physical Review A **96**, 032316 (2017) (cit. on pp. 71, 98).

[141] Ł. Skowronek, E. Størmer, and K. Życzkowski, "Cones of positive maps and their duality relations", Journal of Mathematical Physics **50**, 062106 (2009) (cit. on p. 71).

[142] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, "Volume of the set of separable states", Physical Review A **58**, 883–892 (1998) (cit. on pp. 71, 72).

[143] O. Rudolph, "A separability criterion for density operators", Journal of Physics A: Mathematical and General **33**, 3951 (2000) (cit. on pp. 72, 73).

[144] P. Gawron, Z. Puchała, J. A. Miszczak, Ł. Skowronek, and K. Życzkowski, "Restricted numerical range: A versatile tool in the theory of quantum information", Journal of Mathematical Physics **51**, 102204 (2010) (cit. on p. 72).

[145] S. Lee, D. P. Chi, S. D. Oh, and J. Kim, "Convex-roof extended negativity as an entanglement measure for bipartite quantum systems", Physical Review A **68**, 062304 (2003) (cit. on p. 72).

[146] Y.-C. Ou and H. Fan, "Monogamy inequality in terms of negativity for three-qubit states", Physical Review A **75**, 062308 (2007) (cit. on p. 73).

[147] J. S. Kim, A. Das, and B. C. Sanders, "Entanglement monogamy of multipartite higher-dimensional quantum systems using convex-roof extended negativity", Physical Review A **79**, 012329 (2009) (cit. on p. 73).

[148] J. H. Choi and J. S. Kim, "Negativity and strong monogamy of multiparty quantum entanglement beyond qubits", Physical Review A **92**, 042307 (2015) (cit. on pp. 73, 151).

[149] T. Tian, Y. Luo, and Y. Li, "Generalised monogamy relation of convex-roof extended negativity in multi-level systems", Scientific Reports **6**, 36700 (2016) (cit. on p. 73).

[150] O. Rudolph, "A new class of entanglement measures", Journal of Mathematical Physics **42**, 5306–5314 (2001) (cit. on pp. 73, 76).

[151] G. Pisier, *Factorization of Linear Operators and Geometry of Banach Spaces* (American Mathematical Soc., 1986), 166 pp. (cit. on p. 73).

[152] G. Gour, "Family of concurrence monotones and its applications", Physical Review A **71**, 012318 (2005) (cit. on p. 74).

[153] N. Johnston and D. W. Kribs, "A Family of Norms With Applications In Quantum Information Theory II", Quantum Information and Computation **11**, 104–123 (2011) (cit. on pp. 73, 74).

[154] G. Vidal, "Entanglement of Pure States for a Single Copy", Physical Review Letters **83**, 1046–1049 (1999) (cit. on pp. 75, 98, 127).

[155] M. W. Girard and G. Gour, "Entanglement monotones and transformations of symmetric bipartite states", Physical Review A **95**, 012308 (2017) (cit. on p. 75).

[156] E. M. Rains, "A semidefinite program for distillable entanglement", IEEE Transactions on Information Theory **47**, 2921–2933 (2001) (cit. on pp. 75, 117, 118, 126).

[157] X. Wang and R. Duan, "Improved semidefinite programming upper bound on distillable entanglement", Physical Review A **94**, 050301 (2016) (cit. on pp. 75, 119).

[158] B. Jungnitsch, T. Moroder, and O. Gühne, "Taming Multiparticle Entanglement", Physical Review Letters **106**, 190502 (2011) (cit. on p. 75).

[159] M. Hofmann, T. Moroder, and O. Gühne, "Analytical characterization of the genuine multiparticle negativity", Journal of Physics A: Mathematical and Theoretical **47**, 155301 (2014) (cit. on pp. 75, 77).

[160] O. Gühne, G. Tóth, and H. J. Briegel, "Multipartite entanglement in spin chains", New Journal of Physics **7**, 229 (2005) (cit. on p. 76).

[161] M. Walter, D. Gross, and J. Eisert, "Multi-partite entanglement", arXiv:1612.02437 (2016) (cit. on p. 76).

[162] M. Blasone, F. Dell'Anno, S. De Siena, and F. Illuminati, "Hierarchies of geometric entanglement", Physical Review A **77**, 062304 (2008) (cit. on p. 76).

[163] Z. Puchała, P. Gawron, J. A. Miszczak, Ł. Skowronek, M.-D. Choi, and K. Życzkowski, "Product numerical range in a space with tensor product structure", Linear Algebra and its Applications **434**, 327–342 (2011) (cit. on p. 76).

[164] W. Arveson, "Maximal vectors in Hilbert space and quantum entanglement", Journal of Functional Analysis **256**, 1476–1510 (2009) (cit. on p. 76).

[165] H. A. Carteret, A. Higuchi, and A. Sudbery, "Multipartite generalization of the Schmidt decomposition", Journal of Mathematical Physics **41**, 7932–7939 (2000) (cit. on p. 77).

[166] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach, "Generalized Schmidt Decomposition and Classification of Three-Quantum-Bit States", Physical Review Letters **85**, 1560–1563 (2000) (cit. on p. 77).

[167] F. Sokoli and G. Alber, "Generalized Schmidt decomposability and its relation to projective norms in multipartite entanglement", Journal of Physics A: Mathematical and Theoretical **47**, 325301 (2014) (cit. on p. 77).

[168] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, "Negative quasi-probability as a resource for quantum computation", New Journal of Physics **14**, 113011 (2012) (cit. on p. 77).

[169] M. Howard, J. Wallman, V. Veitch, and J. Emerson, "Contextuality supplies the 'magic' for quantum computation", Nature **510**, 351–355 (2014) (cit. on p. 77).

[170] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, "Wigner Function Negativity and Contextuality in Quantum Computation on Rebits", Physical Review X **5**, 021003 (2015) (cit. on p. 77).

[171] S. Bravyi, G. Smith, and J. A. Smolin, "Trading Classical and Quantum Computational Resources", Physical Review X **6**, 021043 (2016) (cit. on pp. 77, 78).

[172] R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, "Contextuality and Wigner-function negativity in qubit quantum computation", Physical Review A **95**, 052334 (2017) (cit. on p. 77).

[173] M. Ahmadi, H. B. Dang, G. Gour, and B. C. Sanders, "Quantification and manipulation of magic states", Physical Review A **97**, 062332 (2018) (cit. on pp. 77, 78).

[174] D. Gross, "Hudson's theorem for finite-dimensional quantum systems", Journal of Mathematical Physics **47**, 122107 (2006) (cit. on p. 78).

[175] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal Clifford gates and noisy ancillas", Physical Review A **71**, 022316 (2005) (cit. on p. 79).

[176] E. Chitambar, A. Streltsov, S. Rana, M. Bera, G. Adesso, and M. Lewenstein, "Assisted Distillation of Quantum Coherence", Physical Review Letters **116**, 070402 (2016) (cit. on pp. 81, 104, 105, 108, 110).

[177] L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing", Physical Review Letters **108**, 200501 (2012) (cit. on pp. 81, 92).

[178] J. M. Renes and R. Renner, "Noisy Channel Coding via Privacy Amplification and Information Reconciliation", IEEE Transactions on Information Theory **57**, 7377–7385 (2011) (cit. on p. 81).

[179] M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks", IEEE Transactions on Information Theory **59**, 7693–7710 (2013) (cit. on pp. 81, 92).

[180] M. Berta, M. Christandl, and R. Renner, "The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory", Communications in Mathematical Physics **306**, 579 (2011) (cit. on p. 81).

