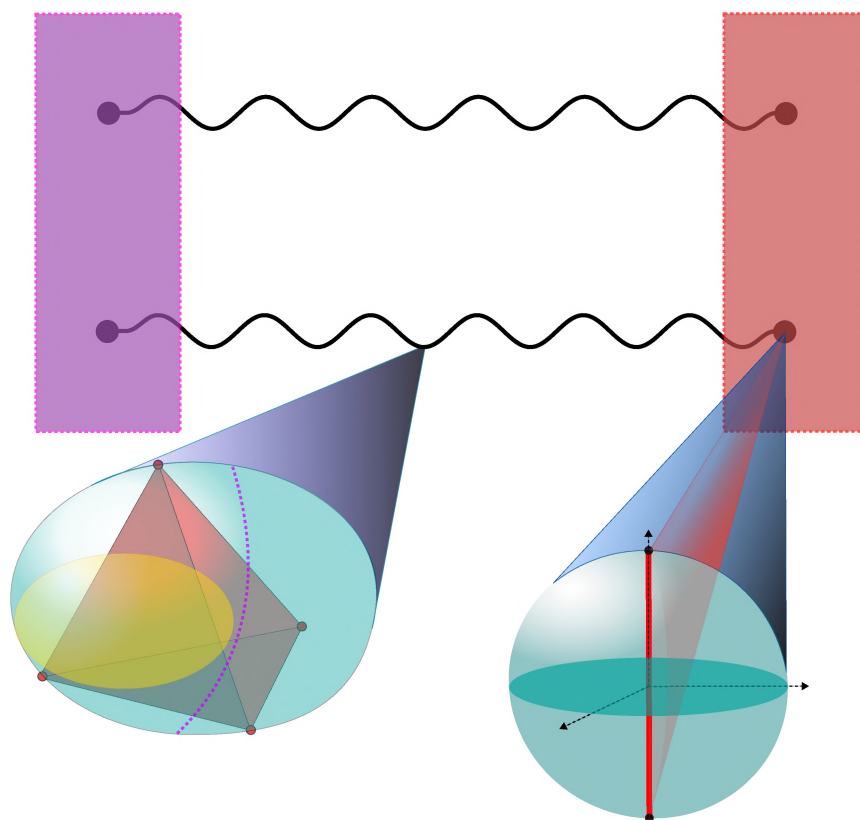


CONVEX STRUCTURES IN QUANTUM INFORMATION

Symmetric and geometric methods



Albert Rico

Convex structures in quantum information

Symmetric and geometric methods

Doctoral Thesis
Physical sciences



Albert Rico

Supervisor: Dr. Felix Huber
Institute of Theoretical Physics

Jagiellonian University, Kraków
Faculty of Physics, Astronomy and Applied Informatics
September 2024

A la meua família, amb amor.

Oświadczenie

Ja niżej podpisany Albert Rico, doktorant Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego oświadczam, że przedłożona przeze mnie rozprawa doktorska pt. “Convex structures in quantum information: Symmetric and geometric methods” jest oryginalna i przedstawia wyniki badań wykonanych przeze mnie osobiście, pod kierunkiem dr. hab. Felix Huber. Prace napisałem samodzielnie. Oświadczam, że moja rozprawa doktorska została opracowana zgodnie z Ustawą o prawie autorskim i prawach pokrewnych z dnia 4 lutego 1994 r. (Dziennik Ustaw 1994 nr 24 poz. 83 wraz z późniejszymi zmianami). Jestem świadom, że niezgodność niniejszego oświadczenia z prawdą ujawniona w dowolnym czasie, niezależnie od skutków prawnych wynikających z ww. ustawy, może spowodować unieważnienie stopnia nabytego na podstawie tej rozprawy.

Kraków, 12 września 2024

Streszczenie

Informacja kwantowa ma na celu ustalenie matematycznych podstaw do opisu przyrody w małych skalach. Aby to zrobić, opiera się na założeniu, że układy fizyczne można opisać za pomocą *stanów* w przestrzeni wektorowej i ich wypukłych kombinacji; a działające na nie operacje można opisać przekrojem zbioru stanów większych układów wymiarowych. Dlatego zarówno podstawowe pytania, jak i zastosowania technologii kwantowych często dotyczą wewnętrznej struktury wybranych wypukłych zbiorów zainteresowań. Jednak w praktyce przestrzeń zarówno stanów, jak i operacji opisujących układ kwantowy rośnie wykładniczo wraz z rozmiarem układu. Jest to nieodłącznie związane z obosiecznym charakterem mechaniki kwantowej: stanowi ona obiecujące ramy dla nowych możliwości, które są nierealne w teorii klasycznej, za cenę znacznego wzrostu złożoności. Dlatego zarówno zrozumienie teorii kwantowej, jak i jej wykorzystanie do nowatorskich celów praktycznych, wymagają nowych technik w celu odkrycia konkretnych rozwiązań leżących w jej złożonej strukturze.

Niniejsza rozprawa koncentruje się na wykorzystaniu dwóch typów narzędzi do odpowiedzi na różnorodne pytania w tej dziedzinie: identyfikowaniu i wykorzystywaniu symetrii w celu znalezienia rozwiązania określonego problemu oraz wykorzystaniu analizy geometrycznej do odkrycia nowych właściwości i połączeń. Pierwsza praca dotyczy detekcji splątania: wprowadzamy metodę nieliniową wykorzystującą identyczne kopie stanu rzeczywistego i standardowych świadków jednokopiowych, która wykazuje właściwości symetryczne i może być realizowana w eksperymentach. W drugiej pracy wykorzystujemy techniki redukcji symetrii do rozwiązywania problemów optymalizacji wypukłej niezmienników pod kopiami tensorowymi operatorów unitarnych. Trzecia praca dotyczy uogólnionych pomiarów kwantowych, dla których ustalamy charakterystykę analogiczną do wielotopu Birkhoffa (w sformułowaniu Ostrowskiego) poprzez relacje majoryzacji operatora. W czwartej pracy identyfikujemy zasoby nielokalne w dwudzielnych, zaszumionych operacjach kwantowych na podstawie eksperymentalnie dostępnych rozkładów prawdopodobieństwa. Analiza ta prowadzi do piątej pracy, w której określamy ilościowo zarówno ogólne, jak i dekoherentne nielokalne zasoby kanałów kwantowych i ujawniamy rolę, jaką w dobrze znanych protokołach kwantowych odgrywa nielokalność, odporna na dekoherencję. Wyniki te rzucają światło na współdzielone zasoby potrzebne do realizacji operacji nielokalnych i ich komponent odporny na skutki dekoherencji występujące w realistycznych urządzeniach.

Abstract

Quantum information aims to establish a mathematical ground to describe nature at small scales. To do so, it is based on the assumption that physical systems can be described by *states* in a vector space and their convex combinations; and the *operations* acting on them can be described by a cross-section of the set of states of larger dimensional systems. Therefore, both fundamental questions and applications to quantum technologies are often concerned with the inner structure of selected convex sets of interest. However, in practice the space of both states and operations describing a quantum system grows exponentially with the system size. This is inherent to the double-edged nature of quantum mechanics: it constitutes a promising framework for new possibilities that are unrealistic in classical theory, at the price of a significant growth of complexity. Therefore, both the understanding of quantum theory and its use for novel practical purposes require new techniques to uncover specific solutions laying within its complex structure.

This dissertation focuses on using two types of tools to answer diverse questions in this realm: identifying and exploiting symmetries to access a the solution to a certain problem, and employing a geometric analysis to uncover new properties and connections. The first work concerns entanglement detection: we introduce a nonlinear method using identical copies of the state in hand and standard single-copy witnesses, which displays symmetric properties and can be realized in experiments. In the second work, we employ symmetry reduction techniques to solve convex optimization problems invariant under tensor copies of unitary operators. The third work concerns generalized quantum measurements, for which we establish a characterization analogous to the Birkhoff polytope (in the Ostrowski formulation) through an operator majorization relation. In the fourth work we identify nonlocal resources in bipartite noisy quantum operations, from experimentally accessible probability distributions. This analysis follows up to the fifth work, where we quantify both the overall and decoherent nonlocal resources of quantum channels, and reveal the role that nonlocality withstanding decoherence plays in well-known quantum protocols. These results shed light into the shared resources needed to implement nonlocal operations, and their component resistant to the effects of decoherence occurring in realistic devices.

Acknowledgements

First of all, I am very grateful to Felix Huber and Karol Życzkowski for giving me the opportunity to develop my PhD degree next to them. Thank you Felix for your guidance towards a focused direction through sharp, dedicated and optimistic advice, balanced with the freedom needed to explore my own approach to science. And thank you Karol for your welcoming attitude, constant support and experienced advice.

Moreover, I am thankful to a number of lovely people who have been around me during these three years and a half. Thank you Moisés Bermejo and Fereshte Shahbeigi for the uncountable hours spent together in front of a blackboard, for all I learned from you, for the shared experiences, and overall for the bond built upon. Thank you Gerard Anglès, Alexssandre de Oliveira and Steffen Kessler for the -scientific and emotional- mutual strong support, for these many scenes lived together, and for making this PhD period a fun journey every single day. Many thanks to the former and current members of the Quantum Information Team and the Jagiellonian University, for contributing to a warm environment with flowing discussions: Adam Burchardt, Ali Emami, Fionnuala Curran, Grzegorz Rajchel-Mieldzioć, Jakub Czartowski, Jakub Zakrzewski, Kamil Korzekwa, Konrad Szymański, Krzysztof Sacha, Marcin Rudziński, Oliver Reardon-Smith, Rafał Bistron, Roberto Salazar, Simon Burton, Staszek Czachórski, Vinayak Jagadish, and Wojciech Bruzda. Special thanks to Kamil for his time and effort on unconditionally helping those of us who are lucky enough to be around him. I am also thankful to Andreas Bluhm for comments and advice, as well as to Anna Sanpera, Benoît Vermersch, Cyril Branciard, Dariusz Chruściński, Guillermo García-Pérez, Jens Siewert, Mario Ziman, Matteo Rosatti, Michał Eckstein, Miguel Navascués, Paweł Horodecki, Remigiusz Augusiak, Sabrina Maniscalco, Simon Morelli and Stefano Pironio for interesting discussions.

Finally, I want to deeply thank my family, friends, and beloved, without whom this PhD would not have been possible.

Contents

Preface	12
1 Basic concepts	15
1.1 Fundamentals	16
1.1.1 Quantum states	16
1.1.2 Quantum operations and measurements	18
1.1.3 Quantum entanglement	21
1.1.4 Entanglement detection	23
1.1.5 Nonlocality	25
1.1.6 Decoherence	26
1.2 Further concepts	29
1.2.1 Semidefinite programming	29
1.2.2 The symmetric group	30
1.2.3 The set of quantum measurements	34
1.2.4 Nonlocality of bipartite channels: special subsets	36
2 Entanglement detection with trace polynomials	38
2.1 Introduction	38
2.2 Nonlinear entanglement witnesses	39
2.2.1 Detection with multiple copies	39
2.2.2 Entanglement concentration	40
2.2.3 An example: detecting noisy GHZ states	41
2.3 Trace polynomial witnesses	47
2.3.1 Witnesses from matrix inequalities	47
2.3.2 Implementation with randomized measurements	49
2.3.3 Expectation value over Haar random states	50
2.3.4 Adapting witnesses to spectra	51
2.3.5 Sampling Werner states	52
2.4 Evaluating positive maps	54
2.5 Conclusions	55
2.6 Further results: positive maps from immanant inequalities	55
2.6.1 State of the art: an overview	55
2.6.2 A new family of positive maps	56

2.6.3	Entanglement detection	58
2.6.4	Further methods of entanglement detection	60
2.6.5	Conclusions	62
3	Unitary invariant semidefinite programs: a user's guide	63
3.1	Introduction	63
3.2	Problem formulation and proof strategy	64
3.2.1	General problem formulation	64
3.2.2	Specific problem formulation	64
3.2.3	Methods	65
3.2.4	Targeted applications	65
3.2.5	Notation	66
3.3	Symmetry reduction	66
3.3.1	General approach	66
3.3.2	Schur-Weyl duality	67
3.3.3	Objective function, constraints and positivity	68
3.3.4	Main SDP symmetry reduction	69
3.4	Variable reduction	69
3.4.1	Good permutations basis	69
3.4.2	Real problems	70
3.5	Scaling	70
3.6	Direct decomposition	72
3.6.1	Objective function and positivity constraint	72
3.7	Quantum Max Cut	74
3.8	Marginal problem for Werner states	74
3.9	Infinite spin lattice with translation invariance	75
3.10	Conclusions	76
4	Discrete dynamics in the set of quantum measurements	77
4.1	Introduction	78
4.2	Discrete dynamics induced by blockwise stochastic matrices	79
4.2.1	Blockwise stochastic matrices and their blockwise product	79
4.2.2	Dynamics in the space of quantum measurements	81
4.2.3	Regions accessible by the blockwise product	82
4.3	Discrete dynamics induced by blockwise bistochastic matrices	84
4.3.1	Bistochastic dynamics and partial order between blockwise probability vectors	85
4.3.2	Majorization relation between quantum measurements	87
4.4	Resource theory of quantum measurements	90
4.5	Concluding remarks	91

4.6	Further results	92
4.6.1	Properties of the blockwise product and its dual version	92
4.6.2	Fixed points under blockwise bistochastic dynamics	95
4.6.3	Nonlinear monotone	97
5	Certifying nonlocal properties of noisy quantum operations	101
5.1	Introduction	101
5.2	Preliminaries	102
5.2.1	Quantum and classical states and channels	102
5.3	Channel nonlocality from local measurements	103
5.3.1	Local measurement protocols with shared entanglement	103
5.3.2	Local measurement protocols without entanglement	105
5.3.3	A Buscemi-like characterization of LOSR operations	105
5.4	The role of dephasing noise	109
5.4.1	Gradual dephasing	109
5.4.2	Complete decoherence	111
5.5	Conclusions and open questions	114
6	Nonlocality of decoherent quantum channels	116
6.1	Introduction	116
6.2	Quantum and classical nonlocality	117
6.2.1	Quantifying nonlocality of quantum channels	117
6.2.2	Quantifying nonlocal resources in the presence of decoherence	118
6.2.3	Quantum to classical measures: examples	120
6.3	Nonlocality resisting decoherence	121
6.3.1	A novel interpretation	121
6.3.2	Bounding the set of LOSR operations	121
6.4	The role of decoherent nonlocality in performing quantum protocols	122
6.4.1	Bell functionals on quantum channels	122
6.4.2	State interconversion	123
6.4.3	Quantum codes	124
6.5	Quantum simulation of stochastic matrices	124
6.6	Conclusions	125
7	Outlook and open problems	126
	Bibliography	129

Preface

Quantum information theory describes quantum physics in non-relativistic regimes, focusing on how information is stored and transferred in small-scale setups. In the last decades, this perspective provided a deep insight into quantum phenomena like entanglement, coherence or nonlocality. Moreover, its information-theoretic description makes quantum information science a suitable framework to develop the field of quantum computation and communication: in principle, protocols obeying quantum mechanics may allow for better performance than their classical counterparts in tasks such as factorizing large numbers [1] or distributing a cryptographic key [2].

Some structures in quantum information theory are based on the relationships between different convex sets. In a broader perspective, the idea of convexity appears in our understanding of nature in several regimes: even in simple experiments like flipping a coin or rolling a dice, our description of the system is given by a convex combination between all possible configurations. In discrete classical systems, which display a finite number of possible configurations, the state space forms a polytope; and the phase space description of statistical and thermodynamic systems accommodates quantities like temperature, volume or pressure, which allow for precise macroscopic predictions. This approach, where probability distributions describe our lack of knowledge, extends to our understanding of quantum mechanics and even more general theories [3]: the state space consists of a convex hull of its extremal points, known as pure states.

As a result, many problems in quantum information deal with delimiting a convex subset of interest. In a very broad perspective, these consist of either determining whether a given point is contained into a certain convex subset; or optimizing an objective function over such subset. Mathematically, these types of problems are often described by feasibility and optimization semidefinite programming problems respectively, since quantum mechanics setups can often be approximated with linear constraints. Physically, in feasibility problems one asks questions like whether a bipartite quantum state is entangled [4], prescribed marginals are compatible with a global state [5], two quantum measurements can be performed simultaneously [6], or a bipartite quantum channel can be performed locally [7]; and in optimization problems one can ask for the maximum fidelity of a state manipulation [8] or desired quantum operation [9], or minimizing the energy of a Hamiltonian with some additional constraints [10]. The specific examples listed above will capture most of the topics of the research work reported in this thesis.

Quantum information accommodates a richer phenomenology than classical information theory, such as superposition. While in principle this property might allow for quantum advantage in technologies, it comes along with an exponential growth of the state space with the system size. Therefore, access to target solutions often requires additional techniques to delimit desired corners in the state space. This thesis focuses in two aspects: First, we identify certain symmetries in a specific setup of study, in order to reduce redundancies and ease the access to the solution [11]. Second, we use a geometric approach to uncover certain structures and their connections [12]. These approaches will be used to tackle the following problems: entanglement detection exploiting symmetry under local unitaries [13] (Chapter 2); optimization problems symmetric under unitary tensor factors (Chapter 3); the dynamics of quantum measurements and characterization of relevant subsets [14] (Chapter 4); verification of nonlocality on quantum operations [15] (Chapter 5); and determining the role of nonlocal resources withstanding decoherence [16] (Chapter 6).

This thesis is oriented to both experts in quantum information theory, and readers outside of

the field who wish to learn the topics discussed. Specifically, the first part of each subsection in Basic concepts 1 will be dedicated to a non-specialized audience. The aim will be to introduce the research topic through a high-level, non-technical explanation providing a rather intuitive overview.

This dissertation is based on the following projects undertaken during my PhD:

- A Albert Rico and Felix Huber,
Entanglement detection with trace polynomials,
[Physical Review Letters](#), vol. **132**, p. 070202 (2024).
- B Albert Rico and Karol Życzkowski,
Discrete dynamics in the set of quantum measurements,
[arXiv:2308.05835](#)
- C Albert Rico, Moisés Bermejo Morán, Fereshte Shahbeigi, Karol Życzkowski,
Certifying nonlocal properties of noisy quantum operations,
[arXiv:2407.02570](#)
- D Albert Rico, Moisés Bermejo Morán, Fereshte Shahbeigi, Karol Życzkowski,
Channel nonlocality under decoherence,
[arXiv:2408.10317](#)
- E Albert Rico and Felix Huber,
Unitary invariant semidefinite programs: a users' guide,
In preparation

1

Basic concepts

“The beauty of nature lies not in its conquering, but in its harmony.”

– Tenzing Norgay

Quantum mechanics aims to describe physical systems at small scales. Although some of its effects are visible to the naked eye, like the sunlight decomposition through a prism, a full description of quantum phenomena is unreachable with the techniques used to describe classical mechanics [17, 18]. Instead, quantum mechanics is an effective mathematical theory that allows to make accurate predictions. Although different formalisms have been established [19, 20], here we will follow the *Copenhagen interpretation*, built upon the following postulates [21].

The first postulate deals with the description of isolated systems, namely with no interactions with the environment:

Postulate 1. *To any isolated physical system, a Hilbert space $\mathcal{H}$ is associated. The system is described by a unit vector $|\psi\rangle \in \mathcal{H}$.*

A Hilbert space is a complex vector space with inner product. Thus the first postulate describes isolated systems in a static way, and requires complex numbers for a complete description. Dynamical properties of isolated physical systems are described by the second postulate:

Postulate 2. *The temporal evolution of an isolated quantum system from an initial state $|\psi\rangle$ to a final state $|\psi'\rangle$ is described by a unitary transformation U , $|\psi'\rangle = U|\psi\rangle$.*

The third postulate describes how information can be obtained from a physical system.

Postulate 3. *Any measurable quantity has associated an Hermitian operator $O = O^\dagger$ acting on the state space $\mathcal{H}$, called observable. It has spectral decomposition $O = \sum_a a P_a$, where each eigenvalue a corresponds to a measurement outcome and each projector P_a projects onto the subspace $\mathcal{H}_a \subset \mathcal{H}$ with eigenvalue a . The probability $p(a)$ of each outcome a depends on the initial state of the system $|\psi\rangle \in \mathcal{H}$ as $p(a) = \langle\psi| P_a |\psi\rangle$ and the state after the measurement is $|\psi_a\rangle = \sqrt{p(a)^{-1}} P_a |\psi\rangle$.*

This procedure where the state of the system is projected onto a subspace is known as the *measurement collapse* [18].