[181] D. Leung and W. Matthews, "On the Power of PPT-Preserving and Non-Signalling Codes", IEEE Transactions on Information Theory **61**, 4486–4499 (2015) (cit. on p. 81).

[182] N. Datta and M. H. Hsieh, "One-Shot Entanglement-Assisted Quantum and Classical Communication", IEEE Transactions on Information Theory **59**, 1929–1939 (2013) (cit. on p. 81).

[183] X. Wang, W. Xie, and R. Duan, "Semidefinite Programming Strong Converse Bounds for Classical Capacity", IEEE Transactions on Information Theory **64**, 640–653 (2018) (cit. on p. 81).

[184] K. Fang, X. Wang, M. Tomamichel, and R. Duan, "Non-asymptotic entanglement distillation", arXiv:1706.06221 (2017) (cit. on pp. 81, 118).

[185] F. G. S. L. Brandão and N. Datta, "One-shot rates for entanglement manipulation under non-entangling maps", IEEE Transactions on Information Theory **57**, 1754–1760 (2011) (cit. on pp. 81, 104).

[186] F. Buscemi and N. Datta, "Distilling entanglement from arbitrary resources", Journal of Mathematical Physics **51**, 102201 (2010) (cit. on pp. 81, 92).

[187] F. Buscemi and N. Datta, "General Theory of Environment-Assisted Entanglement Distillation", IEEE Transactions on Information Theory **59**, 1940–1954 (2013) (cit. on pp. 81, 104, 107, 127).

[188] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and X. Ma, "One-Shot Coherence Dilution", Physical Review Letters **120**, 070403 (2018) (cit. on pp. 81, 90).

[189] K. Bu, U. Singh, S.-M. Fei, A. K. Pati, and J. Wu, "Maximum Relative Entropy of Coherence: An Operational Coherence Measure", Physical Review Letters **119**, 150405 (2017) (cit. on pp. 81, 93).

[190] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, "Quantum majorization and a complete set of entropic conditions for quantum thermodynamics", arXiv:1708.04302 (2017) (cit. on p. 81).

[191] E. Chitambar, "Dephasing-covariant operations enable asymptotic reversibility of quantum resources", Physical Review A **97**, 050301 (2018) (cit. on pp. 81, 90, 99).

[192] N. Johnston, C.-K. Li, and S. Plosker, "The modified trace distance of coherence is constant on most pure states", arXiv:1708.06967 (2017) (cit. on pp. 87, 95).

[193] G. Vidal and J. I. Cirac, "Irreversibility in Asymptotic Manipulations of Entanglement", Physical Review Letters **86**, 5803–5806 (2001) (cit. on p. 90).

[194] G. Vidal, W. Dür, and J. I. Cirac, "Entanglement Cost of Bipartite Mixed States", Physical Review Letters **89**, 027901 (2002) (cit. on p. 90).

[195] F. G. S. L. Brandão and M. B. Plenio, "Entanglement theory and the second law of thermodynamics", Nature Physics **4**, 873–877 (2008) (cit. on p. 90).

[196] X. Wang and R. Duan, "Irreversibility of Asymptotic Entanglement Manipulation Under Quantum Operations Completely Preserving Positivity of Partial Transpose", Physical Review Letters **119**, 180506 (2017) (cit. on pp. 90, 120).

[197] M. Hayashi, *Quantum Information Theory: Mathematical Foundation* (Springer, Nov. 3, 2016), 672 pp. (cit. on p. 92).

[198] M. A. Nielsen, "Conditions for a Class of Entanglement Transformations", Physical Review Letters **83**, 436–439 (1999) (cit. on pp. 93, 94, 122).

[199] H. Zhu, M. Hayashi, and L. Chen, "Axiomatic and operational connections between the l_1-norm of coherence and negativity", Physical Review A **97**, 022342 (2018) (cit. on pp. 93, 94).

[200] S. Ishizaka and M. B. Plenio, "Multiparticle entanglement manipulation under positive partial transpose preserving operations", Physical Review A **71**, 052303 (2005) (cit. on pp. 97, 127).

[201] F. Buscemi and G. Gour, "Quantum relative Lorenz curves", Physical Review A **95**, 012110 (2017) (cit. on p. 97).

[202] H.-K. Lo and S. Popescu, "Concentrating entanglement by local actions: Beyond mean values", Physical Review A **63**, 022301 (2001) (cit. on pp. 98, 102, 127).

[203] S. Du, Z. Bai, and X. Qi, "Coherence Measures and Optimal Conversion for Coherent States", Quantum Inf. Comput. **15**, 1307–1316 (2015) (cit. on p. 98).

[204] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge (Springer-Verlag, Berlin Heidelberg, 1978) (cit. on p. 102).

[205] T. Ogawa and H. Nagaoka, "Strong converse to the quantum channel coding theorem", IEEE Transactions on Information Theory **45**, 2486–2489 (1999) (cit. on p. 102).

[206] A. Winter, "Coding theorem and strong converse for quantum channels", IEEE Transactions on Information Theory **45**, 2481–2485 (1999) (cit. on p. 102).

[207] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, J. A. Smolin, A. Thapliyal, and A. Uhlmann, "Entanglement of Assistance", in *Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science (Springer, Berlin, Heidelberg, 1999), pp. 247–257 (cit. on p. 104).

[208] J. A. Smolin, F. Verstraete, and A. Winter, "Entanglement of assistance and multipartite state distillation", Physical Review A **72**, 052317 (2005) (cit. on pp. 104, 105).

[209] M. Gregoratti and R. F. Werner, "Quantum lost and found", Journal of Modern Optics **50**, 915–933 (2003) (cit. on p. 104).

[210] P. Hayden and C. King, "Correcting quantum channels by measuring the environment", Quantum Information and Computation **5**, 156 (2005) (cit. on p. 104).

[211] A. Winter, "On environment-assisted capacities of quantum channels", Markov Processes and Related Fields, 297 (2007) (cit. on p. 104).

[212] E. Chitambar and M.-H. Hsieh, "Relating the Resource Theories of Entanglement and Quantum Coherence", Physical Review Letters **117**, 020402 (2016) (cit. on p. 104).

[213] A. Streltsov, S. Rana, M. N. Bera, and M. Lewenstein, "Towards Resource Theory of Coherence in Distributed Scenarios", Physical Review X **7**, 011024 (2017) (cit. on pp. 104, 107).

[214] K.-D. Wu, Z. Hou, H.-S. Zhong, Y. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, "Experimentally obtaining maximal coherence via assisted distillation process", Optica **4**, 454–459 (2017) (cit. on p. 104).

[215] K.-D. Wu, Z. Hou, Y.-Y. Zhao, G.-Y. Xiang, C.-F. Li, G.-C. Guo, J. Ma, Q.-Y. He, J. Thompson, and M. Gu, "Experimental Cyclic Interconversion between Coherence and Quantum Correlations", Physical Review Letters **121**, 050401 (2018) (cit. on p. 104).

[216] E. Schrödinger, "Probability relations between separated systems", Mathematical Proceedings of the Cambridge Philosophical Society **32**, 446–452 (1936) (cit. on pp. 104, 130).

[217] L. P. Hughston, R. Jozsa, and W. K. Wootters, "A complete classification of quantum ensembles having a given density matrix", Physics Letters A **183**, 14–18 (1993) (cit. on pp. 104, 130).

[218] J. P. R. Christensen and J. Vesterstrøm, "A note on extreme positive definite matrices", Mathematische Annalen **244**, 65–68 (1979) (cit. on p. 108).

[219] R. Grone, S. Pierce, and W. Watkins, "Extremal correlation matrices", Linear Algebra and its Applications **134**, 63–70 (1990) (cit. on p. 108).