If we take a closer look, quantum systems as described above can in turn be composed of smaller subsystems. This is described by the fourth postulate:

Postulate 4. *The state space of a composite physical system is the tensor product $\otimes$ of the state spaces of the component physical systems. The state of an isolated system composed of n subsystem reads $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_n\rangle$.*

Systems which are not isolated can be described by extending the Hilbert space to contain the environment they interact with. This allows to extend the four basic postulates above to the case of non-isolated systems, whose description is effectively given by a convex combination of possible configurations in an isolated system. The choice of the description employed is given in each case by the physical scenario to be described.

1.1 Fundamentals

1.1.1 Quantum states

What is a classical bit?

A *bit* is a binary unit of information which can take two values, 0 or 1. In classical computation, information is stored in tuples of bits called *bit strings*, for example 0010110. Therefore, a bit string of n bits can encode 2^n different values. In classical computation, bit strings are processed by logical gates implementing boolean functions. For example, the NOT-gate implements a bit flip, namely it outputs 1 if the input bit is 0, and vice-versa. Bits of information are transferred through channels, which can be imperfect [22]. That is, when aiming to send a bit with value 0, it might happen with a certain probability p that by accident a NOT-gate occurs before it arrives to the destination. Then the message is either 0 (with probability $1 - p$) or 1 (with probability p). In probability theory, our lack of knowledge about whether the bit is 0 or 1 can be considered as a probability vector $(1 - p, p)$.

And what is a quantum bit?

In quantum theory, a binary unit of information is described by a *qubit*. According to Postulate 1, it is a normalized vector in the Hilbert space and it is denoted as $|v\rangle \in \mathbb{C}^2$. This means that a qubit can take value $|0\rangle$ and $|1\rangle$, but also linear combinations such as

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \text{or} \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (1.1.1)$$

This phenomenon is known as *superposition*. States of the form of Eq. (1.1.1) have the property that when a measurement is done to determine the value 0 or 1 of the qubit, the result is either 0 or 1 with probability 0.5. Note that this property can also be achieved with a classical bit undergoing a NOT gate with probability $p = 0.5$ as discussed above, where the outcome is deterministic but our knowledge is that it is either 0 or 1 with probability 0.5. The difference is that quantum theory allows to measure whether a system is in the state $|+\rangle$ or $|-\rangle$, and obtain a deterministic outcome (with probability one), whereas measuring whether a state $|+\rangle$ is $|0\rangle$ or $|1\rangle$ leads to an intrinsically probabilistic outcome.

Pure quantum states

In quantum theory, any physical system is associated with a *Hilbert space* $\mathcal{H}$, which is a complex vector space with an scalar product $(\psi, \phi) := \langle \psi | \phi \rangle$ given two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, where $\langle \psi |$ is the transposed conjugated of $|\psi\rangle$. According to the first postulate of Quantum Mechanics, any isolated physical system can be fully described by a *pure quantum state*, which is a normalized vector of the corresponding Hilbert space,

$$|\psi\rangle \in \mathcal{H} : \quad \langle \psi | \psi \rangle = 1. \quad (1.1.2)$$

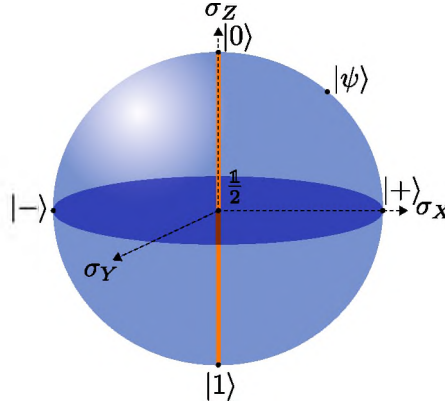


Figure 1.1.1: Bloch sphere representing the state space of one qubit. The surface of the sphere is composed of the vector space of pure states $|\psi\rangle \in \mathbb{C}^2$, spanned by the eigenvectors of the Pauli matrices σ_X , σ_Y and σ_Z . The convex hull of pure states is the space of mixed states ϱ and defines the interior of the sphere, whose center is the maximally mixed state $\mathbb{1}/2$. The three-dimensional Bloch sphere generalizes the one-dimensional space of probability distributions (vertical orange line) through $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| = \text{diag}(p, 1-p)$.

In analogy to the two-dimensional state space of qubits (Fig. 1.1.1), states of a Hilbert space $\mathcal{H}_d = \mathbb{C}^d$ with $d > 2$ are called *qudits*. A qudit $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{i=0}^{d-1} a_i |i\rangle, \quad (1.1.3)$$

where $\sum_{i=0}^{d-1} |a_i|^2 = 1$ and $\{|i\rangle\}$ denotes an orthonormal basis of $\mathcal{H}_d$ called *computational basis*. That is, each column-vector labelled by $i \in \mathbb{Z}_d$ reads

$$|i\rangle = (0, \dots, 0, 1, 0, \dots, 0)^T \quad (1.1.4)$$

where the entry 1 is located at the i 'th component and T denotes transposition.

Mixed quantum states

Systems that are not isolated are described by a statistical ensemble $\{p_i, |\psi_i\rangle\}$. This means that the system is deterministically in one of the states $|\psi_i\rangle$, but we don't know which one and thus we associate to each of them a probability p_i . In a more compact form, we describe these ensembles with a *density matrix* or *mixed quantum state*

$$\varrho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \in \mathcal{B}(\mathcal{H}), \quad (1.1.5)$$

where $\mathcal{B}(\mathcal{H})$ is the set of bounded operators acting on the Hilbert space $\mathcal{H}$ containing $|\psi_i\rangle$. A density matrix satisfies the following properties.

- (i) It is *self-adjoint*, namely $\varrho^\dagger = \varrho$ where $\dagger$ denotes the transpose conjugate, which means that its eigenvalues are real.
- (ii) It is *positive-semidefinite*, $\varrho \geq 0$, namely all its eigenvalues are nonnegative.
- (iii) It is normalized with respect to the trace, $\text{tr}(\varrho) = 1$.

One often writes a quantum state in its so-called *Bloch decomposition*, namely $\varrho = \mathbb{1}/d + \sum_{r=1}^{d^2-1} \alpha_r A_r$ where $\{\mathbb{1}, A_1, \dots, A_{d^2-1}\}$ form an orthonormal basis for $\mathcal{B}(\mathcal{H})$ and $\alpha_r = \text{tr}(\varrho A_r)$ are known as

the *Bloch coefficients*. The diagonal elements ϱ_{ii} of a density matrix, known as *populations* or *occupations*, denote the probability that the system is at the state $|i\rangle$.

Composite systems

Consider a system composed by two subsystems A and B which has associated a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, with local dimensions d_A and d_B . In general, the state of this system is described by a density operator $\varrho_{AB} \in \mathcal{B}(\mathcal{H})$. The state describing the subsystem A , ϱ_A , is given by the partial trace of ϱ_{AB} with respect to the subsystem B ,

$$\varrho_A = \text{tr}_B(\varrho_{AB}) = \sum_{i=0}^{d_B-1} (\mathbb{1} \otimes \langle i|_B) \varrho_{AB} (\mathbb{1} \otimes |i\rangle_B). \quad (1.1.6)$$

The density operator ϱ_A is denoted as the *reduced density operator* corresponding to the subsystem A . An analogous description defines the reduced density operator ϱ_B .

Pure states $|\psi\rangle_{AB}$ in a composite Hilbert space can be decomposed in orthonormal bases $\{|u_i\rangle\}$ and $\{|v_j\rangle\}$ as $|\psi\rangle_{AB} = \sum_i \psi_{ij} |u_i\rangle_A \otimes |v_j\rangle_B$, and often one adopts the shorthand notations $|u_i\rangle |v_j\rangle := |u_i\rangle \otimes |v_j\rangle$ in general and $|ij\rangle := |i\rangle \otimes |j\rangle$ for the computational basis. Physically, the partial trace determines our description of the system, when a part of it is ignored. In particular, any quantum state $\varrho_A \in \mathcal{B}(\mathcal{H}_A)$ can be obtained by the partial trace of a bipartite pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$, via $\varrho_A = \text{tr}_B(|\psi\rangle\langle\psi|)$. The inverse process of constructing a bipartite $|\psi\rangle_{AB}$ whose partial trace is ϱ_A is known as *purification*.

1.1.2 Quantum operations and measurements

What is a measurement in quantum mechanics?

A quantum operation is a process that a quantum system can undergo. In particular, a measurement is an operation whose outcome determines some property of the system [23, 24]. In classical physics, the result of a measurement is assumed to be inherently deterministic, despite the fact that we associate probability distributions with measurement outcomes that are unknown to us (e.g. tossing a coin). Moreover, a measurement in classical physics is assumed to reveal the state of the system without interfering with it (e.g. measuring a length with a ruler, whose outcome in centimeters tells us the length of the system both before and after we measure it). These features are different in quantum theory. The notion of a quantum measurement admits situations in which a given measurement outcome does not describe what the state of the system was *before* such measurement was performed. In this case, each measurement outcome occurs with an intrinsically random probability and determines what is the state *after* the measurement is performed. In the Copenhagen interpretation, this is explained by assuming that the state of the system changes during the measurement process (Postulate 3). This is known as the *collapse* [18] of a quantum state and it is not fully understood within the framework of quantum mechanics itself [23].

Quantum gates

In analogy to classical bits evolving by logical gates, qudits evolve by unitary operations U satisfying $U^\dagger U = U U^\dagger = \mathbb{1}$, also known as quantum gates. Unitary operations preserve the rank of the input states and thus send pure quantum states to pure quantum states, in the form of

$$U |\psi\rangle = |\phi\rangle. \quad (1.1.7)$$

All pure quantum states in a d -dimensional Hilbert space, $|\psi\rangle \in \mathbb{C}^d$, can be obtained with some unitary matrix U from a fixed initial state $|\phi\rangle$ via Eq. (1.1.7). A basis for single-qubit gates is given

by the identity operator $\mathbb{1}$ and the Pauli matrices,

$$\sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (1.1.8)$$

where i is the imaginary unit. Well-known examples of unitary operations are the SWAP and Control-NOT gates,

$$U_{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad U_{C-NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (1.1.9)$$

CPTP maps

A *quantum operation* is the most general process a quantum state $\varrho \in \mathcal{B}(\mathbb{C}^d)$ can undergo. Thus any physical evolution of a system is determined by a quantum operation acting on its state. Mathematically, a quantum operation $\mathcal{E}$ is a linear map

$$\mathcal{E} : \mathcal{B}(\mathbb{C}^{d_0}) \rightarrow \mathcal{B}(\mathbb{C}^{d_1}) \quad (1.1.10)$$

which is *completely positive* (CP) and trace non-increasing, where d_0 and d_1 are the dimensions of the input and output spaces. Complete positivity means that the tensor product of the map $\mathcal{E}$ with the identity satisfies $\mathcal{E}_A \otimes \text{id}_B(X_{AB}) \geq 0$ for any positive semidefinite operator $X_{AB} \geq 0$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ and for any Hilbert space $\mathcal{H}_B$. The trace non-increasing condition means that $\text{tr}(\mathcal{E}(X)) \leq \text{tr}(X)$ for any positive semidefinite operator X . Physically, this requirement is because the sum of the populations described by a density matrix cannot be larger than one. In particular, completely positive and *trace preserving* (TP) maps (CPTP) send normalized density matrices to normalized density matrices, and are also called *quantum channels*.

Dilation of quantum operations

By the so-called Stinespring dilation [25], any quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_0) \rightarrow \mathcal{B}(\mathcal{H}_1)$ can be written as $\mathcal{E}(\varrho) = \text{tr}_e(V\varrho V^\dagger)$ where $V : \mathcal{H}_0 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_e$ is an isometry, namely $V^\dagger V = \mathbb{1}_{d_0}$, and d_0 is the dimension of $\mathcal{H}_0$. The extended subsystem with associated Hilbert space $\mathcal{H}_e$ is called *environment*. In particular, one can find a unitary matrix $U = V \otimes \sum_i |i\rangle : \mathcal{H}_0 \otimes \mathcal{H}_{e'} \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_e$ with $UU^\dagger = U^\dagger U = \mathbb{1}$ and $d_0 d_{e'} = d_1 d_e$ such that

$$\mathcal{E}(\varrho) = \text{tr}_e(U(\varrho_0 \otimes |\psi\rangle\langle\psi|_{e'})U^\dagger) \quad (1.1.11)$$

for any ancillary state $|\psi\rangle \in \mathcal{H}_{e'}$ in an environment. Physically, this procedure gives a recipe to implement a quantum channel $\mathcal{E}$ experimentally with a unitary gate and an ancillary system prepared in an arbitrary state.

Operator-sum representation

A quantum operation $\mathcal{E}$ acting on a quantum state ϱ can be described as [21]

$$\mathcal{E}(\varrho) = \sum_i K_i \varrho K_i^\dagger, \quad (1.1.12)$$

where K_i are called *Kraus operators* and satisfy $\mathbb{1} - \sum_i K_i^\dagger K_i \geq 0$. This condition accounts for the map $\mathcal{E}$ to be trace non-increasing, and it saturates $\sum_i K_i^\dagger K_i = \mathbb{1}$ for CPTP maps (or quantum channels). For example, unitary channels are defined by a single Kraus operator U . Quantum

operations which preserve the identity, $\mathcal{E}(\mathbb{1}) = \mathbb{1}$, are called *unital* and satisfy that $\sum_i K_i K_i^\dagger = \mathbb{1}$. The Kraus operators are not unique: given two sets of Kraus operators $\{K_i\}$ and $\{L_j\}$, one has $\sum_i K_i \varrho K_i^\dagger = \sum_j L_j \varrho L_j^\dagger$ for any quantum state ϱ when $L_j = \sum_i u_{ij} K_i$ for some isometry with entries u_{ij} , namely $\sum_i u_{ij}^* u_{il} = \delta_{il}$. Nonetheless, for any choice of the Kraus operators $\{K_i\}$, one can write the dilation of the quantum channel $\mathcal{E}$ using that $V = \sum_i K_i \otimes |i\rangle$.

Choi-Jamiołkowski isomorphism

An alternative representation of quantum operations relies on the Choi-Jamiołkowski isomorphism: A quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_0) \rightarrow \mathcal{B}(\mathcal{H}_1)$ acts on an input quantum state ϱ_0 via [26, 27]

$$\mathcal{E}(\varrho_0) = \text{tr}_0(J_{01} \cdot (\varrho_0^T \otimes \mathbb{1}_1)), \quad (1.1.13)$$

where T is the matrix transposition, 0 and 1 denote the input and output systems, and $J_{01} \in \mathcal{B}(\mathcal{H}_0 \otimes \mathcal{H}_1)$ is the *Choi matrix*. From this point, and for optimal notation when considering channels acting on composite systems, say A and B , we shall adopt the convention $A := A_0 A_1$ and $B := B_0 B_1$ following [28]. That is, we will denote a single-partite channel by $\mathcal{E}_A$ with Choi matrix $J_{A_0 A_1}$, and a bipartite channel by $\mathcal{E}_{AB}$ with Choi matrix $J_{A_0 A_1 B_0 B_1}$. A map $\mathcal{E}_A$ is completely positive if and only if the Choi matrix is positive semidefinite, $J_{A_0 A_1} \geq 0$; it is trace non-increasing if and only if $\mathbb{1}_{A_0} - \text{tr}_{A_1} J_{A_0 A_1} \geq 0$; and $\mathcal{E}$ is trace preserving if and only if $\text{tr}_{A_1} J_{A_0 A_1} = \mathbb{1}_{A_0}$ [12, 26, 27]. The Choi matrix can be obtained as [12]

$$J_{A_0 A_1} = d_{A_0} \text{id} \otimes \mathcal{E}(|\phi^+\rangle\langle\phi^+|), \quad (1.1.14)$$

where $|\phi^+\rangle = \sum_{i=0}^{d_{A_0}-1} |ii\rangle / \sqrt{d_{A_0}}$ is the Bell state. The Choi matrix can be written from the Kraus operators K_i as

$$(J_{A_0 A_1})^R = \sum_i K_i \otimes K_i^*, \quad (1.1.15)$$

where $*$ denotes the conjugate transpose and the reshuffle R acts as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^R = |i\rangle\langle k| \otimes |j\rangle\langle l|$.

Projective measurements

Here we will explain in further detail Postulate 3, which formalizes the phenomenon of the measurement collapse by introducing the *projective* or *von Neumann* measurement [29]. In this framework, every physical measurable quantity is associated to a Hermitian operator $O = O^\dagger \in \mathcal{B}(\mathcal{H})$, called *observable*, with spectral decomposition

$$O = \sum_a a P_a. \quad (1.1.16)$$

Here P_a are orthogonal projectors resolving the identity, namely satisfying $P_a \geq 0$, $P_a^2 = P_a$ and $\sum_a P_a = \mathbb{1}$. The (real) eigenvalues a of O represent the physical outcomes of the measurement described by O , and they occur with probability $p(a) = \langle\psi| P_a |\psi\rangle$ where $|\psi\rangle$ is the state of the system *before* the measurement was performed. This probabilistic behaviour means that the state $|\psi\rangle$ cannot be revealed by a measurement. Only the probabilities $p(a)$ can be approximated from the relative frequency of occurrence of each outcome a , after a large number of measurements is performed. Moreover, note that different possible initial states $|\psi\rangle$ and $|\psi'\rangle$ might have the same inner product with P_a . Therefore, even if the probabilities $p(a)$ are determined with high accuracy, the state of the system prior to measurement cannot be revealed in general. An exception occurs if one has $p(a) = 1$ for some outcome a , provided that the corresponding subspace is not degenerated.

To describe the state of the system after the measurement is performed, one considers that each projector P_a projects the state $|\psi\rangle$ onto the eigenspace of O corresponding to the outcome a , such

that the state after the measurement reads

$$|\phi_a\rangle = \frac{P_a |\psi\rangle}{\sqrt{p(a)}}. \quad (1.1.17)$$

Due to the probabilistic nature of the quantum measurement, the measurable quantities are in practice *expectation values* $\langle O \rangle_\psi$. These are probabilistic averages of the outcomes a and depend on the state of the system $|\psi\rangle$ before the measurement was performed, as

$$\langle O \rangle_\psi = \langle \psi | O | \psi \rangle = \sum_a p(a) a. \quad (1.1.18)$$

The action of projective measurements on pure states can be extended to mixed states similarly, namely the state after a measurement outcome a occurs reads $\sigma_a = P_a \varrho P_a / p(a)$ since the initial state ϱ is a convex combination of pure states. Then the expectation value is the inner product $\langle O \rangle = \text{tr}(O \varrho)$ where ϱ is the density matrix describing the system before the measurement was performed.

Generalized measurements

Mathematically, one can describe a quantum measurement as a process by which one obtains a set of outcomes a with associated probabilities $p(a)$, which depend linearly on the state of the system prior to the measurement. In full generality, this is described by a set of positive semidefinite operators $M_a = M_a^\dagger \geq 0$ called *measurement effects* summing to the identity, $\sum_a M_a = \mathbb{1}$, from which the probabilities are obtained through

$$p(a) = \text{tr}(M_a \varrho), \quad (1.1.19)$$

where ϱ is the state of the system before the measurement was performed. This is known as a *generalized measurement* or *Positive Operator-Valued measurement (POVM)* [21] and recovers the projective case when M_i are projectors. In the event of a measurement outcome a which occurs with probability $p(a)$, the state after the measurement reads

$$\sigma_a = \frac{E_a \varrho E_a^\dagger}{p(a)}. \quad (1.1.20)$$

The post-measurement state is statistically described as

$$\sigma = \sum_a p(a) \sigma_a = \sum_a E_a \varrho E_a^\dagger, \quad (1.1.21)$$

where we identify the operators E_a as Kraus operators. That is, a generalized quantum measurement can be performed by a quantum channel with Kraus operators E_a satisfying $E_a^\dagger E_a = M_a$, which is the condition for the outcome probability distribution to be given by $p(a)$. Physically, this means that any POVM can be performed by the dilation of the quantum channel through which it is performed, making use of Eq. (1.1.11).