[220] C. Li and B. Tam, "A Note on Extreme Correlation Matrices", SIAM Journal on Matrix Analysis and Applications **15**, 903–908 (1994) (cit. on p. 108).

[221] J. Watrous, "Semidefinite programs for completely bounded norms", Theory of Computing **5**, 217–238 (2009) (cit. on pp. 108, 140).

[222] J. Watrous, "Simpler semidefinite programs for completely bounded norms", Chicago Journal of Theoretical Computer Science **19**, 1–19 (2013) (cit. on pp. 108, 140).

[223] M.-J. Zhao, T. Ma, and S.-M. Fei, "Coherence of assistance and regularized coherence of assistance", Physical Review A **96**, 062332 (2017) (cit. on p. 110).

[224] F. G. S. L. Brandão and M. B. Plenio, "A Reversible Theory of Entanglement and its Relation to the Second Law", Communications in Mathematical Physics **295**, 829–851 (2010) (cit. on pp. 112, 121).

[225] E. M. Rains, "Bound on distillable entanglement", Physical Review A **60**, 179–184 (1999) (cit. on pp. 112, 117, 118).

[226] T. Eggeling, K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf, "Distillability via Protocols Respecting the Positivity of Partial Transpose", Physical Review Letters **87**, 257902 (2001) (cit. on p. 117).

[227] K. Audenaert, M. B. Plenio, and J. Eisert, "Entanglement Cost under Positive-Partial-Transpose-Preserving Operations", Physical Review Letters **90**, 027901 (2003) (cit. on pp. 117, 185).

[228] M. Horodecki, "Entanglement Measures", Quantum Information and Computation **1**, 3–26 (2001) (cit. on p. 117).

[229] W. Matthews and A. Winter, "Pure-state transformations and catalysis under operations that completely preserve positivity of partial transpose", Physical Review A **78**, 012317 (2008) (cit. on p. 117).

[230] M. Horodecki and P. Horodecki, "Reduction criterion of separability and limits for a class of distillation protocols", Physical Review A **59**, 4206–4216 (1999) (cit. on pp. 119, 122, 183).

[231] E. Chitambar, J. I. de Vicente, M. W. Girard, and G. Gour, "Entanglement manipulation and distillability beyond LOCC", arXiv:1711.03835 (2017) (cit. on p. 119).

[232] X. Wang, K. Fang, and R. Duan, "Semidefinite programming converse bounds for quantum communication", arXiv:1709.00200 (2017) (cit. on p. 119).

[233] M. Berta and M. M. Wilde, "Amortization does not enhance the max-Rains information of a quantum channel", New Journal of Physics **20**, 053044 (2018) (cit. on p. 119).

[234] F. G. S. L. Brandão and M. B. Plenio, "A Generalization of Quantum Stein's Lemma", Communications in Mathematical Physics **295**, 791–828 (2010) (cit. on p. 120).

[235] G. Vidal, D. Jonathan, and M. A. Nielsen, "Approximate transformations and robust manipulation of bipartite pure-state entanglement", Physical Review A **62**, 012304 (2000) (cit. on p. 122).

[236] D. Jonathan and M. B. Plenio, "Entanglement-Assisted Local Manipulation of Pure Quantum States", Physical Review Letters **83**, 3566–3569 (1999) (cit. on p. 127).

[237] F. Rozpędek, T. Schiet, L. P. Thinh, D. Elkouss, A. C. Doherty, and S. Wehner, "Optimizing practical entanglement distillation", Physical Review A **97**, 062333 (2018) (cit. on p. 127).

[238] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field", Physical Review **131**, 2766–2788 (1963) (cit. on pp. 128, 129).

[239] E. C. G. Sudarshan, "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams", Physical Review Letters **10**, 277–279 (1963) (cit. on pp. 128, 129).

[240] M. S. Kim, W. Son, V. Bužek, and P. L. Knight, "Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement", Physical Review A **65**, 032323 (2002) (cit. on p. 128).

[241] X.-b. Wang, "Theorem for the beam-splitter entangler", Physical Review A **66**, 024303 (2002) (cit. on p. 128).

[242] M. M. Wolf, J. Eisert, and M. B. Plenio, "Entangling power of passive optical elements", Physical Review Letters **90**, 047904 (2003) (cit. on p. 128).

[243] J. K. Asbóth, J. Calsamiglia, and H. Ritsch, "Computable Measure of Nonclassicality for Light", Physical Review Letters **94**, 173602 (2005) (cit. on pp. 128, 130).

[244] J. S. Ivan, S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda, and R. Simon, "Entanglement and nonclassicality for multimode radiation-field states", Phys. Rev. A **83**, 032118 (2011) (cit. on p. 128).

[245] A. Streltsov, H. Kampermann, and D. Bruß, "Linking quantum discord to entanglement in a measurement", Physical Review Letters **106**, 160401 (2011) (cit. on p. 128).

[246] M. Piani, S. Gharibian, G. Adesso, J. Calsamiglia, P. Horodecki, and A. Winter, "All nonclassical correlations can be activated into distillable entanglement", Physical Review Letters **106**, 220403 (2011) (cit. on p. 128).

[247] W. Ge, M. E. Tasgin, and M. S. Zubairy, "Conservation relation of nonclassicality and entanglement for Gaussian states in a beam splitter", Physical Review A **92**, 052328 (2015) (cit. on p. 128).

[248] F. G. S. L. Brandão and M. B. Plenio, "Thermodynamical formulation of entanglement theory.", Nature Phys. **4**, 873 (2008) (cit. on p. 128).

[249] Y.-R. Zhang, L.-H. Shao, Y. Li, and H. Fan, "Quantifying coherence in infinite-dimensional systems", Physical Review A **93**, 012334 (2016) (cit. on p. 128).

[250] A. Streltsov, E. Chitambar, S. Rana, M. Bera, A. Winter, and M. Lewenstein, "Entanglement and Coherence in Quantum State Merging", Physical Review Letters **116**, 240405 (2016) (cit. on p. 128).

[251] J. Xu, "Quantifying coherence of Gaussian states", Physical Review A **93**, 032111 (2016) (cit. on p. 128).

[252] K. C. Tan, T. Volkoff, H. Kwon, and H. Jeong, "Quantifying the Coherence between Coherent States", Physical Review Letters **119**, 190405 (2017) (cit. on p. 128).

[253] C. Mukhopadhyay, S. Das, S. Bhattacharya, A. S. De, and U. Sen, "Quantum coherence in superposition with non-distinguishable pointers", arXiv:1705.04343 (2017) (cit. on p. 128).

[254] R. Gilmore, "Geometry of symmetrized states", Annals of Physics **74**, 391–463 (1972) (cit. on p. 129).

[255] A. M. Perelomov, "Coherent states for arbitrary Lie group", Communications in Mathematical Physics **26**, 222–236 (1972) (cit. on p. 129).

[256] M. Piani and G. Adesso, "Quantumness of correlations revealed in local measurements exceeds entanglement", Physical Review A **85**, 040301 (2012) (cit. on pp. 130, 135).

[257] A. S. Sørensen and K. Mølmer, "Entanglement and Extreme Spin Squeezing", Physical Review Letters **86**, 4431–4434 (2001) (cit. on p. 131).

[258] R. Jozsa and J. Schlienz, "Distinguishability of states and von Neumann entropy", Physical Review A **62**, 012301 (2000) (cit. on pp. 131, 132).

[259] A. Chefles, R. Jozsa, and A. Winter, "On the existence of physical transformations between sets of quantum states", International Journal of Quantum Information **02**, 11–21 (2004) (cit. on p. 132).