1.1.3 Quantum entanglement

What is quantum entanglement?

In quantum theory, two systems are *entangled* if they do not admit a (deterministic or statistical) independent description of each of them. As an example, suppose that Alice and Bob are spatially separated, and each of them holds a qubit system A and B . In 1935, A. Einstein, B. Podolsky, and N. Rosen realized that according to quantum mechanics, the joint state describing their qubits can

be in a superposition of the form [30]

$$|\psi^+\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_A |1\rangle_B + |1\rangle_A |0\rangle_B \right). \quad (1.1.22)$$

This state is incompatible with a description of the system where the state in Alice's side is $|\psi\rangle_A$ and the state in Bob's side is $|\psi\rangle_B$. Instead, the state of Alice being $|0\rangle$ is linked to the state of Bob being $|1\rangle$, and vice-versa.

This should not be confused with a classical system admitting two possibilities: either the qubit of Alice is in the state $|0\rangle$ and the qubit of Bob is in the state $|1\rangle$, or vice-versa, each of them with a certain probability. This system displays so-called *classical correlations*, and is described by a statistical ensemble

$$\varrho = \frac{1}{2} \left(|0\rangle\langle 0|_A \otimes |1\rangle\langle 1|_B + |1\rangle\langle 1|_A \otimes |0\rangle\langle 0|_B \right). \quad (1.1.23)$$

That is, the system admits the local description $|0\rangle\langle 0|_A \otimes |1\rangle\langle 1|_B$ with probability $1/2$, and the local description $|1\rangle\langle 1|_A \otimes |0\rangle\langle 0|_B$ with probability $1/2$.

The difference between entanglement and classical correlations is that in classical correlations, the state of each independent subsystem has a deterministic description, but we do not have access to it and we describe it as probabilistic and conditional. However, two systems being entangled means that an independent description of each cannot be done, even in probabilistic terms.

Bipartite entanglement

A bipartite quantum state $\varrho_{AB} \in \mathcal{B}(\mathcal{H}_{AB})$ is called *separable* if it admits a decomposition as a convex combination of product states,

$$\varrho_{AB} = \sum_i p_i \varrho_A^i \otimes \varrho_B^i, \quad (1.1.24)$$

where $p_i \geq 0$, $\sum_i p_i = 1$ and $\varrho_A^i, \varrho_B^i \geq 0$, and *entangled* otherwise. In particular, a state ϱ_{AB} is a *product state* if it satisfies Eq. (1.1.24) with a single term, namely $\varrho_{AB} = \varrho_A \otimes \varrho_B$. Since mixed states form the convex hull of product states, a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is separable if and only if it is a product state,

$$|\psi\rangle = |\phi\rangle_A \otimes |\varphi\rangle_B, \quad (1.1.25)$$

where $|\phi\rangle_A \in \mathcal{H}_A$ and $|\varphi\rangle_B \in \mathcal{H}_B$ are pure states. Entanglement is a resource whose free operations are *local operations assisted by classical communication (LOCC)*, which are physically motivated by setups where parties are spatially separated but can share some information about the quantum operations they perform [31]. Whereas quantifying the resource of entanglement requires families of measures for mixed states [32], the case of bipartite pure states is simpler: the amount of entanglement corresponds to the *von Neumann entropy* of the reduced state ϱ_A , $S(\varrho_A) = -\sum_i \lambda_i \log \lambda_i$ where λ_i are the eigenvalues of ϱ_A . For two qubits, the *Bell states*

$$|\phi^\pm\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}} \quad \text{and} \quad |\psi^\pm\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}} \quad (1.1.26)$$

form an orthonormal basis of maximally entangled states, namely that maximize the von Neumann entropy of the reductions, since their marginals are maximally mixed $\varrho_A = \varrho_B = \mathbb{1}/2$.

Multipartite entanglement

Consider a composite system describing n parties, associated to a Hilbert space $\mathcal{H} = \mathbb{C}^{d_1} \otimes \dots \otimes \mathbb{C}^{d_n}$. In analogy to bipartite separability, we say that a quantum state $\varrho \in \mathcal{B}(\mathcal{H})$ is *fully separable* if

$$\varrho = \sum_i p_i \varrho_1^i \otimes \dots \otimes \varrho_n^i. \quad (1.1.27)$$

However, the notion of separability for multipartite systems is more involved. For example, a quantum state can be *biseparable*, namely it satisfies Eq. (1.1.24) for a given bipartition where A contains the first $k < n$ parties and B the remaining $n - k$ parties. Moreover, quantifying entanglement in multipartite systems is much more involved than in the bipartite case, even for pure states [32]. For three qubits, the quantum states

$$|\text{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}} \quad \text{and} \quad |\text{W}\rangle = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}} \quad (1.1.28)$$

cannot be transformed from one another with LOCC and thus different entanglement measures are needed to quantify their entanglement [33].

1.1.4 Entanglement detection

Can entanglement be detected?

The phenomenon of entanglement lies at the core of understanding both the nature of quantum mechanics and its applications [2, 8, 34–38]. However, it is challenging to know whether or not some entanglement is present in a given system [4]. Given a full theoretical description of the state of the system, deciding whether or not a quantum state is entangled is an NP-hard problem [39, 40]. Detecting entanglement is even more challenging in practice, since in experimental setups one does not have access to a complete description of a system [21]. Instead, properties of a quantum state can be partially assessed by measuring an observable, which requires performing multiple measurements and analyzing the relative frequency of their outcomes [23]. For the purpose of detecting whether or not entanglement is present in a system, a common tool are *entanglement witnesses*. These are multipartite observables W whose expectation value $\langle W \rangle$ gives us information about the entanglement of a state: if the expectation value is negative, $\langle W \rangle < 0$, then the state describing the system is entangled.

Positive maps

A main tool to detect entanglement in a quantum state consists of *positive maps* $\mathcal{E}$ which are not completely positive. These satisfy $\mathcal{E}(X) \geq 0$ for all positive semidefinite matrices $X \geq 0$, but $\mathcal{E}_A \otimes \text{id}_B(X_{AB}) \not\geq 0$ for some positive semidefinite matrix $X_{AB} \in \mathcal{B}(\mathcal{H})$ acting on a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. By convexity, a positive map satisfies $\mathcal{E}_A \otimes \text{id}_B(\varrho_{AB}^{\text{sep}}) \geq 0$ for any separable state $\varrho_{AB}^{\text{sep}} = \sum_i p_i \varrho_A^i \otimes \varrho_B^i$. In this sense, a positive map $\mathcal{E}$ *detects* a bipartite state ϱ_{AB} if the output of its action on one subsystem has a negative eigenvalue, $\mathcal{E}_A \otimes \text{id}_B(\varrho_{AB}) \not\geq 0$. It was shown in [41] that a bipartite state ϱ_{AB} is entangled if and only if there exists a positive map detecting it. Paradigmatic examples of positive maps used to detect entanglement are the following:

- *Partial transpose* (T_B): The partial transpose on a subsystem B of a bipartite operator $X_{AB} = \sum_{ijkl} X_{ijkl} |ik\rangle \langle jl|$ is the transposition only on one subsystem, $X_{AB}^{T_B} = \sum_{ijkl} X_{ijkl} |il\rangle \langle jk|$, where we use the shorthand notation $|rs\rangle := |r\rangle \otimes |s\rangle$. The positive partial transpose (PPT)-criterion is simple to evaluate in density matrices, can be approximated experimentally through matrix moments [42, 43] and it certifies undistillability of states [44]. These features make it the main tool to detect entanglement in various setups.

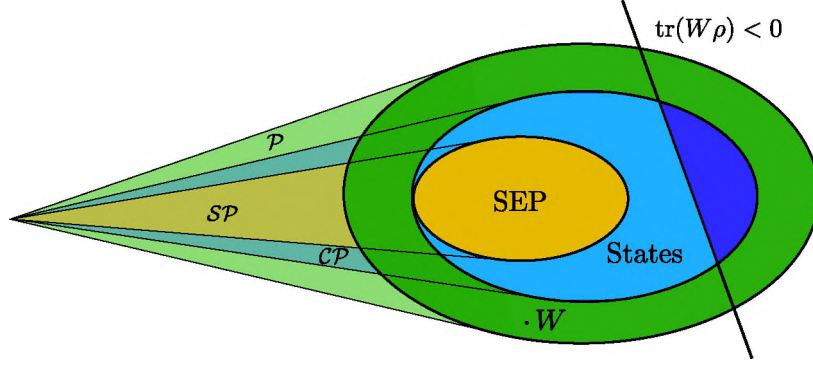


Figure 1.1.2: Entanglement detection with a witness W . The set of block-positive bipartite operators with unit trace contains the set of bipartite states, which in turn contains the set of separable (SEP) operators depicted in orange. Block-positive operators W which are not positive (green set) act as witnesses for entangled states (blue set): each point W defines a hyperplane that separates entangled states ρ that are detected through $\text{tr}(W\rho) < 0$ (dark blue set) from the rest. The set of separable states and the set of normalized block-positive operators are slices of convex cones ($\mathcal{SP}$ and $\mathcal{CP}$) that are dual to each other, in the sense that a state ρ is separable if and only if for all witnesses W one has $\text{tr}(\rho W) \geq 0$; and a non-positive Hermitian operator W is a witness if and only if for all separable states ρ it has a nonnegative expectation value $\text{tr}(W\rho) \geq 0$. Similarly, the set of states is a slice of a self-dual cone ($\mathcal{P}$).

- *Reduction map:* The reduction map is defined as $R(X) = \text{tr}(X)\mathbb{1} - X$ [45]. It can detect bipartite entanglement on a state ρ_{AB} when acting on a subsystem, $R \otimes \text{id}(\rho_{AB})$. The reduction map can only detect states which are also detected by the PPT-criterion, but it is also necessary and sufficient to detect entanglement in two qubits [45].
- *Breuer-Hall map:* Detecting PPT-entangled states is significantly more challenging than detecting non-PPT (NPT) entangled states. For this problem, one can generalize the reduction map to the Breuer-Hall map, $\Phi(X) = \text{tr}(X)\mathbb{1} - X - UXU^\dagger$ where $U = -U^T$ is an anti-unitary matrix [46, 47]. The Breuer-Hall map can indeed detect PPT-entangled states, namely there exist entangled states ρ_{AB} with $\rho_{AB}^{T_A} \geq 0$ but $\Phi \otimes \text{id}(\rho_{AB}) \not\geq 0$.

Entanglement witnesses

An *entanglement witness* is a Hermitian operator W which is *block-positive*, namely it satisfies $\text{tr}(\rho^{\text{sep}} W) \geq 0$ for any separable state ρ^{sep} , and moreover it has a negative eigenvalue, $W \not\geq 0$. Thus detection works as follows: if a witnesses W has a negative expectation value on a state ρ , $\text{tr}(W\rho) < 0$, then ρ is entangled. In this way, every nonseparable state (namely storing bipartite or multipartite entanglement) can be detected by some witness [48]. This is because evaluating an entanglement witness is equivalent to evaluating a positive map, on a positive operator [12]. Indeed, the Choi matrix of a positive (but not completely positive) map $\mathcal{E}$ is a block-positive (but not positive) matrix $W^\mathcal{E}$. Then by the Choi-Jamiołkowski isomorphism one has $\text{tr}(W^\mathcal{E}\rho) = \text{tr}(\text{id} \otimes \mathcal{E}(\phi^+))\rho = \text{tr}(\text{id} \otimes \mathcal{E}^*(\rho))\phi^+$ where $\mathcal{E}^*$ is the adjoint of $\mathcal{E}$. Using that an operator X is positive semidefinite if and only if $\text{tr}(XY)$ is nonnegative for any positive semidefinite operator Y , one has that if $\text{tr}(W^\mathcal{E}\rho) < 0$ then $\mathcal{E} \otimes \text{id}(\rho) \not\geq 0$. That is, a Witness $W^\mathcal{E}$ defining a map $\mathcal{E}$ through the Choi-Jamiołkowski isomorphism evaluates positivity of $\mathcal{E} \otimes \text{id}(\rho)$ through the inner product $\text{tr}(\mathcal{E} \otimes \text{id}(\rho)\phi^+)$. Conversely, if a state ρ is detected by a map $\mathcal{E}$ through $\mathcal{E} \otimes \text{id}(\rho) = \sum_i \lambda_i |\lambda_i\rangle \langle \lambda_i|$ with some negative eigenvalue $\lambda_j < 0$, then it is also detected by a witness defined as $W^\mathcal{E} = \text{id} \otimes \mathcal{E}^*(|\lambda_j\rangle \langle \lambda_j|)$.

Consider the set of block positive operators, the set of positive operators, and the convex hull of tensor products of positive operators, which we denote as $\mathcal{CP}$, $\mathcal{P}$ and $\mathcal{SP}$ following [12]. These three sets form three convex cones containing each other [12], $\mathcal{SP} \subset \mathcal{P} \subset \mathcal{CP}$ (Fig. 1.1.2). Moreover, the

cone $\mathcal{SP}$ is dual to the cone $\mathcal{CP}$: the inner product with block positive operators forms the set of most general linear maps that take nonnegative values on convex combinations of tensor products of positive operators. In this sense the set of positive operators $\mathcal{P}$ is self-dual, namely dual to itself. Physically, one considers the slice of the cones described above depicted in Fig. 1.1.2, given by the unit-trace normalization constraint on states, $\text{tr}(\varrho) = 1$. This defines a hyperplane that intersects also the cone of block-positive operators: one often considers witnesses with unit trace, $\text{tr}(W) = 1$ [49]. Then the dualities between convex cones described above translate to the property that a state ϱ is separable, if and only if it has a nonnegative expectation value $\langle W \rangle_\varrho := \text{tr}(W\varrho) \geq 0$ for all witnesses W [48].

Experimentally, entanglement witnesses are observable operators, and therefore accessible even if the state to be detected is unknown. In contrast, positive maps require some knowledge about the state through full or partial tomography. This makes entanglement witnesses a widely used tool for the detection of entanglement in realistic setups. An example is the swap operator, $V = \sum_{ij} |ij\rangle\langle ji| = |\phi^+\rangle\langle\phi^+|^{T_B}$. Entanglement witnesses can also be constructed from Bell inequalities [50], using the fact that all nonlocal states are entangled.

1.1.5 Nonlocality

Local hidden variables

The notion of entanglement motivated the following question: Can quantum theory be reformulated, so that it admits a local description of reality? This raised the hypothesis that quantum systems might not be intrinsically probabilistic, but described by some deterministic –yet unknown– variables called *local hidden variables*. This hypothesis was ruled out by John Bell in 1964 [51], who showed that quantum mechanics admits correlations between local measurements which cannot be explained with a local hidden variable model. In the framework of modern quantum information, the setting consists of two parties, Alice and Bob, who share a quantum state ϱ_{AB} and perform local measurements. Alice is allowed to choose a measurement M_x with outcome effects $M_{a|x}$, and Bob is allowed to perform a measurement N_y with outcome effects $N_{b|y}$. As an example, consider the value

$$\gamma = \langle M_0 \otimes N_0 \rangle + \langle M_0 \otimes N_1 \rangle + \langle M_1 \otimes N_0 \rangle - \langle M_1 \otimes N_1 \rangle. \quad (1.1.29)$$

with the measurements $M_0 = \sigma_X$, $M_1 = \sigma_Z$, $N_0 = (\sigma_X + \sigma_Z)/\sqrt{2}$ and $N_1 = (\sigma_X - \sigma_Z)/\sqrt{2}$. One can show that if the state ϱ_{AB} can be described by local hidden variables, then $\gamma \leq 2$. However, for the state $|\psi^+\rangle$ of Eq. (1.1.22), one has $\gamma = 2\sqrt{2}$, which shows that this state cannot be described with local hidden variables. This is known as the *CHSH game* and it was the first example of more general measurement setups known as *nonlocal games* [52].

Nonlocality of probability distributions

A bipartite probability distribution describes *independent* events when it is a product of local distributions, $p(a, b|x, y) = p(a|x)p(b|y)$. The convex hull of these forms the *Bell local* distributions (L) [52, 53]. These decompose as

$$p(a, b|x, y) = \sum_{\lambda} p_{\lambda} p(a|x, \lambda) p(b|y, \lambda), \quad (1.1.30)$$

where $p_{\lambda} \geq 0$, $\sum_{\lambda} p_{\lambda} = 1$ in the discrete case, and extend to the case of a continuous distribution p_{λ} with $\int_{\lambda} p_{\lambda} d\lambda = 1$ by replacing the summation by an integral. Distributions that cannot be decomposed in this way are called *Bell nonlocal*. Here λ accounts for hidden variables which determine the correlations of the bipartite system locally. In quantum theory, probabilities $p(a, b|x, y)$ are obtained through Born's rule,

$$p(a, b|x, y) = \text{tr}(M^{a|x} \otimes N^{b|y} \varrho_{AB}) \quad (1.1.31)$$

for some shared state ϱ_{AB} and measurement effects $\{M^{a|x}\}$ and $\{N^{b|y}\}$. Distributions obtained from Eq. (1.1.31) are called *quantum* (Q). A probability distribution is called *nonsignaling* (NS) if it has normalized marginals, i.e., it satisfies

$$p(a|x) = \sum_b p(a, b|x, y) = \sum_b p(a, b|x, y'), \quad (1.1.32a)$$

$$p(b|y) = \sum_a p(a, b|x, y) = \sum_a p(a, b|x', y). \quad (1.1.32b)$$

The sets L, Q and NS are convex and form a chain of inclusions

$$L \subset Q \subset NS. \quad (1.1.33)$$

Thus, they can be separated with linear functionals

$$\gamma(p) := \sum_{xyab} \gamma_{ab,xy} p(ab|xy) \quad (1.1.34)$$

on the space of bipartite probability distributions, called *Bell functionals*. One denotes $\gamma_S := \max_{p \in \mathcal{L}} \gamma(p)$ the maximum value of a Bell functional γ over the set of distributions inside a selected subset S (typically L, Q or NS). For example, the CHSH functional $\gamma_{ab,xy} = (-1)^{a+b}(-1)^{xy}$ satisfies $\gamma_L = 2$, $\gamma_Q = 2\sqrt{2}$ and $\gamma_{NS} = 4$.

Nonlocality of quantum states

A *Bell protocol* is a device independent local measurement protocol where a bipartite probability distribution is obtained from a bipartite quantum state through Eq. (1.1.31). Assuming quantum theory, the only subsets of bipartite distributions that can be obtained in this way are local (L) or quantum (Q). Separable quantum states can only produce local correlations, and therefore observing nonlocal quantum correlations is an indicator of the presence of entanglement [4, 50]. A quantum state ϱ_{AB} which cannot produce nonlocal correlations in the bipartition $A|B$ will be called a *local state* in such bipartition.

Remarkably, not all entangled states can produce correlations outside of L through Eq. (1.1.31), namely there exist local but entangled states [54, 55]. However, a subtle difference exists between separable and entangled states in certain measurement protocols: it was shown by Buscemi that all entangled states ϱ_{AB} produce correlations that cannot be obtained from separable states in a semi-device independent protocol with trusted quantum inputs [56],

$$p(ab|xy) = \text{tr}(\tau_R^x \otimes \varrho_{AB} \otimes \omega_S^y \cdot M_{AR}^a \otimes N_{BS}^b). \quad (1.1.35)$$

Here the inputs are fixed sets of quantum states $\{\tau^x\}$ and $\{\omega^y\}$ used in ancillary systems R and S . Distributions that cannot be obtained in this way from separable states are known as *Buscemi nonlocal*. For example, consider a state that is entangled but cannot produce distributions out of L through Eq. (1.1.31) [54]. Through Eq. (1.1.35), such a state still produces distributions that are inside the local polytope L, but cannot be obtained by any separable state and thus are Buscemi nonlocal. Quantum measurement protocols obtained through Eq. (1.1.35) are called *Buscemi protocols* to distinguish them from *Bell protocols* in Eq. (1.1.31).