[260] W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways", Physical Review A **62**, 062314 (2000) (cit. on pp. 133, 141, 148, 150).

[261] S. Hill and W. K. Wootters, "Entanglement of a Pair of Quantum Bits", Physical Review Letters **78**, 5022–5025 (1997) (cit. on pp. 141, 142, 147).

[262] F. Verstraete, J. Dehaene, and B. De Moor, "Normal forms and entanglement measures for multipartite quantum states", Physical Review A **68**, 012103 (2003) (cit. on pp. 142, 144, 150).

[263] C. Eltschka, T. Bastin, A. Osterloh, and J. Siewert, "Multipartite entanglement monotones and polynomial invariants", Physical Review A **85**, 022301 (2012) (cit. on pp. 142, 144).

[264] A. Osterloh and J. Siewert, "Constructing N-qubit entanglement monotones from antilinear operators", Physical Review A **72**, 012337 (2005) (cit. on pp. 142, 150).

[265] D. Z. Djokovic and A. Osterloh, "On polynomial invariants of several qubits", Journal of Mathematical Physics **50**, 033509 (2009) (cit. on p. 142).

[266] G. Gour and N. R. Wallach, "Classification of Multipartite Entanglement of All Finite Dimensionality", Physical Review Letters **111**, 060502 (2013) (cit. on p. 142).

[267] V. Coffman, J. Kundu, and W. K. Wootters, "Distributed entanglement", Physical Review A **61**, 052306 (2000) (cit. on pp. 142, 147, 150, 151).

[268] R. Lohmayer, A. Osterloh, J. Siewert, and A. Uhlmann, "Entangled Three-Qubit States without Concurrence and Three-Tangle", Physical Review Letters **97**, 260502 (2006) (cit. on pp. 142, 143, 147).

[269] C. Eltschka, A. Osterloh, J. Siewert, and A. Uhlmann, "Three-tangle for mixtures of generalized GHZ and generalized W states", New Journal of Physics **10**, 043014 (2008) (cit. on pp. 142, 147, 151).

[270] E. Jung, M.-R. Hwang, D. K. Park, and J.-W. Son, "Three-tangle for rank-three mixed states: mixture of greenberger-horne-zeilinger, W, and flipped-W states", Physical Review A **79**, 024306 (2009) (cit. on pp. 142, 147).

[271] E. Jung, D. K. Park, and J.-W. Son, "Three-tangle does not properly quantify tripartite entanglement for Greenberger-Horne-Zeilinger--type states", Physical Review A **80**, 010301 (2009) (cit. on pp. 142, 147).

[272] S.-J. He, X.-H. Wang, S.-M. Fei, H.-X. Sun, and Q.-Y. Wen, "Three-Tangle for High-Rank Mixed States", Communications in Theoretical Physics **55**, 251 (2011) (cit. on pp. 142, 147).

[273] J. Siewert and C. Eltschka, "Quantifying Tripartite Entanglement of Three-Qubit Generalized Werner States", Physical Review Letters **108**, 230502 (2012) (cit. on pp. 142, 147).

[274] O. Viehmann, C. Eltschka, and J. Siewert, "Rescaling multipartite entanglement measures for mixed states", Applied Physics B **106**, 533–541 (2012) (cit. on pp. 142, 145, 147, 148).

[275] C. Eltschka and J. Siewert, "Entanglement of Three-Qubit Greenberger-Horne-Zeilinger Symmetric States", Physical Review Letters **108**, 020502 (2012) (cit. on pp. 142, 147).

[276] E. Jung and D. K. Park, "Entanglement of four-qubit rank-2 mixed states", Quantum Information Processing **14**, 3317–3333 (2015) (cit. on p. 142).

[277] A. Osterloh, J. Siewert, and A. Uhlmann, "Tangles of superpositions and the convex-roof extension", Physical Review A **77**, 032310 (2008) (cit. on pp. 142, 143).

[278] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera, "Classification of Mixed Three-Qubit States", Physical Review Letters **87**, 040401 (2001) (cit. on pp. 143, 149).

[279] S. Rodriques, N. Datta, and P. Love, "Bounding polynomial entanglement measures for mixed states", Physical Review A **90**, 012340 (2014) (cit. on pp. 143, 149, 152).

[280] C. Carathéodory, *Theory of functions of a complex variable* (American Mathematical Soc., 1954), 320 pp. (cit. on p. 144).

[281] M. Berger, *Geometry I* (Springer, 1987) (cit. on p. 145).

[282] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, "Four qubits can be entangled in nine different ways", Physical Review A **65**, 052112 (2002) (cit. on pp. 148, 149).

[283] O. Chterental and D. Z. Djokovic, "Normal Forms and Tensor Ranks of Pure States of Four Qubits", in *Linear Algebra Research Advances*, edited by G. D. Ling (Nova Science Publishers, New York, 2007), pp. 133–167 (cit. on p. 148).

[284] L. Chen and Y.-X. Chen, "Classification of ghz-type, W-type, and ghz-W-type multiqubit entanglement", Physical Review A **74**, 062310 (2006) (cit. on p. 148).

[285] L. Chen and Y.-X. Chen, "Range criterion and classification of true entanglement in a $2 \times M \times N$ system", Physical Review A **73**, 052310 (2006) (cit. on p. 148).

[286] A. Streltsov, G. Adesso, M. Piani, and D. Bruß, "Are General Quantum Correlations Monogamous?", Physical Review Letters **109**, 050503 (2012) (cit. on p. 150).

[287] C. Lancien, S. Di Martino, M. Huber, M. Piani, G. Adesso, and A. Winter, "Should Entanglement Measures be Monogamous or Faithful?", Physical Review Letters **117**, 060501 (2016) (cit. on p. 150).

[288] T. J. Osborne and F. Verstraete, "General Monogamy Inequality for Bipartite Qubit Entanglement", Physical Review Letters **96**, 220503 (2006) (cit. on p. 150).

[289] C. Eltschka, A. Osterloh, and J. Siewert, "Possibility of generalized monogamy relations for multipartite entanglement beyond three qubits", Physical Review A **80**, 032313 (2009) (cit. on p. 150).

[290] G. Gour and N. R. Wallach, "All maximally entangled four-qubit states", Journal of Mathematical Physics **51**, 112201 (2010) (cit. on p. 150).

[291] M. F. Cornelio, "Multipartite monogamy of the concurrence", Physical Review A **87**, 032330 (2013) (cit. on p. 150).

[292] C. Eltschka and J. Siewert, "Monogamy Equalities for Qubit Entanglement from Lorentz Invariance", Physical Review Letters **114**, 140402 (2015) (cit. on p. 150).

[293] G. Adesso and F. Illuminati, "Strong Monogamy of Bipartite and Genuine Multipartite Entanglement: The Gaussian Case", Physical Review Letters **99**, 150501 (2007) (cit. on p. 151).

[294] J. S. Kim, "Strong monogamy of quantum entanglement for multiqubit W-class states", Physical Review A **90**, 062306 (2014) (cit. on p. 151).

[295] S. Karmakar, A. Sen, A. Bhar, and D. Sarkar, "Strong monogamy conjecture in a four-qubit system", Physical Review A **93**, 012327 (2016) (cit. on p. 151).

[296] F. Zhang, ed., *The Schur Complement and Its Applications* (Springer, Mar. 4, 2005), 320 pp. (cit. on p. 179).

[297] R. Bhatia, *Positive Definite Matrices* (Princeton University Press, 2007), JSTOR: j.ctt7rxv2 (cit. on p. 179).

[298] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd ed., CMS Books in Mathematics (Springer-Verlag, New York, 2003) (cit. on p. 179).