1.1.6 Decoherence

Quantum phenomena occur at small scales

Quantum mechanics exhibits properties that occur only in a small scale, namely when the system in consideration is composed of a limited number of particles. Instead, observation indicates that

these properties are typically lost when considering larger systems, whose description requires a classical description that emerges as an average of quantum properties. Here we will explain this effect through a single qubit example, and introduce the phenomenon responsible for this. Let us start from Postulate 1, according to which an isolated quantum system L is described by a pure state $|\psi\rangle$. Without loss of generality, for a single qubit one can write

$$|\psi\rangle = \cos\theta |0\rangle + \sin\theta e^{i\varphi} |1\rangle \quad (1.1.36)$$

where $\theta, \varphi \in (0, 2\pi]$. However, in practice physical systems interact with their environment E . This interaction results into a joint entangled state $|\psi\rangle_{LE} = U(|\psi\rangle\langle\psi|_L \otimes |e\rangle\langle e|_E)U^\dagger$ and thus the (reduced) state of the system L needs to be described by a density matrix ϱ . The loss of purity from $|\psi\rangle$ to ϱ is known as *decoherence*.

A well established model of decoherence is the *dephasing model* [57, 58], motivated as follows. In order to learn the properties of the state $|\psi\rangle$ describing our system L , one needs to measure in a fixed basis, say the computational basis $\{|0\rangle, |1\rangle\}$. However, the basis choice limits our knowledge: the parameter θ gives the probabilities of the outcomes 0 and 1 through $p(0) = \cos^2\theta$ and $p(1) = \sin^2\theta$, but the phase $e^{i\varphi}$ cannot be revealed. The dephasing model of decoherence aims to describe the situation in which the interaction with the environment modifies the phase $e^{i\varphi}$ —which cannot be estimated in a computational basis measurement—while the parameter θ —which can be estimated in a computational basis measurement—remains untouched. This results in a state $|\psi'\rangle = \cos\theta |0\rangle + \sin\theta e^{i\varphi'} |1\rangle \neq |\psi\rangle$ that “seems” to be the quantum state initially prepared $|\psi\rangle$ to the measurement protocol, but is not.

Consider the dephasing noise occurring multiple times m . This describes a large number m of subsystems L_j , for example multiple particles. One considers that each of them undergoes a random change of the angle $\varphi \rightarrow \varphi_j$, obtaining a new state $|\psi_j\rangle$ with probability q_j . Statistically, the transformed state is described by a probabilistic ensemble

$$|\psi\rangle\langle\psi| \rightarrow \varrho = \sum_{j=1}^m q_j |\psi_j\rangle\langle\psi_j| = \begin{pmatrix} \cos^2\theta & \varrho_{12} \\ \varrho_{12}^* & \sin^2\theta \end{pmatrix} \quad (1.1.37)$$

with $q_j \geq 0$, $\sum_{j=1}^m q_j = 1$ and $\varrho_{12} = \cos\theta \sin\theta \sum_{j=1}^m q_j e^{-i\varphi_j}$. Indeed, here the measurement probabilities remain untouched but the *coherence* ϱ_{12} carries an average over m different possible phases. In the complex plane it is clear that $|\sum_{\varphi'=1}^m q_{\varphi'} e^{-i\varphi'}| \leq |e^{-i\varphi}|$, and therefore the coherence ϱ_{12} is full or partially lost [57]. In particular, it occurs that ϱ_{12} vanishes for a large m and assuming a flat distribution $q_{\varphi'}$. In this case the state of the system ϱ is described by a purely classical mixture, although it resembles the measurement probabilities of $|\psi\rangle$.

The dephasing model

We will consider in further depth the quantum noise producing decoherence through dephasing. For that we will consider a *dephasing channel* $\mathcal{D}$, which leaves invariant the occupations of the input state in a preferred—say, the computational—basis $\{|i\rangle\}_{i=0}^{d-1}$, $\langle i|\mathcal{D}(\varrho)|i\rangle = \langle i|\varrho|i\rangle$, while damping or inducing phases on the coherences. By imposing that a dephasing channel is a CPTP map, one verifies that it acts as [59–62]

$$\mathcal{D}^G(\varrho) = \varrho \odot G \quad (1.1.38)$$

for some Gram matrix G , i.e., a non-negative matrix with diagonal entries equal to one. Here $\odot$ denotes the Schur product (also called entry-wise or Hadamard product). That is, $\mathcal{D}^G(\varrho)_{ij} = \varrho_{ij} G_{ij}$ where $|\mathcal{D}^G(\varrho)_{ij}| = |\varrho_{ij} G_{ij}| \leq |\varrho_{ij}|$.

Dephasing noise can lead to the full or partial loss of the resource of coherence [57], which plays a crucial role in most quantum protocols [2, 8, 36]. The reason why the definition of the dephasing noise is basis dependent, is because quantum protocols take place by measuring the inputs and outputs in a fixed basis which is determined by physical principles, e.g. the energy

eigenbasis in Hamiltonian dynamics or the basis where measurements are performed in a given protocol [2, 8, 36, 63].

Similarly as for quantum states, the decoherent noise acting on a quantum channel $\mathcal{E}$ through dephasing leaves invariant the diagonal entries of the Choi matrix in a preferred basis and damps the cross-terms [62]. A Choi matrix $J_{A_0 A_1}^{\mathcal{E}}$ describing a channel $\mathcal{E}_A$ is converted into $J_{A_0 A_1}^{\mathcal{E}} \odot G_{A_0 A_1}$ where $G_{A_0 A_1}$ is a Gram matrix with some additional structure [62]. The most general form of such dephasing transformation is realized by pre- and post-processings along with memory [64]. Assuming that there is no memory, dephasing noise acts as $\mathcal{E}' := \mathcal{D}^{G'} \circ \mathcal{E} \circ \mathcal{D}^G$ where $\circ$ means channel composition. Then, the Choi matrix $J_{A_0 A_1}^{\mathcal{E}}$ of a quantum channel $\mathcal{E}_A$ is mapped to [62]

$$J_{A_0 A_1}^{\mathcal{E}'} = J_{A_0 A_1}^{\mathcal{E}} \odot (G_{A_0} \otimes G'_{A_1}). \quad (1.1.39)$$

Complete decoherence: from quantum to classical

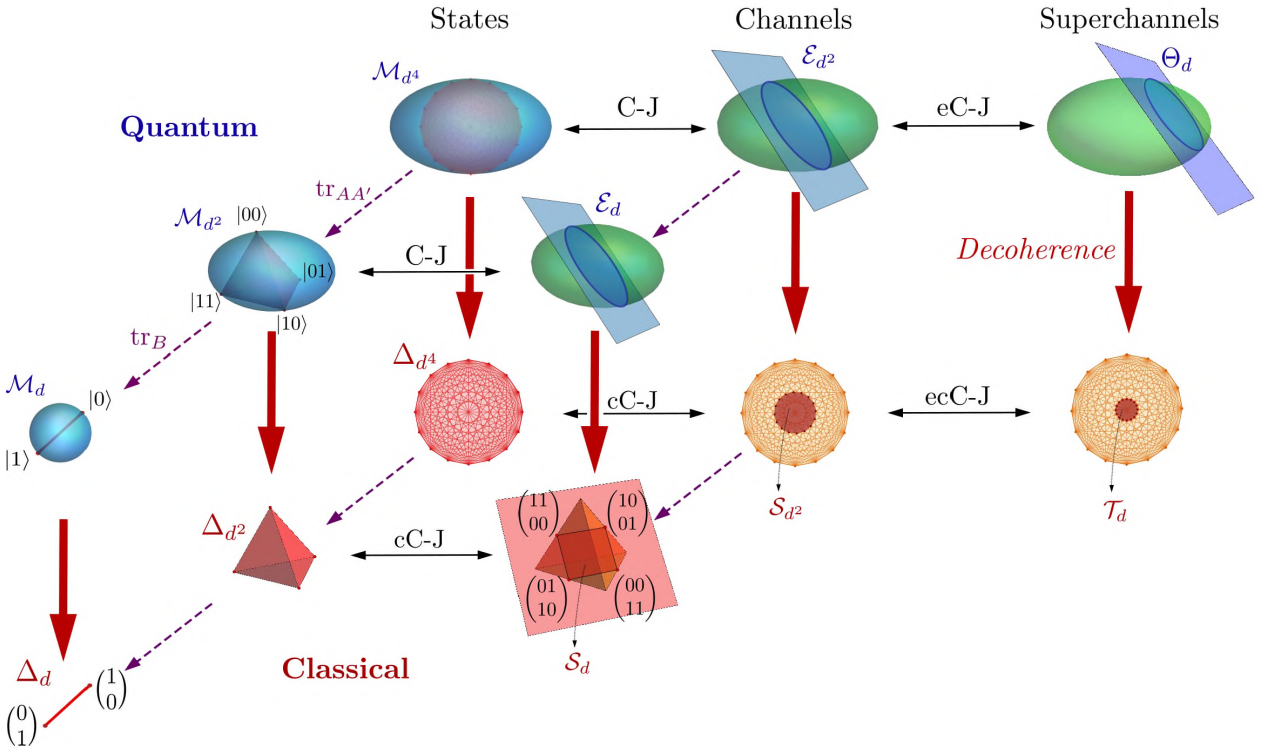


Figure 1.1.3: Complete decoherence of states, channels and superchannels. Blue convex sets in the left represent sets of quantum states consisting of 4, 2 and 1 subsystems of size d related by partial trace (directed dashed arrow). Bidirectional solid arrows denote the Choi-Jamiołkowski isomorphism (C-J): a cross section of the set $\mathcal{M}_{d^2}$ defines through Eq. (5.2.2) the set $\mathcal{E}_d$ of quantum channels (blue section of the green convex set). An extension of the Choi-Jamiołkowski isomorphism (eC-J) links the set of two-qudit channels $\mathcal{E}_{d^2}$ with one-qudit superchannels Θ_d through Eq. (5.2.4c). Under decoherence (red thick arrows) these objects become classical polytopes depicted in red and contained inside larger orange polytopes: a quantum state $\varrho \in \mathcal{M}_d$ is mapped to a vector $p = \text{diag}(\varrho) \in \Delta_d$, a quantum channel $\mathcal{E} \in \mathcal{E}_d$ is mapped into a stochastic matrix $S \in \mathcal{S}_d$, while a quantum superchannel from γ_d is transformed into an element of the set $\mathcal{T}_d$ of classical supermaps. The sets $\mathcal{M}_{d^2}$ of bipartite quantum states and $\mathcal{S}_{d^2}$ of bipartite stochastic matrices are both obtained from the set $\mathcal{M}_{d^4}$ of four-partite states and inherit some of its properties.

Consider now complete decoherence occurring by a completely dephasing channel $\mathcal{D}^{\mathbb{1}}$, acting as in Eq. (1.1.38) where $G = \mathbb{1}$. Then, a quantum state $\varrho \in \mathcal{M}_d$ loses all off-diagonal terms and becomes a probability vector $p = \text{diag}(\varrho) \in \Delta_d$. Similarly, when complete decoherence occurs in

both the input and output systems of a quantum channel $\mathcal{E}_A$, the resulting operation is given by the diagonal of the Choi matrix [62]. That is, one has

$$J^{\mathcal{D}^1 \circ \mathcal{E} \circ \mathcal{D}^1} = \iota[\text{diag}(J^{\mathcal{E}})] = \sum_{i=0}^{d_{A_0} d_{A_1} - 1} |i\rangle\langle i| J^{\mathcal{E}} |i\rangle\langle i|, \quad (1.1.40)$$

where ι embeds a vector into the diagonal of a matrix. The CPTP condition on $\mathcal{E}_A$ implies that the vector $\text{diag}(J^{\mathcal{E}})$ can be reshaped into a $d_{A_1} \times d_{A_0}$ stochastic matrix $S^{\mathcal{E}}$, which is a linear map between the d_{A_1} - and d_{A_0} -dimensional probability simplices. The stochastic map $S^{\mathcal{E}}$, known as classical action [62, 63, 65–67], will be called here the *decoherent action* of $\mathcal{E}$ [62] to emphasize that it corresponds to a possibly nonlocal channel arising from the effects of decoherence. The decoherent action can be obtained from Kraus operators K_i of $\mathcal{E}$ through $S^{\mathcal{E}} = \sum_i K_i \odot K_i^*$, where $*$ denotes the complex conjugation [12].

The effect of complete decoherence on quantum sets discussed above is summarized in Fig. 1.1.3. These sets are obtained from the set $\mathcal{M}_{d^4}$ of four-partite states in different ways. Bipartite and monopartite states are obtained by partial trace. The probability simplices of corresponding sizes are obtained from these sets under decoherence. Bipartite and monopartite quantum channels are obtained from four-partite and bipartite states through the cross-section of Eq. (5.2.2). Then a classical map (stochastic matrix) is obtained due to decoherence as in Eq. (1.1.40) from the diagonal entries of the Choi matrix J describing the bipartite channel. The transition from channels to superchannels can be seen as an extended Choi-Jamiołkowski isomorphism, as superchannels are isomorphic to one-way quantum nonsignaling channels through Eq. (5.2.4c). These decohere to one-way nonsignalling distributions through Eq. (5.2.5).

Thus one identifies classical subsets embedded in quantum subsets at different levels: general distributions, conditional distributions and nonsignaling distributions are embedded in quantum states, channels and superchannels. In this way, recent studies on the geometry of the sets of local and nonlocal correlations [68, 69] are directly linked to the research on geometry of quantum states [12, 70, 71] and geometry of quantum entanglement [72, 73]. Furthermore, analysis of the relative volume of the set of bipartite local correlations [74–76] can be compared with the problem of estimating the relative volume of the set of bipartite separable states [77–81].

1.2 Further concepts

1.2.1 Semidefinite programming

Overview and applications

Convex optimization problems consist of maximizing (or minimizing) a convex function over a convex set. A particular case consists of *semidefinite programs (SDP)*, where a linear objective function is maximized over a convex set (called *feasible set*) satisfying prescribed linear constraints. SDPs are efficient to compute and their outcome provides with a certificate: the output is either an example maximizing the objective function (if the problem is feasible), or a certificate of infeasibility (if the problem is infeasible). SDPs are useful in many areas like control systems [82], and in particular for problems in quantum information theory. Here one often aims to find convex relaxations to NP-Hard problems and asymptotically convergent solutions called *complete hierarchies*. Examples of the use of SDP are the separability problem [83, 84], bounding the set of quantum correlations [85], quantum steering [86], and assessing compatibility of POVM measurements [6].

Primal and dual formulations

The *primal* formulation of a semidefinite program reads [87, 88]

$$\begin{aligned} & \underset{X \in \mathbb{C}^{d \times d}}{\text{minimize}} && \text{tr}(CX) \\ & \text{such that} && \text{tr}(A_i X) = b_i \quad \forall i \\ & && X \geq 0. \end{aligned} \tag{1.2.1}$$

Here the variable X is required to be in the cone of positive semidefinite matrices, $C = C^\dagger$ is Hermitian, $b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$ is a vector with m real components and $A_i \in \mathbb{R}^{d \times d}$ are symmetric matrices. The dual problem associated to the primal formulation of an SDP reads

$$\begin{aligned} & \underset{y \in \mathbb{R}^m}{\text{maximize}} && b^T y \\ & \text{such that} && C - \sum_{i=1}^m y_i A_i \geq 0, \end{aligned} \tag{1.2.2}$$

where $y = (y_1, \dots, y_m)^T$ is the variable of the dual problem whose size m coincides with the size of b . An SDP is *feasible* if it has solution, namely a *feasible point*. So-called *feasibility problems* consist of asking whether the constraints in (1.2.1) can be satisfied (namely whether there exists a feasible point). These are obtained by setting $C = 0$ in the primal formulation, Eq. (1.2.1); or $b_i = 0$ in the dual formulation, Eq. (1.2.2).

Weak duality and certificates

One verifies that if both the primal and the dual formulations are feasible, then $\epsilon := \text{tr}(CX) + b^T y = \text{tr}(X(C - \sum_{i=1}^m y_i A_i)) \geq 0$. The fact that ϵ is nonnegative is known as *weak duality*, and it allows to assert feasibility as follows: a solution for the dual problem provides a lower bound to the primal solution, and a solution for the primal problem conversely upper bounds the solution for the dual problem. Therefore, a contradiction between these bounds implies that the problem has no solution. For example, consider Eq. (1.2.2) as a feasibility problem, namely setting $b_i = 0$. If there exists C such that $\text{tr}(CX) < 0$ in Eq. (1.2.1), then $\epsilon < 0$ which is a contradiction, and therefore the dual problem Eq. (1.2.2) must be infeasible. In this case, C acts as a so-called *certificate* of infeasibility. This technique is used in entanglement detection problems, as follows [89]: One sets $\{A_i\}$ to be an orthonormal basis of Hermitian matrices, such that the solution components y_i are for example the Bloch coefficients. An entanglement detection problem can thus be formulated as a feasibility problem ($b_i = 0$) in the variable y of Bloch coefficients, by imposing certain necessary constraints for separability. An example is the complete hierarchy established by Doherty-Parrilo-Spedalieri (DPS) [83]. Then, if there exists C Hermitian such that $\text{tr}(CX) < 0$, such matrix C acts as an entanglement witness [4], which provides a certificate of infeasibility [83, 84, 89].

1.2.2 The symmetric group

Some basics of group theory

Here we will sketch a few basic concepts about group theory which play a relevant role in this thesis. A *group* $G = \{g, *\}$ is a set of elements g equipped with a product $*$, satisfying that

1. $g * h \in G$ for all $g, h \in G$
2. $(g * h) * f = g * (h * f)$ for all $g, h, f \in G$
3. G contains the identity, e , namely $e * g = g * e = g$ for all $g \in G$

4. for all $g \in G$ there exists an inverse g^{-1} such that $g * g^{-1} = e$.

An example is the *general linear group* $GL(d)$ of $d \times d$ invertible matrices.

The *conjugation* of an element $a \in G$ with $g \in G$ is defined as gag^{-1} . This defines for each $a \in G$ the *conjugacy class* of a , namely the set of elements that can be obtained by conjugating a with some element g , $Cl(a) := \{b \in G \mid b = gag^{-1}, g \in G\}$. The *center* of a group G is the set of elements $c \in G$ which commute with all the group elements, namely

$$c * g = g * c \quad \forall g \in G. \quad (1.2.3)$$

Some basics of representation theory

In Physics, it is often the case that identifying group properties in a physical system allows for a better understanding. In particular, most physical theories are built upon vector spaces where to identify and use such properties. Therefore, one often considers a *representation* of a group G on a vector space V , which is a linear map T from each element $g \in G$ to a matrix $T(g)$ acting on a vector space V , which preserves the product,

$$T(g * h) = T(g) \cdot T(h) \quad \forall g, h \in G, \quad (1.2.4)$$

where $\cdot$ is the standard matrix product and is often omitted for simplicity. Thus the matrices $T(g)$ are isomorphic to the group elements themselves: they form a closed set under matrix multiplication, satisfy associativity, include the identity matrix $\mathbf{1}$, and are invertible. For example, the diagonal action of the general linear group $GL(d)$ on n copies of the Hilbert space $(\mathbb{C}^d)^{\otimes n}$ is described by simultaneous matrix multiplication in its diagonal action,

$$T(g)(|v_1\rangle \otimes \dots \otimes |v_n\rangle) = g|v_1\rangle \otimes \dots \otimes g|v_n\rangle, \quad g \in GL(d). \quad (1.2.5)$$

Consider a group G acting on a vector space V . The *invariant subspace* of each of these representations T is the subspace $W \subset V$ which is invariant under the action of η , namely all its vectors $w \in W$ satisfy that $T(g)w \in W$ for all $g \in G$. The *character* χ of a representation T (of a group G) associates to each element $g \in G$, the trace of its representation,

$$\chi_T(g) := \text{tr}(T(g)). \quad (1.2.6)$$

A representation T is *irreducible* if the vector space V where it acts has only two invariant subspaces: the 0-dimensional vector space, and itself.

Permuting tensor factors

Let us focus on the *permutation group* or *symmetric group* S_n composed of $n!$ elements $\sigma \in S_n$ which are permutations of n elements $1, 2, \dots, n$. The permutations can be decomposed as $\sigma = (\alpha_1)(\alpha_2)\dots(\alpha_t)$. Each *cycle* (α_i) is the single generator of a cyclic permutation subgroup, and the number and length of the cycles determine the conjugacy class of σ . For example, the 3-element permutation acting as $\sigma[1, 2, 3] = [2, 1, 3]$ will be written as $\sigma = (12)(3)$, and it belongs to the same conjugacy class as $(1)(23)$. Here we will adopt the convention that all permutations start by 1, and short-hand denote the identity permutation as $(1)(2)\dots(n) := (\text{id})$.