Proofs and auxiliary results



CHAPTER 7

Theorem 7.16. *For any pure state φ and any m, we have*

$$P_{\mathrm{DIO}}(\varphi \to \Psi_m, 0) = P_{\mathrm{(S)IO}}(\varphi \to \Psi_m, 0). \tag{A.1}$$

Proof. We will assume without loss of generality that the coefficients of $|\varphi\rangle$ are non-negative and arranged in non-increasing order. Let

$$k \coloneqq \arg\min_{1 \le j \le m} \frac{1}{j} \sum_{i=m-j+1}^{d} \varphi_i^2 \tag{A.2}$$

such that $P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) = \frac{m}{k} \sum_{i=m-k+1}^{d} \varphi_i^2$. Note that if $k = m$, then $1 \ge P_{\mathrm{DIO}}(\varphi \to \Psi_m, 0) \ge P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) = 1$. In the following, we will therefore assume that $1 \le k \le m-1$. Further, let us begin by considering strictly positive ε.

Define

$$\begin{aligned} \varphi_A &\coloneqq \sum_{i,j=1}^{m-k} |i\rangle\langle i|\, \varphi\, |j\rangle\langle j| \\ \varphi_B &\coloneqq \sum_{i,j=m-k+1}^{d} |i\rangle\langle i|\, \varphi\, |j\rangle\langle j| \end{aligned} \tag{A.3}$$

and notice that $P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) = \frac{m}{k} \operatorname{Tr} \varphi_B$.

Taking the dual of the SDP for the maximum distillation probability $P_{\mathrm{DIO}}(\rho \to \Psi_m, \varepsilon)$, we get

$$\begin{aligned} \underset{X,P,\lambda}{\text{minimise}} \quad & \operatorname{Tr} X \\ \text{subject to} \quad & m\Delta(\rho) + m\Delta(P) - P - m\Delta(X) + \lambda\rho - m(1-\varepsilon)\lambda\Delta(\rho) \succeq 0 \\ & P, X \succeq 0 \\ & \lambda \in \mathbb{R}_+. \end{aligned} \tag{A.4}$$

We will take the ansatz

$$\begin{aligned} P &= \mu_1 \varphi_B - \mu_2 \varphi_A + \left[(1-\varepsilon)\lambda - 1\right] \varphi \\ X &= \mu_1 \Delta(\varphi_B) \end{aligned} \tag{A.5}$$

for some coefficients $\mu_1, \mu_2 \in \mathbb{R}_+$. Our aim will now be to show that there exists a suitable choice of λ, μ_1, and μ_2 such that (X, P, λ) is feasible for the SDP (A.4).

To this end, denote by Θ the all-ones matrix of appropriate size, and notice that the first condition in (A.4) reduces to

$$\begin{aligned} 0 &\geq m\Delta(\varphi) + m\Delta(P) - P - m\Delta(X) + \lambda\varphi - m(1-\varepsilon)\lambda\Delta(\varphi) \\ &= \mu_2\left(\varphi_A - m\Delta(\varphi_A)\right) - \mu_1\varphi_B + (1+\lambda\varepsilon)\varphi \\ &= \begin{pmatrix} \left[\mu_2 + 1 + \lambda\varepsilon\right]\Theta_A - m\mu_2\mathbb{1}_A & \left[1+\lambda\varepsilon\right]\Theta_O \\ \left[1+\lambda\varepsilon\right]\Theta_O^T & \left[1+\lambda\varepsilon-\mu_1\right]\Theta_B \end{pmatrix} \circ \varphi \\ &=: Q \circ \varphi \end{aligned} \tag{A.6}$$

where $\circ$ denotes the Schur product. Showing that $Q \preceq 0$ will then imply the desired relation $Q \circ \varphi \preceq 0$ by the Schur product theorem. From the generalised Schur complement condition, $Q \preceq 0$ if and only if the following all hold [296,297]:

[296] Zhang, 2005

[297] Bhatia, 2007

$$\begin{aligned} &\text{(i)} && \left(\mu_2 + 1 + \lambda\varepsilon\right)\Theta_A - m\mu_2\mathbb{1}_A \preceq 0, \\ &\text{(ii)} && \left(1 + \lambda\varepsilon - \mu_1\right)\Theta_B \preceq 0, \\ &\text{(iii)} && \left(\mu_2 + 1 + \lambda\varepsilon\right)\Theta_A - m\mu_2\mathbb{1}_A - (1+\lambda\varepsilon)^2\Theta_O\left(\left[1+\lambda\varepsilon-\mu_1\right]\Theta_B\right)^{-1}\Theta_O^T \preceq 0 \end{aligned} \tag{A.7}$$

where M^{-1} denotes the Moore-Penrose pseudoinverse, which in particular satisfies $M^{-1} = M/\operatorname{Tr}(M)^2$ for any rank-one Hermitian M [298]. From the first two conditions, we have

[298] Ben-Israel & Greville, 2003

$$\begin{aligned} &\text{(i)} && \mu_2 \geq (1+\lambda\varepsilon)\frac{m-k}{k}, \\ &\text{(ii)} && \mu_1 \geq 1 + \lambda\varepsilon, \end{aligned} \tag{A.8}$$

and the third condition reduces to

$$\begin{aligned} 0 &\geq \left(\mu_2 + 1 + \lambda\varepsilon\right)\Theta_A - m\mu_2\mathbb{1}_A - \frac{(1+\lambda\varepsilon)^2}{\left[1+\lambda\varepsilon-\mu_1\right]\operatorname{Tr}\Theta_B^2}\Theta_O\Theta_B\Theta_O^T \\ &= \left(\mu_2 + 1 + \lambda\varepsilon\right)\Theta_A - m\mu_2\mathbb{1}_A - \frac{(1+\lambda\varepsilon)^2}{1+\lambda\varepsilon-\mu_1}\Theta_A \\ &= \left[\mu_2 - \frac{(1+\lambda\varepsilon)\mu_1}{1+\lambda\varepsilon-\mu_1}\right]\Theta_A - m\mu_2\mathbb{1}_A. \end{aligned} \tag{A.9}$$

Noting that the coefficient $\mu_2 - \frac{(1+\lambda\varepsilon)\mu_1}{1+\lambda\varepsilon-\mu_1}$ can never be negative when conditions (i) and (ii) are satisfied, we have

$$\text{(iii)} \quad m\mu_2 \geq (m-k)\left[\mu_2 - \frac{(1+\lambda\varepsilon)\mu_1}{1+\lambda\varepsilon-\mu_1}\right]. \tag{A.10}$$

To ensure that $P \succeq 0$, notice that since

$$P = \begin{pmatrix} \left[-\mu_2 + (1-\varepsilon)\lambda - 1\right]\Theta_A & \left[(1-\varepsilon)\lambda - 1\right]\Theta_O \\ \left[(1-\varepsilon)\lambda - 1\right]\Theta_O^T & \left[\mu_1 + (1-\varepsilon)\lambda - 1\right]\Theta_B \end{pmatrix} \circ \varphi, \tag{A.11}$$

the positivity of P is equivalent to the positivity of the matrix

$$P' \coloneqq \begin{pmatrix} -\mu_2 + (1-\varepsilon)\lambda - 1 & (1-\varepsilon)\lambda - 1 \\ (1-\varepsilon)\lambda - 1 & \mu_1 + (1-\varepsilon)\lambda - 1 \end{pmatrix}. \tag{A.12}$$

We therefore have two additional conditions:

$$\begin{aligned} &\text{(iv)} \quad 0 \leq \operatorname{Tr} P' = 2(1-\varepsilon)\lambda + \mu_1 - \mu_2 - 2 \\ &\text{(v)} \quad 0 \leq \det P' = \mu_2 - (1-\varepsilon)\lambda\mu_2 - \mu_1\left[1 - (1-\varepsilon)\lambda + \mu_2\right]. \end{aligned} \tag{A.13}$$

Let us stress that conditions (i)–(v) together with $\mu_1 \geq 0$ constitute a set of sufficient conditions for a given choice of (X, P, λ) of the form (A.5) to satisfy the feasibility conditions of SDP (A.4).