Each irreducible representation of the symmetric group, which we denote as $R_\lambda(\sigma)$ for each permutation σ , is characterized by a partition $\lambda \vdash n$ of n , namely a vector $\lambda = (\lambda_1, \dots, \lambda_{|\lambda|})$ with $|\lambda|$ components summing to n , $\lambda_1 + \dots + \lambda_{|\lambda|} = n$ where λ_i are positive integers ordered nonincreasingly, $\lambda_i \geq \lambda_{i+1}$. Each partition is associated to a *Young tableau* with $|\lambda|$ rows of length λ_i each, where $|\lambda|$ is often called the *height* of the tableau. For clarity, we will sometimes complete the vector $\lambda = (\lambda_1, \dots, \lambda_{|\lambda|}, 0, \dots, 0)$ to have n components. For example, the integer $n = 2$ admits

Partition	Tableau	Character	Subspace $\mathcal{V}_\lambda$	Dimension				
$\lambda = (3, 0, 0)$	<table><tr><td></td><td></td><td></td></tr></table>				$\chi_3(\sigma) = 1$	Symmetric $\mathcal{V}_3$	$\text{tr}(P_3) = 4$	
$\lambda = (2, 1, 0)$	<table><tr><td></td><td></td></tr><tr><td></td><td></td></tr></table>					$\chi_{2,1}((\text{id})) = 2$ $\chi_{2,1}((\alpha)) = -1$ $\chi_{2,1}((\alpha)(\beta)) = 0$	Standard $\mathcal{V}_{2,1}$	$\text{tr}(P_{2,1}) = 4$
$\lambda = (1, 1, 1)$	<table><tr><td></td></tr><tr><td></td></tr><tr><td></td></tr></table>				$\chi_{1^3}(\sigma) = \text{sign}(\sigma)$	Antisymmetric $\mathcal{V}_{1^3}$	$\text{tr}(P_{1^3}) = 1$	

Table 1.1: Schur-Weyl decomposition of the three-qutrit Hilbert space $(\mathbb{C}^3)^{\otimes 3}$. Each partition $\lambda = (\lambda_1, \lambda_2, \lambda_3) \vdash 3$ corresponds to a Young tableau whose i 'th row has length equal to the λ_i 'th component of λ . The character $\chi_\lambda(\sigma)$ depends on the conjugacy class of each permutation σ , namely its cycle type. For example, $\chi_{2,1}((12)(3)) = 0$ holds for $\lambda = (2, 1)$ because $(12)(3)$ decomposes as $(\alpha) = (12)$ and $(\beta) = (3)$. For each partition λ , the characters define through Eq. (1.2.12) the Young projector P_λ onto the corresponding isotypic subspace composed of copies of the irreducible subspace $\mathcal{V}_\lambda \subset (\mathbb{C}^3)^{\otimes 3}$. The dimension of each irreducible subspace is given by $\dim(\mathcal{V}_\lambda) = \text{tr}(P_\lambda)$.

two partitions, $\lambda_1 = (2, 0)$ and $\lambda_2 = (1, 1)$ since $2 = 2 + 0 = 1 + 1$; and the integer $n = 3$ admits three bipartitions: $(3, 0, 0)$, $(2, 1, 0)$ and $(1, 1, 1)$.

In this thesis, an important representation of the symmetric group S_n is the representation acting on the n -qudit Hilbert space of homogeneous local dimension d , $(\mathbb{C}^d)^{\otimes n}$. This will be denoted as $\eta_d(\sigma)$ with $\sigma \in S_n$, and its action on n qudits is given by

$$\eta_d(\sigma) |u_1\rangle \otimes \dots \otimes |u_n\rangle = |u_{\sigma^{-1}(1)}\rangle \otimes \dots \otimes |u_{\sigma^{-1}(n)}\rangle \in (\mathbb{C}^d)^{\otimes n}. \quad (1.2.7)$$

for all tensor product vectors. An example for two qudits $\mathbb{C}^d \otimes \mathbb{C}^d$ is the swap operator, $V |u\rangle \otimes |v\rangle = |v\rangle \otimes |u\rangle$.

Schur-Weyl duality

The Schur-Weyl duality states that the n -qudit Hilbert space can be block diagonalized under the joint action of the symmetric group S_n and the general linear group $GL(d)$ as follows [90, 91]. There exists an isomorphism between the n -qudit Hilbert space $(\mathbb{C}^d)^{\otimes n}$ and a direct sum of subspaces with a tensor product structure,

$$(\mathbb{C}^d)^{\otimes n} \simeq \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathcal{U}_\lambda \otimes \mathcal{S}_\lambda, \quad (1.2.8)$$

where $\mathcal{U}_\lambda$ and $\mathcal{S}_\lambda$ are the invariant subspaces of an irreducible representation labelled by λ of the unitary group $U(d)$ and the symmetric group S_n , respectively.

The Schur-Weyl duality implies that any operator A acting on $(\mathbb{C}^d)^{\otimes n}$ that is invariant under the conjugate diagonal action of the symmetric group or the general linear group, namely permutations $\eta_d(\sigma)$ with $\sigma \in S_n$ or unitaries $U^{\otimes n}$ with $U \in U(d)$, can be block-diagonalized in some basis as [91]

$$A = \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} Q_\lambda(u) \otimes R_\lambda(a), \quad (1.2.9)$$

where $Q_\lambda(u)$ and $R_\lambda(a)$ are irreducible representations of unitary group algebra, $u \in \mathbb{C}U(d)$, and the symmetric group algebra, $a \in \mathbb{C}S_n$. Moreover, since the group actions of $U(d)$ and S_n in

the Hilbert space are the unique commutant of each other, the irreducible representation $Q_\lambda(u)$ in (1.2.9) can only be the identity. Thus we have

$$A = \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathbb{1}_{s_\lambda} \otimes R_\lambda(a), \quad (1.2.10)$$

where $a = \sum_{\sigma \in S_n} a_\sigma \sigma \in \mathbb{C}S_n$ and the dimension s_λ of the identity corresponds to the multiplicity of H_λ and is given by the Schur polynomial evaluated on $|\lambda|$ ones,

$$s_\lambda := s_\lambda(1, \dots, 1) = \prod_{1 \leq i < j \leq |\lambda|} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \quad (1.2.11)$$

For each partition $\lambda \vdash n$ one defines a centrally primitive idempotent $p_\lambda = \sum_{\sigma} \chi_\lambda(\sigma) \sigma / n!$, which commutes with all permutations $p_\lambda \sigma = \sigma p_\lambda$ and thus is invariant under the average $\sum_{\sigma \in S_n} \sigma p_\lambda \sigma^{-1} = p_\lambda$. Its representation on $(\mathbb{C}^d)^{\otimes n}$ is known as a *Young projector*,

$$P_\lambda = \frac{1}{n!} \sum_{\sigma} \chi_\lambda(\sigma) \eta_d(\sigma), \quad (1.2.12)$$

which projects onto a subspace of $(\mathbb{C}^d)^{\otimes n}$. Each Young projector P_λ projects onto the corresponding invariant subspace $\mathcal{V}_\lambda \subset (\mathbb{C}^d)^{\otimes n}$. By construction, the Young projectors are orthogonal and satisfy the identity resolution

$$\sum_{\lambda \vdash n} P_\lambda = \mathbb{1}_{d^n}. \quad (1.2.13)$$

Some examples

The two-qubit Hilbert space decomposes as $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathcal{V}_{(2,0)} \oplus \mathcal{V}_{(1,1)}$ in two subspaces: the three-dimensional *symmetric* subspace $\mathcal{V}_{(2,0)}$, defined by the Young projector

$$P_{(2,0)} = \frac{1}{2} \left(\eta_2(\text{id}) + \eta_2(12) \right) = |00\rangle\langle 00| + |11\rangle\langle 11| + |\psi^+\rangle\langle \psi^+|, \quad (1.2.14)$$

where $|00\rangle$, $|11\rangle$ and $|\psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ are eigenvalues of the swap $V = \eta_2(12)$ with eigenvalue 1 and thus are permutationally symmetric; and the one-dimensional *antisymmetric subspace* $\mathcal{V}_{(1,1)}$, defined by the Young projector

$$P_{(1,1)} = \frac{1}{2} \left(\eta_2(\text{id}) - \eta_2(12) \right) = |\psi^-\rangle\langle \psi^-|, \quad (1.2.15)$$

where the *singlet* vector $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is permutationally antisymmetric as $V|\psi^-\rangle = -|\psi^-\rangle$.

A less trivial example is the decomposition of the three-qutrit Hilbert space $(\mathbb{C}^3)^{\otimes 3}$, shown in Table 1.1. The Young projectors of the symmetric, standard and antisymmetric subspaces will be used in Chapter 2 and are given by

$$P_3 = \frac{1}{6} \eta_3 \left([(\text{id}) + (12) + (13) + (23) + (123) + (132)] \right), \quad (1.2.16)$$

$$P_{2,1} = \frac{1}{6} \eta_3 \left([2(\text{id}) - (123) - (132)] \right), \quad (1.2.17)$$

$$P_{1^3} = \frac{1}{6} \eta_3 \left([(\text{id}) - (12) - (13) - (23) + (123) + (132)] \right). \quad (1.2.18)$$

1.2.3 The set of quantum measurements

From probability distributions to quantum measurements

A discrete classical probabilistic measurement with n possible outcomes is described by a column probability vector $p = (p_1, \dots, p_n)^T$ with probabilities $p_j \geq 0$ normalized to $\sum_{j=1}^n p_j = 1$. In the quantum setup, a measurement is a set of positive operators P_j summing to the identity. As a natural generalization of the classical case, quantum measurements on a system of dimension d can be described as a *blockwise probability vector* [14, 92], namely a column vector

$$\mathbf{P} = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix}, \quad (1.2.19)$$

with n components P_j being Hermitian, positive semidefinite matrices of order d , $P_j \geq 0$, satisfying the identity resolution

$$\sum_{j=1}^n P_j = \mathbf{1}_d. \quad (1.2.20)$$

Thus a blockwise probability vector $\mathbf{P} = (P_1, \dots, P_n)^\dagger$ corresponds to a generalized quantum measurement, whose effects P_j play the role of vector components over an operator algebra. The classical notion of probability vectors with n real nonnegative components is recovered for $d = 1$. This can be obtained by taking the expectation value of each effect with respect to a quantum state ϱ acting on $\mathbb{C}^d$, which gives the probability $p_j = p(j) = \text{tr}(P_j \varrho)$ of obtaining the outcome j . A similar notion of blockwise vector was introduced in [93].

The quantum analog of the probability simplex

The probability simplex Δ_n of n -point probability vectors

$$\Delta_n = \left\{ \sum_{j=1}^n v_j p_j \right\}_{p_j \geq 0, \sum_j p_j = 1} \subset \mathbb{R}^{n-1}, \quad (1.2.21)$$

is given by the convex hull of the vectors v_j ,

$$v_1 = (1, 0, \dots, 0)^T, v_2 = (0, 1, 0, \dots, 0)^T, \dots, v_n = (0, 0, \dots, 1)^T \quad (1.2.22)$$

which form its extremal points. The extremal points of the set of quantum measurements with n effects in dimension d induce a more involved geometry [94–101]. Here we present two generalizations of the standard convex geometry of the probability simplex.

Following [92, 102], we consider a decomposition of a blockwise probability vector $\mathbf{P}$ in the form of

$$\mathbf{P} = \sum_{j=1}^n v_j \otimes P_j, \quad (1.2.23)$$

where $P_j \geq 0$ and $\sum_{j=1}^n P_j = \mathbf{1}_d$. In this sense, the set of blockwise probability vectors can be seen as a generalization of the probability simplex Δ_n , as it can be obtained from the vectors v_j as

$$\left\{ \sum_{j=1}^n v_j \otimes P_j \right\}_{P_j \geq 0, \sum_j P_j = \mathbf{1}} \subset \mathbb{R}^{n-1} \otimes (\mathbb{C}^d \times \mathbb{C}^d). \quad (1.2.24)$$

The set $\Delta_{n,d}$ of blockwise probability vectors with n blocks of size d each contains the central point

$$\mathbf{P}_u = (\mathbf{1}_d/n, \dots, \mathbf{1}_d/n)^\dagger \quad (1.2.25)$$

and d particular extremal points

$$\mathbf{V}_j = v_j \otimes \mathbb{1}_d, \quad (1.2.26)$$

known as *fuzzy* POVMs [103], which generate the set of quantum measurements with n outcomes in a d -dimensional system, $\Delta_{n,d}$, through the non-standard convex combination of Eq. (1.2.23).

Concatenation of measurements

Another important generalization of convexity of the standard probability simplex Δ_n can be illustrated as follows: if two probability vectors p and q are contained in Δ_n , then their convex combination $c = \sqrt{a}p\sqrt{a} + \sqrt{1-a}q\sqrt{1-a}$ also belongs to Δ_n for any $a \in [0, 1]$. Here a positive number a is written as a product of its square roots to pave the road for the following generalization. The set $\Delta_{n,d}$ of blockwise probability vectors is *matrix convex* [104–106]: given two such vectors $\mathbf{P}$ and $\mathbf{Q}$ their *matrix-convex combination* $\mathbf{C}$, with components

$$C_j = \sqrt{A}U P_j U^\dagger \sqrt{A} + \sqrt{1-A}V Q_j V^\dagger \sqrt{1-A}, \quad (1.2.27)$$

belongs to $\Delta_{n,d}$ for any any unitary $U, V \in U(d)$ and any Hermitian matrix A such that $0 \leq A \leq \mathbb{1}$. To show this it is enough to realize that all components are positive, $C_j \geq 0$, and normalized, $\sum_{j=1}^n C_j = \mathbb{1}$. The concatenation of two measurements [14, 107] will be formally described by their matrix convex combination in Section 4.2.

A low-dimensional example

Understanding the convex geometric structure of the set of quantum measurements of arbitrary dimension d and number of effects n is an ongoing challenge [96, 99, 108], this task is much simpler in the case of $n = 2$, for the set of two-effect measurements $\Delta_{2,d}$. Its vectors consist of two effects, $\mathbf{P} = (P, \mathbb{1} - P)^\dagger$, and thus it is fully described by the set of positive semidefinite matrices of order d having eigenvalues between 0 and 1,

$$\Delta_{2,d} \simeq \{P : 0 \leq P \leq \mathbb{1}_d\}, \quad (1.2.28)$$

where $\simeq$ denotes isomorphism between the two sets. This means that $\Delta_{2,d}$ is defined by a fraction of the positive cone, which is obtained by the convex hull of pairs of orthogonal projectors.

Let us focus on the simplest nontrivial set of blockwise probability vectors $\Delta_{2,2}$, which describes a single-qubit measurement with two effects P and $\mathbb{1} - P$, and denote as λ_+ and λ_- the maximal and minimal eigenvalues of P respectively. Since P can be viewed as a quantum state up to a normalization factor $t := \text{tr}(P) = \lambda_+ + \lambda_-$, the difference of eigenvalues $\lambda_+ - \lambda_-$ can be viewed as the length of a generalized Bloch vector $\vec{\tau}$ with components $\tau_i = \text{tr}(P\sigma_i)$ over the Pauli basis $\{\sigma_i\}$,

$$\tau := |\vec{\tau}| = \lambda_+ - \lambda_- = 2\lambda_+ - t. \quad (1.2.29)$$

Here the length τ is bounded between 0 and 1, since one has $0 \leq \lambda_+ \leq 1$, and the variable $t = \text{tr}(P)$ is bounded between $0 \leq t \leq 2$.

The set $\Delta_{2,2}$ has the structure of two anti-parallel 4-dimensional cones glued at $t = \text{tr}(P) = 1$ along the Bloch ball $\Omega_2 \subset \mathbb{R}^3$ (see Figure 1.2.1), which is formed by the convex hull of the Bloch sphere $\partial\Omega_2 \subset \mathbb{R}^2$ containing all single-qubit pure states. The extremal points of the set $\Delta_{2,2}$ are the $\hat{0}$ and $\mathbb{1}$ operators, located in the apexes of the hypercone with $t = 0$ and $t = 2$ respectively, together with the sphere $\partial\Omega_2$. Therefore, $\Delta_{2,2}$ can be written as a convex hull of these sets,

$$\Delta_{2,2} = \text{conv}\left(\{\hat{0}\} \cup \partial\Omega_2 \cup \{\mathbb{1}\}\right) = \text{conv}\left(\{\hat{0}, U|\psi\rangle\langle\psi|U^\dagger, \mathbb{1}\}_U\right), \quad (1.2.30)$$

where $|\psi\rangle\langle\psi|$ is a fixed rank-one projector and the set is obtained by all possible rotations by unitary matrices $U \in \mathcal{U}(2)$. Therefore, the extremal points of $\Delta_{2,2}$ are composed of all projectors of rank

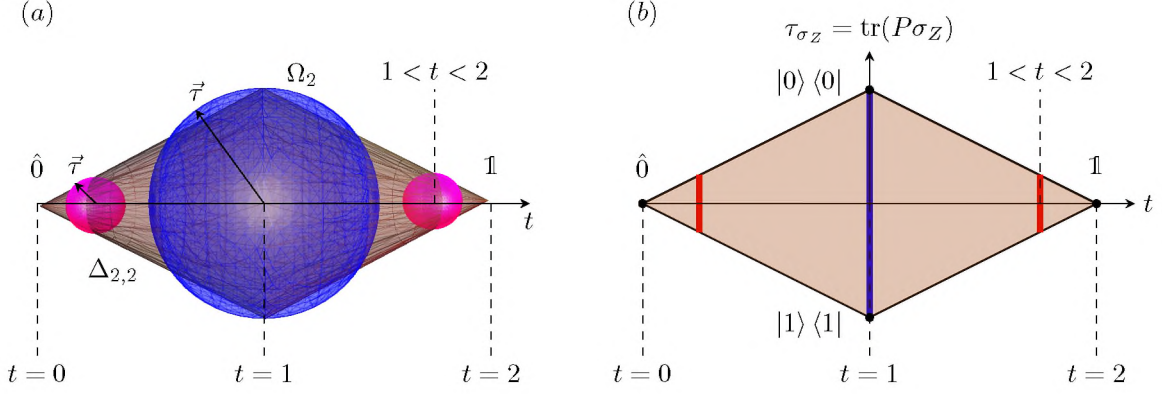


Figure 1.2.1: Visualization of the set of single-qubit measurements with two effects, $\Delta_{2,2}$. (a) The dark set $\Delta_{2,2}$ corresponds to the intersection of two antiparallel 4-dimensional hypercones with axial coordinate $t = \text{tr}(P)$ and 3-dimensional spherical sections parameterized by $\vec{\tau}$ (blue and red). The blue section defined by $t = 1$ defines the Bloch ball Ω_2 . Blockwise vectors $\mathbf{P}$ with $\text{tr}(P) \neq 1$ are described by the two distinct red sections located at $1 < t < 2$ and its complementary. (b) The dark set corresponds to the two-dimensional projection of $\Delta_{2,2}$ onto the plane with coordinates t and $\tau_{\sigma_Z} = \text{tr}(P\sigma_Z)$, so that its blue and red sections are lines of length $2|t - 1|$ parameterized by τ_{σ_Z} . The four extremal points (black dots) of this projection of $\Delta_{2,2}$ are $P = \hat{0}$ and $P = \mathbb{1}$, in the apexes, and $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$, in the edges of the intersection of the two hypercones.

0, 1 and 2.

As a result of Eq. (1.2.30), the hypervolume of the set $\Delta_{2,2}$ represented in Figure 1.2.1 is the hypervolume of two 4-dimensional hypercones of height $h := \max_P t/2 = 1$ and spherical base of radius $a := \max_P \tau = 1$, namely

$$V(\Delta_{2,2}) = 2 \times \frac{1}{4} h \left(\frac{4}{3} \pi a^3 \right) = \frac{2\pi}{3}. \quad (1.2.31)$$

This implies that the ratio between hypervolumes of $\Delta_{2,2}$ and the set of positive semidefinite operators P satisfying $0 \leq \text{tr}(P) \leq 2$ is $1/8$. In other words, sampling P from a flat Hilbert-Schmidt measure in the space of positive semidefinite operators bounded by $\text{tr}(P) \leq 2$, $\mathbf{P} = (P, \mathbb{1} - P)$ will be a valid measurement in $\Delta_{2,2}$ with probability $1/8$.