We will now make the ansatz

$$\begin{aligned} \lambda &= \frac{\mu_1 - \mu_2}{\varepsilon\mu_2} \\ \mu_1 &= \frac{k\mu_2(\mu_2 - 1)}{m - k} \end{aligned} \tag{A.14}$$

leaving μ_2 as a free variable for now. With this choice, conditions (i)–(iii) are always satisfied for any $\mu_2 \in \mathbb{R}$, while the other conditions reduce to

$$\begin{aligned} &\text{(iv)} \quad 2(k\mu_2 - m) + \varepsilon(2k - (2k+m)\mu_2 + k\mu_2^2) \geq 0 \\ &\text{(v)} \quad \mu_2\left(\varepsilon km(1-\mu_2)^2 - (m - k\mu_2)^2\right) \leq 0 \end{aligned} \tag{A.15}$$

We will now make a choice of μ_2 which can be verified to satisfy the above the inequalities for any $\varepsilon < \frac{k}{m}$ as

$$\mu_2 = \frac{m + m^2\sqrt{\varepsilon}}{k - \varepsilon m} \tag{A.16}$$

where we note that $\mu_2 \geq 1$ and $\lim_{\varepsilon\to 0} \mu_1 = \lim_{\varepsilon\to 0} \mu_2 = \frac{m}{k}$. Since all conditions (i)–(v) are satisfied for our choice of (X, P, λ) with the given μ_1 and μ_2 for any $\varepsilon < \frac{k}{m}$, the triple (X, P, λ) is a valid feasible solution for the SDP (A.4). This means in particular that for any $0 < \varepsilon < \frac{k}{m}$ we have

$$P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \leq \operatorname{Tr} X = \mu_1 \operatorname{Tr} \varphi_B. \tag{A.17}$$

Using the fact that SIO $\subset$ DIO as well as that $P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \geq P_{\mathrm{DIO}}(\varphi \to \Psi_m, 0)$ for any ε, this then gives

$$\begin{aligned} P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) &\leq P_{\mathrm{DIO}}(\varphi \to \Psi_m, 0) \\ &\leq \lim_{\varepsilon\to 0} P_{\mathrm{DIO}}(\varphi \to \Psi_m, \varepsilon) \\ &\leq \frac{m}{k} \operatorname{Tr} \varphi_B \\ &= P_{\mathrm{SIO}}(\varphi \to \Psi_m, 0) \end{aligned} \tag{A.18}$$

which completes the proof. ∎

Proposition 7.17. *For $\mathcal{O} \in \{DIO, MIO\}$ and any pure state φ with $\varphi_1 \geq \ldots \geq \varphi_n > 0$, it holds that*

$$P_{\mathcal{O}}(\varphi \to \Psi_2, \varepsilon) = \begin{cases} 1 & \text{if } \varepsilon \geq \varepsilon_0(\varphi_1), \\ 2(1-\varphi_1^2)\left(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon}\right)^2 & \text{otherwise,} \end{cases}$$

$$\text{where } \varepsilon_0(\varphi_1) \coloneqq \begin{cases} 0 & \text{if } \varphi_1 \leq \frac{1}{\sqrt{2}}, \\ \frac{1}{2} - \varphi_1\sqrt{1-\varphi_1^2} & \text{otherwise.} \end{cases}$$

Proof. Since the case $\varepsilon = 0$ follows from Thm. 7.16, we will assume $\varepsilon > 0$. For the case of $\varphi_1 \leq \frac{1}{\sqrt{2}}$, we have

$$P_{\mathcal{O}}(\varphi \to \Psi_2, \varepsilon) \geq P_{\text{SIO}}(\varphi \to \Psi_2, 0) = 1. \tag{A.19}$$

From the result in Thm. 7.11, we know that if $\varepsilon \geq \frac{1}{2} - \varphi_1\sqrt{1-\varphi_1^2}$, then $P_{\text{DIO}}(\varphi \to \Psi_2, \varepsilon) = 1$. In the following, we therefore only consider the case $\varepsilon < \frac{1}{2} - \varphi_1\sqrt{1-\varphi_1^2}$. We prove this result by explicit constructing feasible solutions in both primal and dual SDPs. The primal SDP under DIO is given by

$$P_{\text{DIO}}(\varphi \to \Psi_2, \varepsilon) = \max\Big\{ \operatorname{Tr} G\varphi \;\Big|\; \operatorname{Tr} C\varphi \geq (1-\varepsilon)\operatorname{Tr} G\varphi, \quad 0 \leq C \leq G \leq \mathbb{1},\ \Delta(G) = 2\Delta(C),\ G = \Delta(G)\Big\}. \tag{A.20}$$

We take the ansatz

$$G = \mathbb{1} - x\,|0\rangle\langle 0|, \quad C = \frac{1}{2}G + y\sum_{i=2}^{n}\varphi_1\varphi_n(|1\rangle\langle i| + |i\rangle\langle 1|). \tag{A.21}$$

Then we have

$$\operatorname{Tr} G\rho = 1 - x\varphi_1^2, \quad \operatorname{Tr} C\rho = \frac{1}{2}(1 - x\varphi_1^2) + 2y\varphi_1^2(1-\varphi_1^2), \tag{A.22}$$

and

$$\operatorname{spec}(C) = \operatorname{spec}(G - C) \tag{A.23}$$

$$= \left\{ \underbrace{\frac{1}{2}, \cdots, \frac{1}{2}}_{n-2 \text{ fold}}, \frac{1}{2} - \frac{x}{4} - \sqrt{y^2\varphi_1^2(1-\varphi_1^2) + \frac{x^2}{16}}, \frac{1}{2} - \frac{x}{4} + \sqrt{y^2\varphi_1^2(1-\varphi_1^2) + \frac{x^2}{16}} \right\}. \tag{A.24}$$

Then we have the relaxation

$$\begin{aligned} P_{\mathrm{DIO}}(\varphi \to \Psi_2, \varepsilon) \geq \text{maximise} \quad & 1 - x\varphi_1^2 \\ \text{subject to} \quad & 4y\varphi_1^2(1-\varphi_1^2) = (1-2\varepsilon)(1-x\varphi_1^2), \\ & 1 - x \geq 4y^2\varphi_1^2(1-\varphi_1^2), \\ & 0 \leq x \leq 1. \end{aligned} \tag{A.25}$$

By choosing

$$x = \frac{1 - 2(1-\varphi_1^2)(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon})^2}{\varphi_1^2}, \quad y = \frac{(\sqrt{1-\varepsilon}+\sqrt{\varepsilon})^2}{2\varphi_1^2(1-2\varepsilon)}, \tag{A.26}$$

we can verify that this is a feasible solution. Thus we have

$$P_{\mathrm{DIO}}(\varphi \to \Psi_2, \varepsilon) \geq 2(1-\varphi_1^2)\left(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon}\right)^2. \tag{A.27}$$

As for the dual problem, we consider the dual SDP under MIO,

$$\begin{aligned} P_{\mathrm{MIO}}(\varphi \to \Psi_2, \varepsilon) = \text{minimise} \quad & \mathrm{Tr}\, Y \\ \text{subject to} \quad & (1-(1-\varepsilon)x)\varphi + X + \Delta(Z) \preceq Y, \\ & x\varphi - X - 2\Delta(Z) \preceq 0, \\ & x \geq 0, X \geq 0, Y \geq 0 \end{aligned} \tag{A.28}$$