1.2.4 Nonlocality of bipartite channels: special subsets

In the framework of quantum channel nonlocality, we distinguish the following three classes of channels.

- *Local operations assisted by shared randomness* (LOSR): Operations $\mathcal{E}_{AB}$ which can be represented as the convex combination of *local operations* – tensor product of quantum channels,

$$\mathcal{E}_{AB} = \sum_{\lambda} p_{\lambda} \mathcal{E}_A^{\lambda} \otimes \mathcal{E}_B^{\lambda}, \quad (1.2.32)$$

where $p_{\lambda} \geq 0$, $\sum_{\lambda} p_{\lambda} = 1$ and $\mathcal{E}_{A(B)}^{\lambda}$ are quantum channels. These are the free operations in the resource theory of both Bell and Buscemi nonlocality [7, 109–111] and are analogous to separable bipartite quantum states. Channels outside the set of LOSR are termed *nonlocal* and can be compared to entangled states.

- *Localizable*, or *local operations and shared entanglement* (LOSE): Operations $\mathcal{E}_{AB}$ where Alice and Bob can do local operations assisted by a shared quantum state τ_{EF} [112, 113]. Namely,

$$\mathcal{E}(\varrho_{A_0 B_0}) = \text{tr}_{E_1 F_1} (\mathcal{E}_{AE} \otimes \mathcal{E}_{BF}(\varrho_{A_0 B_0} \otimes \tau_{E_0 F_0})). \quad (1.2.33)$$

- *Causal*, or *quantum nonsignaling* (qNS): Operations $\mathcal{E}_{AB}$ where Bob's (Alice's) output cannot be influenced by Alice's (Bob's) local operations [112, 114]. The marginals of their four-partite Choi matrices $J_{A_0 B_0 A_1 B_1}^{\mathcal{E}}$ satisfy

$$J_{A_0 B_0 B_1}^{\mathcal{E}} = \text{id}_{A_0} \otimes J_{B_0 B_1}^{\mathcal{E}} / d_{A_0}, \quad (1.2.34a)$$

$$J_{A_0 B_0 A_1}^{\mathcal{E}} = J_{A_0 A_1}^{\mathcal{E}} \otimes \text{id}_{B_0} / d_{B_0}, \quad (1.2.34b)$$

where partial trace is denoted by omitting the traced subsystems, for example $J_{A_0 B_0 B_1}^{\mathcal{E}} = \text{tr}_{A_1}(J_{A_0 B_0 A_1 B_1}^{\mathcal{E}})$. Analogous to nonsignaling correlations in Eq. (1.1.32), nonsignaling channels have also well-defined marginal channels proportional to $J_{A_0 A_1}^{\mathcal{E}}$ and $J_{B_0 B_1}^{\mathcal{E}}$.

We have the chain of inclusions (see Fig. 5.3.2)

$$\text{LOSR} \subset \text{LOSE} \subset \text{qNS}. \quad (1.2.35)$$

2

Entanglement detection with trace polynomials

In this chapter we provide a systematic method for nonlinear entanglement detection based on trace polynomial inequalities. In particular, this allows to employ multi-partite witnesses for the detection of bipartite states, and vice versa. We identify witnesses for which linear detection of an entangled state fails, but for which nonlinear detection succeeds. With the trace polynomial formulation a great variety of witnesses arise from immanant inequalities, which can be implemented in the laboratory through randomized measurements.

2.1 Introduction

The experimental detection of entanglement is an ongoing challenge [115–117], for which a key tool are entanglement witnesses [4]. These detect some entangled states by virtue of having a negative expectation value, separating them from the set of fully separable states, i.e. from convex combinations of product states. However, the construction of witnesses is not a straightforward task [4] and frequently relies on making use of specific structure in the state to be detected [118–125].

Nonlinear entanglement detection has recently become a focus of attention due to the development of the randomized measurement toolbox [126], making local unitary invariants experimentally accessible through single copy measurements. From a conceptual perspective, this replaces (complete or partial) state tomography by a set of local unitary invariants from which one aims to detect entanglement in the state.

Yet, the currently available techniques are limited and a systematic development of nonlinear witnesses is desirable. The aim of this paper is to provide such a systematic method that is not only suitable for the randomized measurement framework, but also makes a broader use of known constructions. The basic task we study is the following: *given an entanglement witness W , can it be employed also in a nonlinear fashion as to detect entanglement in multiple copies $\rho^{\otimes k}$?*

Here we answer this question in the affirmative. In particular we show that: i) having access to multiple copies of the state it is possible to detect entanglement locally, where the size of states and witnesses can be different [Observation 1 and Figs. 2.2.1 and 2.2.2]; ii) that there exist pairs of states and witnesses for which linear entanglement detection fails, but for which non-linear detection succeeds [Observation 2]; and iii) that there is a large class of nonlinear witnesses arising from trace polynomial inequalities [Fig. 2.3.1], which can analytically be treated in the group ring $\mathbb{C}S_n$ and that are experimentally accessible through randomized measurements [126].

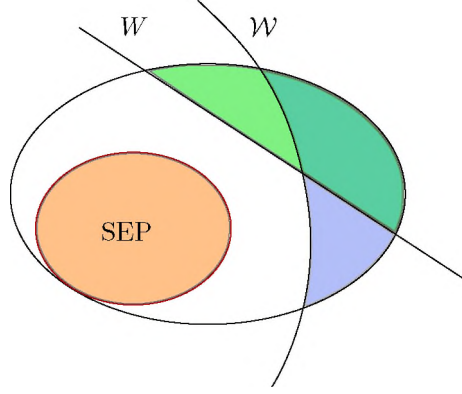


Figure 2.1.1: Linear versus nonlinear entanglement detection. A linear witness W defines a hyperplane in the state space, separating some entangled states (green area) from the rest; a non-linear witness $\mathcal{W}$ cuts the state space analogously in a nonlinear hypersurface, thus detecting a different set of entangled states (blue area).

2.2 Nonlinear entanglement witnesses

2.2.1 Detection with multiple copies

A quantum state is called separable if it can be written as convex combination of product states,

$$\varrho = \sum_i p_i \varrho_i^{(1)} \otimes \varrho_i^{(2)} \otimes \dots \otimes \varrho_i^{(n)}, \quad (2.2.1)$$

and entangled otherwise. Entanglement detection with a witness works in the following way: to detect an entangled state ϱ_{ent} , a witness is an observable W such that $\text{tr}(W \varrho_{\text{ent}}) < 0$ holds, while $\text{tr}(W \varrho_{\text{sep}}) \geq 0$ for all separable states $\varrho_{\text{sep}} \in \text{SEP}$. In this way, W acts as a hyperplane, separating a subset of entangled states from the rest. We call this *linear detection*, in contrast to nonlinear detection introduced below. This is illustrated in Fig. 2.1.1.

Beyond linear methods, nonlinear corrections to linear witnesses W can detect strictly more states [127–130] and witnesses $\mathcal{W}$ acting on larger Hilbert spaces have been considered [131–134] to detect entanglement in multiple copies in the form of $\varrho^{\otimes k}$. To use a witness in the multicopy scenario, one similarly asks that $\text{tr}(\mathcal{W} \varrho_{\text{ent}}^{\otimes k}) < 0$ holds, while $\text{tr}(\mathcal{W} \varrho_{\text{sep}}^{\otimes k}) \geq 0$ for all separable states $\varrho_{\text{sep}} \in \text{SEP}$ [135]. Setting $k = 1$ then recovers the standard use of witnesses.

How can one find such nonlinear witness $\mathcal{W}$? One approach is to take a tensor product of witnesses

$$\mathcal{W} = W_1 \otimes \dots \otimes W_n. \quad (2.2.2)$$

Naturally, the expectation values need to be computed in a manner such that the dimensions of $\mathcal{W}$ and $\varrho^{\otimes k}$ match (see Fig. 2.2.3 for a detailed explanation). For example, for a tripartite state ϱ_{ABC} , take the tensor product of three bipartite witnesses,

$$\mathcal{W} = U_{AA'} \otimes V_{BB'} \otimes W_{CC'}. \quad (2.2.3)$$

Then, the expression

$$\langle \mathcal{W} \rangle_{\varrho^{\otimes 2}} = \text{tr}(\mathcal{W}(\varrho_{ABC} \otimes \varrho_{A'B'C'})) \quad (2.2.4)$$

is nonnegative if ϱ is separable.

At first sight, it may not be clear why Eq. (2.2.4) can detect entanglement, since the witnesses act along the cuts $A|A'$, $B|B'$ and $C|C'$ and $\varrho_{ABC}^{\otimes 2}$ is separable in the cut $ABC|A'B'C'$. This apparent contradiction is resolved by realizing that the tensor product of witnesses for $A|A'$, $B|B'$, $C|C'$ is not necessarily a witness for $ABC|A'B'C'$. The following shows that this method can

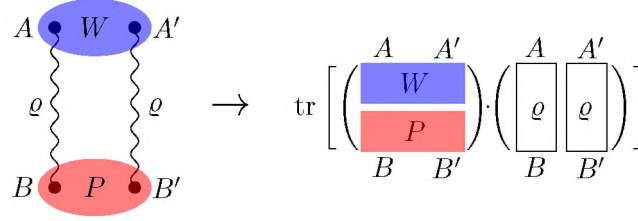


Figure 2.2.1: Sketch of the entanglement concentration scheme. Expressions of the type $\langle W \otimes P \rangle_{\rho^{\otimes 2}}$ can be obtained by first measuring two copies of Bob's subsystems (red bubble), and then two copies of Alice's subsystems (blue bubble). When $\rho = |\psi\rangle\langle\psi|$ is a pure state such that $|\psi\rangle = \mathbb{1} \otimes S |\phi^+\rangle$ with S invertible, Bob can teleport his part of ρ to Alice by using $P = |\varphi\rangle\langle\varphi|$ with $|\varphi\rangle = \mathbb{1} \otimes S^{\dagger-1} |\phi^+\rangle$. Then standard linear witness evaluation $\text{tr}(W|\psi\rangle\langle\psi|)$ on Alice's subsystem is obtained as a particular case of the nonlinear method proposed in this work.

indeed work:

Observation 1. Take the 2-qubit witnesses [50]

$$\begin{aligned} W &= \mathbb{1} - X \otimes X - Z \otimes Z \quad \text{and} \\ V &= |\phi^+\rangle\langle\phi^+|^{T_2}, \end{aligned} \tag{2.2.5}$$

where $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, and the Greenberger-Horne-Zeilinger state $\varphi = |\text{GHZ}\rangle\langle\text{GHZ}|$, where $|\text{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$. Then

$$\text{tr}((W_{AA'} \otimes W_{BB'} \otimes V_{CC'})(\varphi_{ABC}^{\otimes 2})) = -1/2. \tag{2.2.6}$$

Thus the Ansatz of Eq. (2.2.3) for a non-linear witness can detect entanglement. Moreover, this method can be used when linear detection fails. In the above example this is clear because the two-qubit reduction of the GHZ state is separable, and thus cannot be detected with any two-qubit witness. An example for $k = n$, in which nonlinear detection can be compared to linear one more naturally, is the Bell state $|\psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. One has $\langle W \rangle_{\psi^+} = \langle V \rangle_{\psi^+} = 1$ but $\langle W \otimes V \rangle_{\psi^+ \otimes 2} = -1/2$, and therefore $|\psi^+\rangle$ becomes detectable with the witnesses W and V .

2.2.2 Entanglement concentration

If in Eq. (2.2.2) we replace some witnesses W_i by positive semidefinite operators P_i , the expectation value with respect to k copies of an n -qudit separable state remains nonnegative. The following observation shows that this way, a witness with nonnegative expectation value with respect to a certain entangled state can detect its entanglement.

Observation 2. There exist pairs of states ρ and witnesses W for which $\text{tr}(W\rho) \geq 0$, but for which there is $P \geq 0$ such that

$$\text{tr}((W \otimes P)\rho^{\otimes k}) < 0. \tag{2.2.7}$$

In short: linear entanglement detection with W fails, but nonlinear detection with $W \otimes P$ succeeds.

One can understand the entanglement concentration scheme as a variant of the entanglement swapping protocol. To see this, suppose Alice and Bob share two copies of a full Schmidt rank bipartite pure state $|\psi\rangle$. Since $|\psi\rangle$ is pure, it can be written as

$$|\psi\rangle = \mathbb{1} \otimes S |\phi^+\rangle = S^T \otimes \mathbb{1} |\phi^+\rangle \tag{2.2.8}$$

where S is a $d \times d$ complex matrix with $\text{tr}(S^\dagger S) = 1$. By assumption S is square and full rank, and therefore it has inverse S^{-1} .

In this case, Bob can project onto $|\varphi\rangle\langle\varphi|$ with $|\varphi\rangle = \mathbb{1} \otimes S^{\dagger-1} |\phi^+\rangle$, thus obtaining a state ξ with certain probability p . It will be useful to consider the outcome state up to normalization, $\zeta = p\xi$, which reads

$$\zeta = \text{tr}_{BB'}((\mathbb{1}_{AA'} \otimes |\varphi\rangle\langle\varphi|_{BB'}) (|\psi\rangle\langle\psi|_{AB} \otimes |\psi\rangle\langle\psi|_{A'B'})). \quad (2.2.9)$$

To compute ζ it will be convenient to use the row-column notation sketched in Fig. 2.2.1, in which local parties are displayed in rows and copies are displayed in columns. The first factor inside the trace contains the measurement operators each party performs, and the second factor contains the copies of the state. This notation is further explained in Fig. 2.2.3. Then it becomes clear that ζ reads

$$\zeta = \frac{1}{d^3} \sum_{i,j,k,l,r,s=0}^{d-1} \text{tr}_{BB'} \left(\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ |i\rangle\langle j| & S^{\dagger-1} |i\rangle\langle j| S^{-1} \end{pmatrix} \begin{pmatrix} S^T |k\rangle\langle l| S^* & |r\rangle\langle s| \\ |k\rangle\langle l| & S |r\rangle\langle s| S^{\dagger} \end{pmatrix} \right) \quad (2.2.10)$$

$$= \frac{1}{d^3} \sum_{i,j,k,l,r,s=0}^{d-1} \begin{matrix} S^T |k\rangle\langle l| S^* & |r\rangle\langle s| \\ \langle j|k\rangle\langle l|i\rangle & \langle j|r\rangle\langle s|i\rangle \end{matrix} \quad (2.2.11)$$

$$= \frac{1}{d^3} \sum_{r,s=0}^{d-1} S^T |r\rangle\langle s| S^* \otimes |r\rangle\langle s| = \frac{1}{d^2} |\psi\rangle\langle\psi|, \quad (2.2.12)$$

which means that the final state on Alice's side is $|\psi\rangle$ with probability $p = 1/d^2$ (given that Bob's outcome is $|\varphi\rangle$). This procedure is similar to the entanglement swapping protocol [8, 36], except for the fact that we consider any pure quantum state with full Schmidt rank instead of restricting ourselves to maximally entangled states. When Alice obtains $|\psi\rangle$, she can detect its entanglement linearly with a witness W fulfilling $\langle\psi|W|\psi\rangle < 0$.

This can be summarized as follows: if $|\psi\rangle$ is full Schmidt rank and it can be detected linearly by a witness W_ψ , then it can also be detected using entanglement concentration with $\mathcal{W} = |\varphi\rangle\langle\varphi| \otimes W_\psi$.

2.2.3 An example: detecting noisy GHZ states

Consider the n -qudit GHZ state affected by white noise,

$$\varrho = p|\text{GHZ}\rangle\langle\text{GHZ}| + \frac{1-p}{d^n} \mathbb{1}, \quad (2.2.13)$$

whose experimental generation and detection plays a central role in quantum computing and communication setups [136–138]. We use the n -qudit witness

$$W = \mathbb{1} - n!P_{1^n}, \quad (2.2.14)$$

where P_{1^n} is the projector onto the fully antisymmetric subspace of n qudits. The state ϱ has a nonpositive partial transpose (and thus is entangled) for $p > 1/(1 + d^{n-1})$ [139, 140]. Since $\text{tr}(|\text{GHZ}\rangle\langle\text{GHZ}|P_{1^n}) = 0$, we have

$$\text{tr}(W\varrho) \geq 0 \quad (2.2.15)$$

for all values of p and thus single-copy detection fails.

However, given $\mathcal{W} = (\mathbb{1} - k!P_{1^k}) \otimes P_{1^k}^{\otimes n-1}$, one verifies that

$$\text{tr}(\mathcal{W}\varrho^{\otimes k}) < 0 \quad (2.2.16)$$

for values of p shown in Fig. 2.2.2. Eq. (2.2.16) is evaluated using the following result.

Lemma 1. *Let $\mathcal{W} = (\mathbb{1} - k!P_{1^k}) \otimes P_{1^k}^{\otimes n-1}$. The expectation value of $\mathcal{W}$ with respect to k copies of*

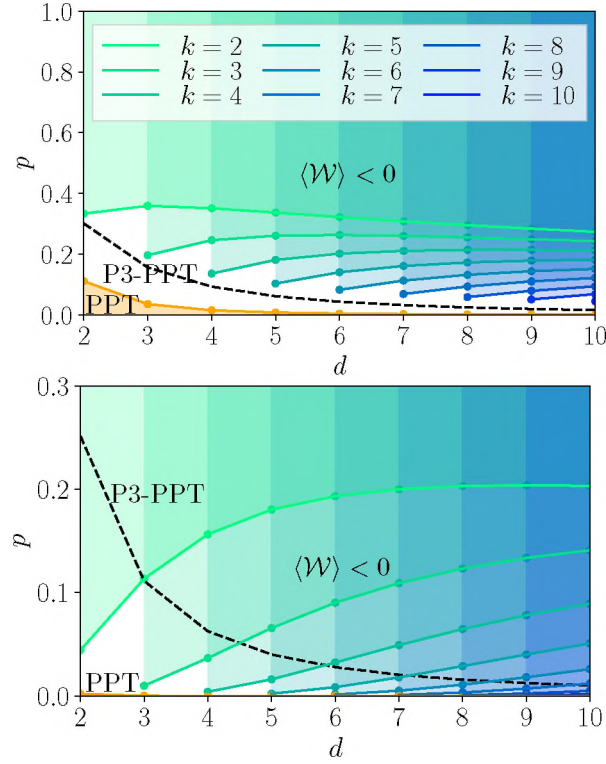


Figure 2.2.2: Nonlinear detection of noisy 4- (top) and 10-partite (bottom) GHZ states with LU-invariants. Shown are the detection curves with $k = 2, 3, \dots, 10$ copies of noisy n -partite GHZ states shared among an even number of parties n with local dimension d (top $n = 4$; bottom $n = 10$). We assume $d \geq k$ as otherwise the anti-symmetrizer vanishes. All states with positive partial transpositions (PPT) are in the orange region, and hence states outside are entangled; the green-blue regions are states that are detected by Eq. (2.2.16) with $W = \mathbb{1} - k!P_{1^k}$, which does not detect the GHZ state linearly, and $\mathcal{W} = W \otimes P_{1^k}^{\otimes n-1}$. Entangled states that are not detected by $\mathcal{W}$ are in the white region, suggesting that detection gains robustness with the number of copies. The dashed line denotes the detection threshold by the P3-PPT criterion proposed in [43]. The evaluation of $\langle \mathcal{W} \rangle$, the P3-PPT criterion, and the positive partial transpose criterion is detailed in Section 2.2.3.

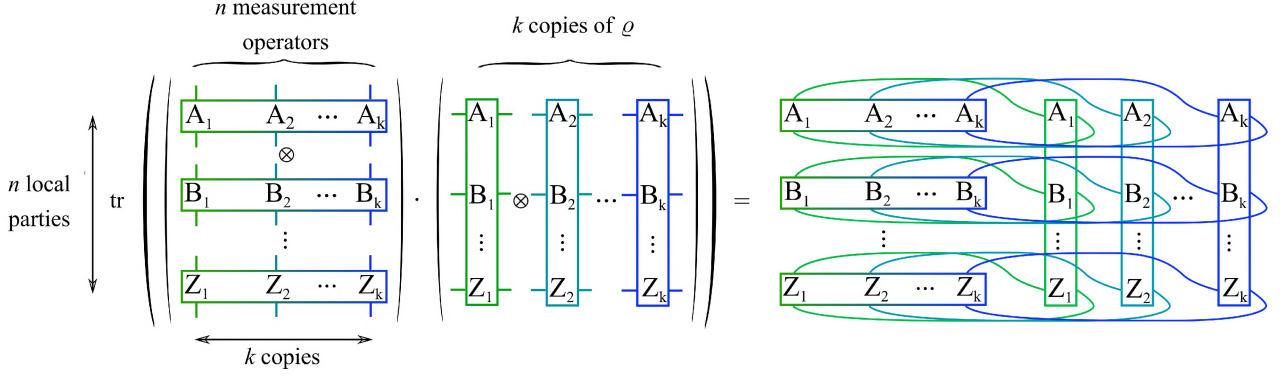


Figure 2.2.3: Graphical representation of the row-column notation for tensor products. The full space is composed of n local parties $A, \dots, Z$, each of which is copied k times. We denote the i 'th copy of each party as $A_i, \dots, Z_i$. In our expressions we compute the expectation values of n operators (witnesses or projectors) acting across k copies of each local subsystem, e.g. $A_1 \otimes A_2 \dots A_k$, with respect to k operators (k copies of a quantum state) acting on a single copy of all n parties, e.g. $A_1 \otimes B_1 \dots Z_1$. By assigning rows to local parties and columns to copies, this is represented by placing the operators in rows and columns correspondingly (left hand side). This representation can be understood as a tensor network notation of the expectation value to be evaluated by index contraction (right hand side).

noisy n -partite GHZ state with local dimension d reads

$$\text{tr}(\mathcal{W}\varrho^{\otimes k}) = \frac{(d)_{(k)}}{k!^{n-1}} \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n}\right)^r \left(\frac{p}{d}\right)^{k-r} (d+1-k)^{(r)n-2} \Delta_r \quad (2.2.17)$$

with

$$\Delta_r = d^r - \varsigma_n(k-r)!(d+1-k)^{(r)}, \quad (2.2.18)$$

where $(a)_{(b)} = a(a-1)\dots(a+1-b)$ is the falling factorial and $(a)^{(b)} = a(a+1)\dots(a+b-1)$ is the rising factorial, and we define $\varsigma_n = (1 + (-1)^n)/2$.

Proof. First note that we can split

$$\text{tr}(\mathcal{W}\varrho^{\otimes k}) = T_1 - k!T_2 \quad (2.2.19)$$

with

$$T_1 = \text{tr}(\mathbb{1} \otimes P_{1^k}^{\otimes n-1} \varrho^{\otimes k}), \quad T_2 = \text{tr}(P_{1^k}^{\otimes n} \varrho^{\otimes k}). \quad (2.2.20)$$

Let us first compute T_1 . To this end, it will be convenient to use the notation of Fig. 2.2.1 for tensor products mixing copies and subsystems. We arrange operators at different positions such that elements on the same row act on different *copies* of a local party, e.g. A_1 and A_2 ; and operators in same column act on different *subsystems* of the same copy, e.g. A_1 and B_1 . This is illustrated in 2.2.3. For example,

$$\begin{aligned} \text{tr}\left(\begin{pmatrix} X \\ Y \end{pmatrix} \cdot \begin{pmatrix} M^{\otimes 2} \\ N^{\otimes 2} \end{pmatrix}\right) &= \text{tr}\left(\begin{pmatrix} X \\ Y \end{pmatrix} \cdot \begin{pmatrix} M \\ N \end{pmatrix}^{\otimes 2}\right) \\ &:= \text{tr}((X_{A_1 A_2} \otimes Y_{B_1 B_2}) \cdot (M_{A_1} \otimes M_{A_2} \otimes N_{B_1} \otimes N_{B_2})). \end{aligned} \quad (2.2.21)$$

With this notation we write T_1 as

$$T_1 = \text{tr} \left(\underbrace{\begin{pmatrix} \mathbb{1}_d^{\otimes k} \\ P_{1^k} \\ \vdots \\ P_{1^k} \end{pmatrix}}_{(1)} \cdot \underbrace{\begin{pmatrix} \mathbb{1}_d & & |t\rangle\langle s| \\ \frac{(1-p)}{d^n} \mathbb{1}_d & & |t\rangle\langle s| \\ \vdots & & \vdots \\ \mathbb{1}_d & & |t\rangle\langle s| \end{pmatrix}}_{(2)}^{\otimes k} \right). \quad (2.2.22)$$

Note when expanding above expression, each term from (2) contains r copies of identity with a factor of $((1-p)/d^n)^r$, and $k-r$ copies of the (unnormalized) GHZ state with a factor of $(p/d)^{k-r}$. Since both $\mathbb{1}_d^{\otimes k} \otimes P_{1^k}^{\otimes n-1}$ in (1) are symmetric under permutation of the copies (i.e. permutations of the *columns* in Fig. 2.2.3), we obtain the standard binomial expression,

$$T_1 = \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \sum_{\vec{t}, \vec{s} \in \mathbb{Z}_d^{k-r}} \text{tr} \begin{pmatrix} \mathbb{1}_d^{\otimes k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle\langle\vec{s}| \\ P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle\langle\vec{s}| \\ \vdots \\ P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle\langle\vec{s}| \end{pmatrix}. \quad (2.2.23)$$

Define

$$C_r = d^r \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \quad (2.2.24)$$

and note that the first row in Eq. (2.2.23) reads

$$\text{tr} \left(\mathbb{1}_d^{\otimes k} \cdot (\mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle\langle\vec{s}|) \right) = \begin{cases} d^r & \text{if } \vec{t} = \vec{s} \\ 0 & \text{if } \vec{t} \neq \vec{s} \end{cases}. \quad (2.2.25)$$

Then one has

$$\begin{aligned} T_1 &= \sum_{r=0}^k C_r \sum_{\vec{t} \in \mathbb{Z}_d^{k-r}} \text{tr} \left(P_{1^k} \cdot (\mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle\langle\vec{t}|) \right)^{n-1} \\ &= \sum_{r=0}^k C_r \sum_{\vec{t} \in \mathbb{Z}_d^{k-r}} \text{tr} \left(\text{tr}_{1\dots r}(P_{1^k}) \cdot |\vec{t}\rangle\langle\vec{t}| \right)^{n-1}. \end{aligned} \quad (2.2.26)$$

The partial trace of P_{1^k} with respect to any arbitrary choice of r subsystems is known to be [141, Proposition 4]

$$\text{tr}_{1\dots r}(P_{1^k}) = \prod_{j=0}^{r-1} (d+1-k+j) \frac{(k-r)!}{k!} P_{1^{k-r}}, \quad (2.2.27)$$

where $P_{1^{k-r}}$ is the antisymmetric projector acting on $k-r$ subsystems. Therefore we have

$$T_1 = \sum_{r=0}^k C_r \prod_{j=0}^{r-1} (d+1-k+j)^{n-1} \left(\frac{(k-r)!}{k!} \right)^{n-1} \sum_{\vec{t} \in \mathbb{Z}_d^{k-r}} \text{tr} (P_{1^{k-r}} |\vec{t}\rangle\langle\vec{t}|)^{n-1}. \quad (2.2.28)$$

We will calculate separately the factor $\sum_{\vec{t} \in \mathbb{Z}_d^{k-r}} \text{tr} (P_{1^{k-r}} |\vec{t}\rangle\langle\vec{t}|)^{n-1}$. Due to the properties of the antisymmetric subspace, one has $P_{1^{k-r}} |\vec{t}\rangle = 0$ unless all components of $\vec{t} = (t_1, t_2, \dots, t_{k-r})$ are

different. Therefore, $P_{1^{k-r}} |\vec{t}\rangle$ is nonzero for

$$d(d-1)(d-2)\cdots(d-(k-r+1)) = \prod_{j=1}^{k-r} (d+1-j) \quad (2.2.29)$$

choices of $\vec{t}$. In case it is non zero, we have

$$\langle \vec{t} | P_{1^{k-r}} | \vec{t} \rangle = \frac{1}{(k-r)!} \sum_{\sigma \in S_{k-r}} (-1)^{\text{sign}(\sigma)} \langle \vec{t} | \sigma | \vec{t} \rangle = \frac{1}{(k-r)!}, \quad (2.2.30)$$

where $\langle \vec{t} | \sigma | \vec{t} \rangle$ vanishes if $\sigma \neq \text{id}$. Therefore, we have

$$\sum_{\vec{t} \in \mathbb{Z}_d^{k-r}} \text{tr} (P_{1^{k-r}} |\vec{t}\rangle \langle \vec{t}|)^{n-1} = \frac{\prod_{j=1}^{k-r} (d+1-j)}{(k-r)!^{n-1}}. \quad (2.2.31)$$

With this we are left with

$$T_1 = \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} d^r \prod_{i=0}^{r-1} (d+1-k+i)^{n-1} \frac{(k-r)!^{n-1}}{k!^{n-1}} \frac{\prod_{j=1}^{k-r} (d+1-j)}{(k-r)!^{n-1}}. \quad (2.2.32)$$

By shifting the indices, we have

$$\prod_{i=0}^{r-1} (d+1-k+i) \prod_{j=1}^{k-r} (d+1-j) = \prod_{j=1}^k (d+1-j) \quad (2.2.33)$$

and therefore T_1 reads

$$\begin{aligned} T_1 &= \frac{\prod_{i=1}^k (d+1-i)}{k!^{n-1}} \sum_{r=0}^k d^r \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \prod_{j=0}^{r-1} (d+1-k+j)^{n-2} \\ &= \frac{(d)_{(k)}}{k!^{n-1}} \sum_{r=0}^k d^r \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} (d+1-k)^{(r)}^{n-2}. \end{aligned} \quad (2.2.34)$$

Now we want to compute T_2 . Following a similar procedure as the one described above, we have

$$\begin{aligned} T_2 &= \text{tr} \left(P_{1^k} \left(\frac{(1-p)}{d^n} \mathbb{1}_d \begin{pmatrix} |t\rangle \langle s| \\ \vdots \end{pmatrix} + \frac{p}{d} \sum_{t,s=0}^{d-1} \begin{pmatrix} |t\rangle \langle s| \\ \vdots \end{pmatrix} \right)^{\otimes k} \right) \\ &= \sum_{r=0}^k \frac{C_r}{d^r} \sum_{\vec{t}, \vec{s} \in \mathbb{Z}_d^{k-r}} \text{tr} \begin{pmatrix} P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle \langle \vec{s}| \\ \vdots \\ P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle \langle \vec{s}| \end{pmatrix} \\ &= \sum_{r=0}^k \frac{C_r}{d^r} \sum_{\vec{t}, \vec{s} \in \mathbb{Z}_d^{k-r}} \text{tr} (P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle \langle \vec{s}|). \end{aligned} \quad (2.2.35)$$

By noting that

$$\text{tr} (P_{1^k} \cdot \mathbb{1}_d^{\otimes r} \otimes |\vec{t}\rangle \langle \vec{s}|) = \text{tr}(\text{tr}_{1\dots r}(P_{1^k}) |\vec{t}\rangle \langle \vec{s}|), \quad (2.2.36)$$

we have

$$T_2 = \sum_{r=0}^k \frac{C_r}{d^r} \prod_{j=0}^{r-1} (d+1-k+j)^n \frac{(k-r)!^n}{k!^n} \sum_{\vec{t}, \vec{s} \in \mathbb{Z}_d^{k-r}} \text{tr} (P_{1^{k-r}} |\vec{t}\rangle \langle \vec{s}|)^n. \quad (2.2.37)$$

For each pair of vectors $\vec{t}$ and $\vec{s}$, the factor $\text{tr} (P_{1^{k-r}} |\vec{t}\rangle \langle \vec{s}|)$ reads

$$\text{tr} (P_{1^{k-r}} |\vec{t}\rangle \langle \vec{s}|) = \frac{1}{(k-r)!} \sum_{\sigma \in S_{k-r}} (-1)^{\text{sign}(\sigma)} \langle \vec{s} | \sigma | \vec{t} \rangle. \quad (2.2.38)$$

where σ permutes the indices of the vectors $\vec{s} \in \mathbb{Z}_d^{k-r}$. Recall that the components of $\vec{s}$ are pairwise different, and the same holds for $\vec{t}$. Consequently, Eq. (2.2.38) is equal to $(-1)^{\text{sign}(\sigma)} (k-r)!^{-1}$, where $\sigma \in S_{k-r}$ is the unique permutation such that $\sigma | \vec{t} \rangle = | \vec{s} \rangle$.

Thus to evaluate Eq. (2.2.37), we need to sum Eq. (2.2.38) over $\{\vec{t} \in \mathbb{Z}_d^{k-r}\}$ with pairwise different components and over all the permutations of each vector $\vec{t}$ of length $k-r$. Those two sets have cardinalities $\prod_{j=1}^{k-r} (d+1-j)$ and $(k-r)!$. Furthermore, note that $\sum_{\sigma \in S_{k-r}} ((-1)^{\text{sign}(\sigma)})^n$ equals $(k-r)!$ if n is even, and 0 if n is odd. Those observations together lead to

$$\sum_{\vec{t}, \vec{s} \in S_{k-r}} \text{tr} (P_{1^{k-r}} |\vec{t}\rangle \langle \vec{s}|)^n = \frac{\sum_{\vec{t} \in \mathbb{Z}_d^{k-r}, \sigma \in S_{k-r}} ((-1)^{\text{sign}(\sigma)})^n}{(k-r)!^n} = \varsigma_n \frac{\prod_{j=1}^{k-r} (d+1-j)}{(k-r)!^{n-1}}, \quad (2.2.39)$$

where the function $\varsigma_n = (1 + (-1)^n)/2$ vanishes if n is odd. This allows us to compute T_2 ,

$$\begin{aligned} T_2 &= \varsigma_n \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \prod_{i=0}^{r-1} (d+1-k+i)^n \prod_{j=1}^{k-r} (d+1-j) \frac{(k-r)!}{k!^n} \\ &= \varsigma_n \prod_{i=1}^k (d+1-i) \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \frac{(k-r)!}{k!^n} \prod_{j=0}^{r-1} (d+1-k+j)^{n-1} \\ &= \varsigma_n (d)_{(k)} \sum_{r=0}^k \binom{k}{r} \left(\frac{1-p}{d^n} \right)^r \left(\frac{p}{d} \right)^{k-r} \frac{(k-r)!}{k!^n} (d+1-k)^{(r)n-1}. \end{aligned} \quad (2.2.40)$$

□

We compare this result to the P3-PPT method proposed in [43], which consists of evaluating the LU-invariant expression

$$\zeta(\varrho) = \text{tr}(\varrho^{T_X^3}) - \text{tr}(\varrho^{T_X^2})^2 \quad (2.2.41)$$

for each set $X \subset \{1, \dots, n\}$ of local parties. By the symmetry of the GHZ and maximally mixed states, one observes that $\zeta(\varrho)$ does not depend on the choice of the subsystem X . One can evaluate the P3-PPT criterion shown in Fig. 2.2.2 by evaluating the second and third moments of the partial transpose,

$$\text{tr}(\varrho^{T_X^2}) = p^2 + \frac{2p(1-p)}{d^n} + \frac{(1-p)^2}{d^n} \quad (2.2.42)$$

and

$$\text{tr}(\varrho^{T_X^3}) = \frac{p^3}{d^2} + \frac{2p^2(1-p)}{d^n} + \frac{p(1-p)^2}{d^{2n}} + \frac{1-p}{d^n} \text{tr}(\varrho^{T_X^2}). \quad (2.2.43)$$

2.3 Trace polynomial witnesses

2.3.1 Witnesses from matrix inequalities

How can one find suitable nonlinear witnesses that are simple to work with experimentally? Here we focus on trace polynomials, due to the fact that these are experimentally accessible through the randomized measurement toolbox [126] (see Section 2.3.2). These arise in the following way: matrix inequalities such as the Hadamard inequality,

$$\prod_{i=1}^n A_{ii} \geq \det(A) \quad (2.3.1)$$

for any $n \times n$ positive semidefinite matrix A , have been a long-standing topic of investigation [142, 143]. In particular, the so-called *immanant inequalities* generalize Eq. (2.3.1) and provide a large supply of n -partite linear witnesses of the form [144]

$$W = \sum_{\sigma \in S_n} a_{\sigma} \eta_d(\sigma), \quad (2.3.2)$$

where $a_{\sigma} \in \mathbb{C}$ and $\eta_d(\sigma)$ is the representation of $\sigma \in S_n$ permuting the n tensor factors of $(\mathbb{C}^d)^{\otimes n}$,

$$\eta_d(\sigma) |i_1\rangle \otimes \dots \otimes |i_n\rangle = |i_{\sigma^{-1}(1)}\rangle \otimes \dots \otimes |i_{\sigma^{-1}(n)}\rangle. \quad (2.3.3)$$

How to transform matrix inequalities to witnesses in the symmetric group algebra over the complex numbers, $\mathbb{C}S_n = \{\sum_{\sigma \in S_n} a_{\sigma} \sigma \mid a_{\sigma} \in \mathbb{C}\}$, is sketched as follows.

We shortly sketch the connections established in Refs. [144–146], which rely on some notions of representation theory. We will use the notation above, namely denote by S_n the symmetric group permuting n elements. Irreducible representations of S_n are labeled by partitions $\lambda = (\lambda_1, \dots, \lambda_n)$ with integers $\lambda_1 \geq \dots \geq \lambda_n$ satisfying $\lambda_1 + \dots + \lambda_n = n$. The character associated to λ is then the trace of the representation λ , denoted $\chi_{\lambda}(\sigma) = \text{tr}(\lambda(\sigma))$.

Given an $n \times n$ matrix A , one can then define the immanant associated to λ as

$$\text{Imm}_{\lambda}(A) = \sum_{\sigma \in S_n} \chi_{\lambda}(\sigma) \prod_{i=1}^n A_{i\sigma(i)}. \quad (2.3.4)$$

For example, the symmetric and antisymmetric representations are labelled by $(n, 0, \dots, 0) = (n)$ and $(1, \dots, 1) = (1^n)$, respectively. Their associated characters are $\chi_{(n)} = 1$ and $\chi_{(1^n)}(\sigma) = \text{sign}(\sigma)$. The corresponding matrix functions are the permanent and the determinant,

$$\text{Imm}_{(n)}(A) = \text{per}(A) \quad \text{and} \quad (2.3.5)$$

$$\text{Imm}_{(1^n)}(A) = \det(A). \quad (2.3.6)$$

With this in mind, let $\eta_d(\sigma)$ be the representation of the symmetric group on $(\mathbb{C}^d)^{\otimes n}$ permuting its tensor factors as

$$\eta_d(\sigma) |v_1\rangle \otimes \dots \otimes |v_n\rangle = |v_{\sigma^{-1}(1)}\rangle \otimes \dots \otimes |v_{\sigma^{-1}(n)}\rangle. \quad (2.3.7)$$

Maassen and Kümmerer [144] established the following relation between inequalities for immanants and entanglement witnesses.

Proposition 1 ([144]). *If for all $A \geq 0$ it holds that*

$$a \text{Imm}_{\lambda}(A) - b \text{Imm}_{\mu}(A) \geq 0, \quad (2.3.8)$$

Name	Inequality	Witness
Hadamard	$\prod_i A_{ii} \geq \det(A)$	$\mathbb{1} - n!P_{1^n}$
Schur	$\frac{\text{Imm}_\lambda(A)}{\chi_\lambda} \geq \det(A)$	$\frac{P_\lambda}{\chi_\lambda^2} - P_{1^n}$
Hook	$\frac{\text{Imm}_t(A)}{\chi_t} \geq \frac{\text{Imm}_{t-1}(A)}{\chi_{t-1}}$	$\frac{P_t}{\chi_t^2} - \frac{P_{t-1}}{\chi_{t-1}^2}$
Marcus	$\text{per}(A) \geq \prod_i A_{ii}$	$n!P_n - \mathbb{1}$
Permanent	$\text{per}(A) \geq \frac{\text{Imm}_\lambda(A)}{\chi_\lambda}$	$P_n - \frac{P_\lambda}{\chi_\lambda^2}$

Table 2.1: Immanant inequalities and their corresponding witnesses. Here are listed some standard immanant inequalities and their corresponding entanglement witnesses. The symbol χ_λ denotes the character of the identity in the irreducible representation λ of S_n . We denote $\text{Imm}_t(A) := \text{Imm}_{(t, 1^{n-t})}(A)$, $\chi_t := \chi_{(t, 1^{n-t})}$ and $P_t := P_{(t, 1^{n-t})}$. The permanent dominance inequality is conjectured for large matrices [142].

then the operator

$$W = \sum_{\sigma \in S_n} \left(\frac{a}{\chi_\lambda(\sigma)} - \frac{b}{\chi_\mu(\sigma)} \right) \eta_d(\sigma) \quad (2.3.9)$$

has non-negative expectation value on the set of separable states.