Taking

$$x = \frac{(\sqrt{1-\varepsilon}+\sqrt{\varepsilon})^2}{\sqrt{1-\varepsilon}\sqrt{\varepsilon}(1-2\varepsilon)}, \quad Y = 2\left(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon}\right)^2 (\mathbb{1} - |0\rangle\langle 0|)\varphi(\mathbb{1} - |0\rangle\langle 0|), \tag{A.29}$$

$$Z = \frac{2}{1-2\varepsilon}|0\rangle\langle 0|\varphi|0\rangle\langle 0|, \quad X = Y - \Delta(Z) - (1-(1-\varepsilon)x)\varphi, \tag{A.30}$$

we can verify that $\{x, X, Y, Z\}$ is a valid feasible solution. Thus

$$P_{\mathrm{MIO}}(\varphi \to \Psi_2, \varepsilon) \leq 2(1-\varphi_1^2)\left(\frac{\sqrt{1-\varepsilon}+\sqrt{\varepsilon}}{1-2\varepsilon}\right)^2. \tag{A.31}$$

Combining Eqs. (A.27) and (A.31), we have the desired result. ■

CHAPTER 8

Theorem 8.12. *The one-shot entanglement cost under the classes of operations* PPT, P-PPT, *and* P-PPT$_+$ *is given by*

$$\begin{aligned} E_{c,\text{P-PPT}_+}^{(1),\varepsilon}(\rho) &= \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \log\left(R_{\text{PPT}_+}^{\text{PPT}_+}(\widetilde{\rho})+1\right)\right\rceil_{\log} \\ E_{c,\text{PPT}}^{(1),\varepsilon}(\rho) &= E_{c,\text{P-PPT}}^{(1),\varepsilon}(\rho) = \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \Upsilon_{\text{PPT}}(\widetilde{\rho})\right\rceil_{\log} \end{aligned} \tag{A.32}$$

where

$$\begin{aligned} \Upsilon_{\text{PPT}}(\rho) &\coloneqq \log \min_{\omega\in\text{PPT}_+}\left\{\lambda\in\mathbb{R}_+ \;\middle|\; \frac{\rho+(\lambda-1)\omega}{\lambda}\in\text{PPT},\ \frac{(\lambda+1)\omega-\rho}{\lambda}\in\text{PPT}\right\} \\ &= \log \min_{Z\in\mathbb{H}_+}\left\{ m\operatorname{Tr}Z \;\middle|\; -(m-1)Z^{T_B}\le\rho^{T_B}\le(m+1)\operatorname{Tr}Z^{T_B}\right\}. \end{aligned} \tag{A.33}$$

For convenience, we will split the proof of the Theorem into separate lemmas.

Lemma A.4. $E_{c,\text{P-PPT}_+}^{(1),\varepsilon}(\rho) = \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \log\left(R_{\text{PPT}_+}^{\text{PPT}_+}(\widetilde{\rho})+1\right)\right\rceil_{\log}.$

Proof. Since Ψ_m is invariant under any unitary of the form $U\otimes U^*$, it is in particular invariant under the twirling $\mathcal{T}$, and we can without loss of generality consider only operations of the form $\Lambda=\Lambda\circ\mathcal{T}$, giving

$$\Lambda(\rho) = \operatorname{Tr}(\rho\Psi_m)X + [1-\operatorname{Tr}(\rho\Psi_m)]\,Z. \tag{A.34}$$

for some operators X, Z. The condition that Λ is a CPTP map then constraints X and Z to be density operators. To constrain the operation Λ to be in P-PPT$_+$, we restrict without loss of generality to isotropic input states of the form

$$I_f \coloneqq f\Psi_m + \frac{1-f}{m^2-1}(\mathbb{1}-\Psi_m). \tag{A.35}$$

Recalling that $I_f\in\text{SEP}$ iff $I_f\in\text{PPT}_+$ iff $0\le f\le\frac{1}{m}$ [230], we have that every PPT isotropic state can be written as

[230] Horodecki & Horodecki, PRA 1999

$$qI_{\frac{1}{m}} + (1-q)\frac{\mathbb{1}-\Psi_m}{m^2-1} \tag{A.36}$$

with $q\in[0,1]$, and therefore for any $\Lambda\in\text{PPT}_+$ it is sufficient and necessary to impose that

$$\begin{aligned} \Lambda\left(I_{\frac{1}{m}}\right) &= \frac{X+(m-1)Z}{m}\in\text{PPT}_+ \\ \Lambda\left(\frac{\mathbb{1}-\Psi_m}{m^2-1}\right) &= Z\in\text{PPT}_+. \end{aligned} \tag{A.37}$$

Noticing that $\Lambda(\Psi_m) = X$, we then have

$$\begin{aligned}
&\max\left\{ F(\rho, \Lambda(\Psi_m)) \;\middle|\; \Lambda \in \text{P-PPT}_+ \right\} \\
&= \max\left\{ F(\rho,\omega) \;\middle|\; \exists\, \sigma \in \mathsf{PPT}_+ \text{ s.t. } \frac{\omega + (m-1)\sigma}{m} \in \mathsf{PPT}_+ \right\} \\
&= \max\left\{ F(\rho,\omega) \;\middle|\; R^{\mathsf{PPT}_+}_{\mathsf{PPT}_+}(\omega) \leq m-1 \right\}
\end{aligned} \tag{A.38}$$

and the result follows by definition of entanglement cost. ∎

Lemma A.5. $E^{(1),\varepsilon}_{c,\text{P-PPT}}(\rho) = \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \Upsilon_{\mathsf{PPT}}(\widetilde{\rho}) \right\rceil_{\log}.$

Proof. The argument is similar to Lemma A.4. To constrain the operation Λ to be in P-PPT, we again restrict without loss of generality to isotropic normalised operators

$$I_f \coloneqq f\Psi_m + \frac{1-f}{m^2-1}\left(\mathbb{1} - \Psi_m\right) \tag{A.39}$$

where we note that $I_f \in \mathsf{PPT}$ iff $-\frac{1}{m} \leq f \leq \frac{1}{m}$. This means that every isotropic PPT operator of trace 1 can be written as either

$$qI_{\frac{1}{m}} + (1-q)\frac{\mathbb{1}-\Psi_m}{m^2-1} \tag{A.40}$$

or

$$qI_{-\frac{1}{m}} + (1-q)\frac{\mathbb{1}-\Psi_m}{m^2-1}. \tag{A.41}$$

with $q \in [0,1]$. We therefore get the constraints:

$$\begin{aligned}
\Lambda\left(I_{\frac{1}{m}}\right) &= \frac{X + (m-1)Z}{m} \in \mathsf{PPT} \\
\Lambda\left(I_{-\frac{1}{m}}\right) &= \frac{(m+1)Z - X}{m} \in \mathsf{PPT} \\
\Lambda\left(\frac{\mathbb{1}-\Psi_m}{m^2-1}\right) &= Z \in \mathsf{PPT}
\end{aligned} \tag{A.42}$$

Recalling that the complete positivity of Λ imposes $Z \in \mathbb{D}$ and hence $Z \in \mathsf{PPT}_+$, we then get

$$\begin{aligned}
&\max\left\{ F(\rho, \Lambda(\Psi_m)) \;\middle|\; \Lambda \in \text{P-PPT} \right\} \\
&= \max\left\{ F(\rho,\omega) \;\middle|\; \exists\, \sigma \in \mathsf{PPT}_+ \text{ s.t. } \frac{\rho + (m-1)\sigma}{m} \in \mathsf{PPT},\ \frac{(m+1)\sigma - \rho}{m} \in \mathsf{PPT} \right\} \\
&= \max\left\{ F(\rho,\omega) \mid \Upsilon_{\mathsf{PPT}}(\rho) \leq m \right\}
\end{aligned} \tag{A.43}$$

as required. ∎

Lemma A.6. $E_{c,\mathrm{PPT}}^{(1),\varepsilon}(\rho) = \left\lceil \min_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \Upsilon_{\mathrm{PPT}}(\widetilde{\rho}) \right\rceil_{\log}.$

Proof. We again consider operations of the form

$$\Lambda(\rho) = \mathrm{Tr}(\rho\Psi_m)X + [1 - \mathrm{Tr}(\rho\Psi_m)]\, Z. \tag{A.44}$$

The problem of entanglement cost under PPT operations has already been considered in [227], where it was shown that any operation of the form (A.44) is PPT if and only if $-(m-1)Z^{T_B} \leq X^{T_B} \leq (m+1)Z^{T_B}$. ■

[227] Audenaert et al., PRL 2003

CHAPTER 9

Lemma 9.6. *Given a state of the form*

$$|\psi\rangle = \sum_{i=1}^{d} c_i\, |i\rangle\, |\underline{2^{d-i}}\rangle \tag{A.45}$$

where we can take $|c_1| \geq |c_2| \geq \ldots \geq |c_d|$ without loss of generality, the closest k-producible state with respect to the fidelity-based geometric measure of entanglement can be chosen as a state $|\psi^c\rangle \in \mathcal{B}^{(k)}$.
In other words, there exists $|\psi^c\rangle \in \mathcal{B}^{(k)}$ such that

$$\max_{|\varsigma\rangle\in\mathcal{P}^{(k)}} F(|\psi\rangle, |\varsigma\rangle) = \max_{|\sigma\rangle\in\mathcal{B}^{(k)}} F(|\psi\rangle, |\sigma\rangle) = F(|\psi\rangle, |\psi^c\rangle). \tag{A.46}$$

Proof. Recall that, by Thm. 1, a state in $\mathcal{B}^{(k)}$ has p nonzero coefficients ($2 \leq p \leq d$) iff it is $p+1$-partite entangled. Therefore, we have that any $|\omega\rangle \in \mathcal{B}^{(k)}$ is given by

$$|\omega\rangle = \sum_{j\in\mathcal{J}} d_j\, |j\rangle\, |\underline{2^{d-j}}\rangle \tag{A.47}$$

where $\mathcal{J}$ is a subset of at most $k-1$ elements of $\{1,\ldots,d\}$. Hence

$$\begin{aligned} F(|\psi\rangle, |\omega\rangle) &= |\langle\psi|\omega\rangle|^2 \\ &= \left| \sum_{i=1}^{d} \sum_{j\in\mathcal{J}} \overline{c_i} d_j \langle i|j\rangle \langle \underline{2^{d-i}} | \underline{2^{d-j}} \rangle \right|^2 \\ &= \left| \sum_{j\in\mathcal{J}} \overline{c_j} d_j \right|^2 \\ &\leq \sum_{j\in\mathcal{J}} |c_j|^2 \sum_{j\in\mathcal{J}} |d_j|^2 \\ &= \sum_{j\in\mathcal{J}} |c_j|^2 \\ &\leq \sum_{i=1}^{k-1} |c_i|^2 \end{aligned} \tag{A.48}$$

using the Cauchy-Schwarz inequality and the fact that the coefficients $|c_i|$ are arranged in non-increasing order, so the choice of the first $k-1$ coefficients maximises the expression. This bound is saturated by the renormalised state given by

$$|\psi^c\rangle = \frac{\sum_{i=1}^{k-1} c_i |i\rangle |\underline{2^{d-i}}\rangle}{\sqrt{\sum_{j=1}^{k-1} |c_i|^2}} \in \mathcal{B}^{(k)} \tag{A.49}$$

and so it is tight.

Therefore, we need to show that for any $|\phi\rangle \in \mathcal{P}^{(k)}$ we have $|\langle\psi|\phi\rangle|^2 \leq \sum_{j=1}^{k-1} |c_i|^2$.

For a general k-producible state, we have to consider several different cases corresponding to different partitions. Consider a general k-producible state $|\phi^{(n)}\rangle$, where the qudit is entangled with n qubits and the remaining $(d-n)$ qubits are in some arrangement — we do not explicitly assume anything about the state of the $(d-n)$ qubits. The state has the following form:

$$\begin{aligned} |\phi^{(n)}\rangle &= |\text{qudit} + n \text{ qubits}\rangle \otimes |d-n \text{ qubits}\rangle \\ &= |\phi_A^{(n)}\rangle \otimes |\phi_B^{(n)}\rangle . \end{aligned} \tag{A.50}$$

We necessarily have that $n \leq k-1$, because otherwise $|\phi^{(n)}\rangle$ would not be k-producible.

Note that for a general bipartite state $|\alpha_{AB}\rangle$, it follows from the Schmidt decomposition that [118]

[118] Shimony, Ann. NY Ac. 1995

$$\max_{\text{separable } |\eta_{AB}\rangle} |\langle\eta_{AB}|\alpha_{AB}\rangle|^2 = \lambda_{\max}\Big(\text{Tr}_A \left(|\alpha_{AB}\rangle \langle\alpha_{AB}| \right) \Big) \tag{A.51}$$

where $\lambda_{\max}$ denotes the largest eigenvalue. In our case, we can treat $|\psi\rangle$ as a bipartite state where the subsystem A is comprised of the qudit and n qubits,

and $|\phi^{(n)}\rangle$ is the corresponding product state. We then have

$$\begin{aligned}
\mathrm{Tr}_{(\text{qudit})}\left(|\psi\rangle\langle\psi|\right) &= \sum_{i,j,m=1}^{d} c_i\overline{c_j}\,\langle m|i\rangle\,\langle j|m\rangle\,|\underline{2^{d-i}}\rangle_d\,\langle\underline{2^{d-j}}| \\
&= \sum_{i=1}^{d} |c_i|^2\,|\underline{2^{d-i}}\rangle_d\,\langle\underline{2^{d-i}}| \\
&=: \rho' \\
\mathrm{Tr}_{(n\text{ leftmost qubits})}(\rho') &= \sum_{j=1}^{n} |c_j|^2\,|00\ldots0\rangle\langle 00\ldots0| + \\
&\quad + \sum_{i=n+1}^{d} |c_i|^2\,|\underline{2^{d-i}}\rangle_{d-n}\,\langle\underline{2^{d-i}}| \\
&=: \rho^{(n)}
\end{aligned} \tag{A.52}$$

where we have introduced a subscript in the notation $|2^{d-i}\rangle_{d-n}$ to indicate that there are $d-n$ qubits left over. Notice that the only possible nonzero eigenvalues of $\rho^{(n)}$ are given by: $\sum_{i=1}^{n}|c_i|^2, |c_{n+1}|^2, |c_{n+2}|^2, \ldots, |c_d|^2$. Since the coefficients $|c_i|$ are arranged in non-increasing order by assumption, it follows that

$$\max_{|\phi^{(n)}\rangle\in\mathcal{P}^{(k)}} |\langle\phi^{(n)}|\psi\rangle|^2 \leq \lambda_{\max}\left(\rho^{(n)}\right) = \sum_{i=1}^{n}|c_i|^2 \leq \sum_{j=1}^{k-1}|c_i|^2 \tag{A.53}$$

as required. ■