Proof. Let $f(A)$ be a linear combination of permutations of entries of A which is nonnegative for $A \geq 0$, and consider A as a Gram matrix of vectors $|v_1\rangle, \dots, |v_n\rangle$. Then

$$\begin{aligned} 0 \leq f(A) &= \sum_{\sigma \in S_n} c_\sigma \prod_{i=1}^n A_{i\sigma^{-1}(i)} \\ &= \sum_{\sigma \in S_n} c_\sigma \prod_{i=1}^n \langle v_i | v_{\sigma^{-1}(i)} \rangle \\ &= \sum_{\sigma \in S_n} c_\sigma \text{tr} \left(\eta_d(\sigma) \bigotimes_{i=1}^n |v_i\rangle \langle v_i| \right). \end{aligned} \quad (2.3.10)$$

Proposition 1 follows by assigning $f(A) = a \text{Imm}_\lambda(A) - b \text{Imm}_\mu(A)$ and considering the convex combination $\varrho = \sum_j p_j \bigotimes_{i=1}^n |v_i^{(j)}\rangle \langle v_i^{(j)}|$, which can be seen as an unnormalized separable state. $\square$

For example, it holds that the permanent of a positive semidefinite matrix is larger than the determinant,

$$\text{per}(A) \geq \det(A) \quad \text{for all } A \geq 0. \quad (2.3.11)$$

From this inequality one obtains the witness $W = P_{(n)} - P_{1^n}$, where $P_{(n)}$ is the projector onto the fully symmetric subspace of n qudits and $P_{(1^n)}$ is the rank-1 projector onto the fully antisymmetric subspace. Well-known immanant inequalities [142] like Eq. (2.3.11) and their associated witnesses are listed in Table 2.1. For example, from Eq. (2.3.1) one obtains the n -qudit witness of Eq. (2.2.14).

The nonlinear usage of linear witnesses introduced in this work allows us to construct nonlinear witnesses of the form $\mathcal{W} = W_1 \otimes \dots \otimes W_n$, as in Eq. (2.2.2), from witnesses arising from immanant inequalities. By doing so, one obtains expressions with terms of the form (2.3.19), which are invariant under local unitaries and thus detect the whole orbit of a quantum state ϱ .

Beyond using linear witnesses in the nonlinear fashion introduced in this work, one can also go out of this scheme and use the same tools to design other n -qudit nonlinear witnesses by considering

immanant inequalities involving matrices of different sizes, products of matrices and combinations of standard immanant inequalities [142]. Below we sketch graphically some examples for the bipartite case. Given a partition λ of a natural number r , we denote $\tilde{P}_\lambda = r!P_\lambda/\chi_\lambda^2$ with $\chi_\lambda = \chi_\lambda(\text{id})$. The following witnesses can be constructed from different matrix inequalities.

Assumption	Matrix inequality	Witness expression
$A \geq 0, \quad B \geq 0$	$\text{Imm}_\lambda(A)/\chi_\lambda \geq \text{Imm}_\gamma(A)/\chi_\gamma$ $\text{Imm}_\mu(B)/\chi_\mu \geq \text{Imm}_\nu(B)/\chi_\nu$	$\text{tr} \left[\left(\begin{array}{ c } \hline \tilde{P}_\lambda \\ \hline \end{array} - \begin{array}{ c } \hline \tilde{P}_\gamma \\ \hline \end{array} \right) \left(\begin{array}{ c } \hline \varrho \\ \hline \end{array} \cdots \begin{array}{ c } \hline \varrho \\ \hline \end{array} \right) \right]$

(2.3.12)

$\begin{pmatrix} A_{11} & A_{12} \\ A_{12}^\dagger & A_{22} \end{pmatrix} \geq 0$	$\text{per}(A_{11}) \text{per}(A_{22}) \leq \text{per}(A)$	$\text{tr} \left[\left(\begin{array}{ c } \hline \tilde{P}_{2k} \\ \hline \end{array} - \begin{array}{ c } \hline \tilde{P}_k \\ \hline \end{array} \right) \left(\begin{array}{ c } \hline \varrho \\ \hline \end{array} \cdots \begin{array}{ c } \hline \varrho \\ \hline \end{array} \right) \right]$
---	--	--

(2.3.13)

$A \geq 0, \quad B \geq 0$	$\det(A \circ B) \geq \det(A) \prod_{i=1}^k B_{ii}$	$\text{tr} \left[\left(\begin{array}{ c } \hline P_{1^k} \\ \hline \end{array} - \begin{array}{ c } \hline P_{1^k} \\ \hline \end{array} \right) \left(\begin{array}{ c } \hline \varrho \\ \hline \end{array} \cdots \begin{array}{ c } \hline \varrho \\ \hline \end{array} \right) \right]$
----------------------------	---	---

(2.3.14)

Here some notation needs to be clarified. In Eq. (2.3.12) $P_\lambda, P_\mu, P_\gamma$ and P_ν are the k -partite Young projectors corresponding to partitions of k labelled by λ, μ, γ and ν respectively. In Eq. (2.3.13), P_{2k} is the $(2k)$ -partite symmetrizer with local dimension d , which acts across all copies and all parties ($2k$ local subspaces in total). In Eq. (2.3.14), P_{1^k} inside the large box is the antisymmetrizer of k parties with local dimension d^2 . That is, it reads

$$P_{1^k} = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{\text{sign}(\sigma)} \sigma_A \otimes \sigma_B, \quad (2.3.15)$$

where each σ_A permutes k copies of the party A (which has dimension d) and σ_B permutes k copies of the party B (which has dimension d).

2.3.2 Implementation with randomized measurements

By recycling linear witnesses or positive operators $O_i \in \mathbb{C}S_k$ arising from immanant inequalities and using them in the nonlinear way of Eq. (2.2.2), the resulting nonlinear witness $\mathcal{W} = O_1 \otimes \dots \otimes O_n$ has interesting features. Its expectation value $\langle \mathcal{W} \rangle_{\varrho^{\otimes k}}$ is an homogeneous polynomial of degree k in ϱ , acting on the copies of subsystems as *trace polynomials* [145, 147]. Since by the Schur-Weyl duality the action of the symmetric group S_k commutes with that of k -fold tensor products $X^{\otimes k}$, the expression $\langle \mathcal{W} \rangle_{\varrho^{\otimes k}}$ is local unitary invariant. Similarly to the recently introduced p_3 -PPT condition and other local unitary invariant quantities [42, 43], these can be measured experimentally using the randomized measurements approach, in which a sequence of local projective measurements in random bases is performed iteratively to estimate the expectation value of k -copy observables [126, 148].

Trace polynomial witnesses can be evaluated with classical post-processing by matrix contraction, thus avoiding storing tensor products of matrices. To see this, here we sketch how trace polynomial witnesses introduced in this work can be evaluated with randomized measurements, using the tools introduced in [42, 43, 126].

Let $\{W_i\}$ be n witnesses for k -qudit mixed states, and let $\mathcal{W} = \bigotimes_{i=1}^n W_i$ be a nonlinear witness constructed from the linear witnesses W_i . The expectation value $\text{tr}(\mathcal{W} \varrho^{\otimes k})$ can be approximated as follows. First one applies random local unitaries $U_1 \otimes \dots \otimes U_n$ to ϱ and measures in the computational basis, obtaining a classical outcome string $q_1, \dots, q_n$. By performing m such measurements with different local unitaries one obtains m outcome strings $q_1^{(r)}, \dots, q_n^{(r)}$ with $1 \leq r \leq m$, from which to

construct the classical shadows

$$\tilde{\varrho}^{(r)} = \bigotimes_{i=1}^n \tilde{\varrho}_i^{(r)} \quad \text{with} \quad (2.3.16)$$

$$\tilde{\varrho}_i^{(r)} = (d+1)U_i^{(r)\dagger} |q_i^{(r)}\rangle \langle q_i^{(r)}| U_i^{(r)} - \mathbb{1}_d, \quad (2.3.17)$$

so that $\varrho^{(r)}$ has the same expectation value as ϱ over the unitary ensemble and measurements [149]. By replacing each j 'th copy of ϱ by an operator $\tilde{\varrho}^{(r_j)}$, the expectation value $\text{tr}(\mathcal{W}\varrho^{\otimes k})$ can be approximated as

$$\tilde{\theta}(\varrho) = \frac{1}{k!} \binom{m}{k}^{-1} \sum_{r_1 \neq r_2 \dots \neq r_k} \text{tr} \left(\mathcal{W} \varrho^{(r_1)} \otimes \dots \otimes \varrho^{(r_k)} \right) \quad (2.3.18)$$

with an error which decays as $1/\sqrt{m}$ for large m (see [42, 43] and references therein).

Expression (2.3.18) can be evaluated without computing the tensor product of classical shadows, if $\mathcal{W} = W_1 \otimes \dots \otimes W_n$ is a trace polynomial witness constructed from k -partite linear witnesses $W_i = \sum_{\sigma \in S_k} a_\sigma \eta_d(\sigma)$ in the symmetric group algebra. To see this, note that in this case Eq. (2.3.18) decomposes as $\tilde{\theta}(\varrho) = \sum_s c_s \tilde{\theta}_s(\varrho)$ in terms of the form

$$\tilde{\theta}_s(\varrho) = \text{tr} \left(\bigotimes_{i=1}^n \eta_d(\sigma_i) \cdot \varrho^{(r_1)} \otimes \dots \otimes \varrho^{(r_k)} \right), \quad (2.3.19)$$

where $s = (\sigma_1, \dots, \sigma_n)$. By recalling that each operator $\eta_d(\sigma_i)$ with $\sigma_i \in S_k$ permutes k tensor factors in the party i and that each shadow $\varrho^{(r_j)}$ is a product state (2.3.16), one can now argue analogously as in [42]: for any permutation $\sigma = (\alpha_1 \dots \alpha_l) \dots (\xi_1 \dots \xi_t) \in \mathbb{C}S_k$ with cycles denoted by brackets and given k matrices $X_1, \dots, X_k$, it holds that

$$\text{tr}(\eta_d(\sigma) X_1 \otimes \dots \otimes X_k) = \text{tr}(X_1 \dots X_l) \dots \text{tr}(X_{k-t} \dots X_k). \quad (2.3.20)$$

Therefore, by assigning $X_j = \varrho_i^{(r_j)}$, the quantity $\tilde{\theta}(\varrho)$ can be evaluated as a matrix contraction of the tensor factors (2.3.17) at each local party i , instead of computing a tensor product of k shadows.

2.3.3 Expectation value over Haar random states

Insight into the strenght of this approach can be made using Haar integration (Theorem 4.3 in [150]), which allows to compute the average expectation value of nonlinear trace polynomial witnesses with respect to Haar random states $|\psi\rangle \in \mathbb{C}^d$ with $d = d_1 \dots d_n$ over the Haar measure $\mathbf{d}_\mu$,

$$\mathbb{E}[\langle \mathcal{W} \rangle_\psi] = \int_{U \in \mathbf{U}(d)} \text{tr} \left(\mathcal{W} U |\psi\rangle \langle \psi| U^\dagger \right) \mathbf{d}_\mu(U). \quad (2.3.21)$$

For example, entanglement concentration of $|\psi\rangle$ using the witness $W = \mathbb{1} - k!P_{1^3}$ in Eq. (2.2.14) yields a negative value in average when $|\psi\rangle$ is shared among an even number of parties.

To see this, we will evaluate the average expectation value of the nonlinear witness $\mathcal{W} = (\mathbb{1} - k!P_{1^3}) \otimes P_{1^3}^{\otimes n-1}$ taken over k copies of Haar-random states shared among n inhomogeneous systems, $|\psi\rangle \in \mathbb{C}^d$ with $d = d_1 \dots d_n$. To have nonvanishing antisymmetrizers P_{1^k} acting on $\mathbb{C}^{d_i \otimes k}$, we will assume $k \leq d_{\min}$ where $d_{\min}$ is the minimal local dimension. The expression to evaluate is

$$\mathbb{E}[\langle \mathcal{W} \rangle] = \int_{U \in \mathbf{U}(d)} \text{tr} \left(\mathcal{W} (U |\psi\rangle \langle \psi| U^\dagger)^{\otimes k} \mathbf{d}_\mu(U) \right). \quad (2.3.22)$$

Due to Theorem 4.3 in [150], it holds that

$$\int_{U \in \mathbf{U}(d)} (U |\psi\rangle \langle \psi| U^\dagger)^{\otimes k} \mathbf{d}_\mu(U) = \binom{d+k-1}{k}^{-1} \frac{1}{k!} \sum_{\sigma_d \in S_k} \sigma_d, \quad (2.3.23)$$

where each permutation $\sigma_d = \sigma_{d_1} \otimes \cdots \otimes \sigma_{d_n}$ permutes k tensor factors of dimension $d = d_1 \cdots d_n$. Therefore we have

$$\mathbb{E}[\langle \mathcal{W} \rangle] \binom{d+k-1}{k} = \frac{1}{k!} \sum_{\sigma \in S_k} \text{tr}(\sigma_{d_1}) \prod_{i=2}^n \text{tr}(P_{1^k} \sigma_{d_i}) - k! \prod_{i=1}^n \text{tr}(P_{1^k} \sigma_{d_i}). \quad (2.3.24)$$

Now we need to make use of the following property of the antisymmetrizer,

$$P_{1^k} \sigma = (-1)^{\text{sign}(\sigma)} P_{1^k}. \quad (2.3.25)$$

To see this, first note that (1) the product of two permutations $\sigma, \pi \in S_k$ is again a permutation $\tau \in S_k$ with $\text{sign}(\tau) = \text{sign}(\sigma) \text{sign}(\pi)$, and (2) given $\sigma, \pi, \pi' \in S_k$ with $\pi \neq \pi'$, one has $\sigma\pi \neq \sigma\pi'$. Now we can insert $1 = (-1)^{\text{sign}(\sigma)} (-1)^{\text{sign}(\sigma)}$ to Eq. (2.3.25) and obtain

$$\sigma P_{1^k} = \frac{1}{k!} \sum_{\pi \in S_k} (-1)^{\text{sign}(\pi)} (-1)^{\text{sign}(\sigma)} (-1)^{\text{sign}(\sigma)} \sigma \pi \quad (2.3.26)$$

$$= \frac{1}{k!} (-1)^{\text{sign}(\sigma)} \sum_{\tau \in S_k} (-1)^{\text{sign}(\tau)} \tau \quad (2.3.27)$$

$$= (-1)^{\text{sign}(\sigma)} P_{1^k}. \quad (2.3.28)$$

With this one can reduce Eq (2.3.24) to

$$\mathbb{E}[\langle \mathcal{W} \rangle] \binom{d+k-1}{k} = \frac{1}{k!} \sum_{\sigma \in S_k} \text{tr}(\sigma) \text{tr}(P_{1^k})^{n-1} \left((-1)^{\text{sign}(\sigma)} \right)^{n-1} - k! \sum_{\sigma \in S_k} \text{tr}(P_{1^k})^n \left((-1)^{\text{sign}(\sigma)} \right)^n. \quad (2.3.29)$$

Now we need to distinguish between n being even and odd. If n is even, then we have

$$\mathbb{E}[\langle \mathcal{W} \rangle] \propto \frac{1}{k!} \text{tr}(P_{1^k})^{n-1} \left(\sum_{\sigma \in S_k} \text{tr}(\sigma) (-1)^{\text{sign}(\sigma)} - k!^2 \text{tr}(P_{1^k}) \right) \quad (2.3.30)$$

$$= \text{tr}(P_{1^k})^n (1 - k!) < 0, \quad (2.3.31)$$

since by definition $\sum_{\sigma \in S_k} \text{tr}(\sigma) (-1)^{\text{sign}(\sigma)} = k! \text{tr}(P_{1^k})$ holds. If n is odd, one can similarly show that $\mathbb{E}[\langle \mathcal{W} \rangle] = 0$.

2.3.4 Adapting witnesses to spectra

This approach allows to design separability criteria based on the spectra of the state in hand and its reductions. The question of what states are better detected can be addressed using the fact that the global and local spectra of quantum states influence the inner product between Young projectors and tensor copies of the state in hand [151, 152]. As an example, the following inequality involving a bipartite quantum state ϱ_{AB} and two local projectors P and Q [152],

$$\text{tr}((P \otimes Q) \varrho_{AB}^{\otimes k}) \geq \text{tr}(P \varrho_A^{\otimes k}) + \text{tr}(Q \varrho_B^{\otimes k}) - 1, \quad (2.3.32)$$

suggests that states with reduction ϱ_A maximizing the overlap $\text{tr}(P_{1^k} \varrho_A^{\otimes k})$, also maximize the negative term in Eq. (2.3.14). In this direction, nonlinear witnesses from trace polynomials give rise to nonlinear separability criteria like the following one.

Observation 3. Let ϱ be a bipartite state with local dimension d such that $\text{rank}(\varrho) < d$. The state ϱ is entangled if it has at least one maximally mixed local reduction $\varrho_A \propto \mathbb{1}$.

Proof. Let us start from the witness of Eq. (2.3.14),

$$\mathcal{W} = P_{1^k}^{(d^2)} - P_{1^k}^{(d)} \otimes \mathbb{1}_d^{\otimes k}. \quad (2.3.33)$$

Here $P_{1^k}^{(d^2)}$ is the antisymmetrizer acting on k copies of the full bipartite Hilbert space $\mathbb{C}^d \otimes \mathbb{C}^d$; $P_{1^k}^{(d)}$ is the antisymmetrizer acting on k copies of the Hilbert space $\mathbb{C}^d$ of the party A ; and $\mathbb{1}_d^{\otimes k}$ is the identity operator acting on k copies of the Hilbert space $\mathbb{C}^d$ of the party B . We will restrict ourselves to $k \leq d$ copies, as otherwise $P_{1^k}^{(d)} = 0$ and therefore $\mathcal{W}$ is positive semidefinite.

Now let ϱ have spectrum $\{r_1, \dots, r_{d^2}\}$ ordered by $r_1 \geq r_2 \geq \dots \geq r_{d^2}$. As a particular case of Theorem 1 in [152] we have the following bound,

$$\text{tr}(P_{1^k}^{(d^2)} \varrho^{\otimes k}) \leq (k+1)^{d^2(d^2-1)/2} k \prod_{i=1}^k r_i, \quad (2.3.34)$$

where $\{r_i\}$ are the largest k eigenvalues of ϱ . By assumption, $\text{rank } \varrho < d$ and consider a number of copies between $\text{rank } \varrho$ and d , $\text{rank } \varrho < k \leq d$. Then there exists at least a zero eigenvalue $r_{i^*} = 0$ with $i^* \leq k$ and the nonnegative term of $\langle \mathcal{W} \rangle_{\varrho^{\otimes k}}$ vanishes,

$$\text{tr}(P_{1^k}^{(d^2)} \varrho^{\otimes k}) = 0. \quad (2.3.35)$$

In this case, the expectation value of $\mathcal{W}$ is either 0 if $\text{tr}(P_{1^k}^d \varrho_A^{\otimes k}) = 0$, or strictly negative if $\text{tr}(P_{1^k}^d \varrho_A^{\otimes k}) > 0$. By Corollary 2 in [152], if ϱ_A is maximally mixed, then there exists a number of copies $k_0 \leq d$ such that for all $k \geq k_0$, its projection onto the antisymmetric subspace is nonzero,

$$\text{tr}(P_{1^k} \varrho_A^{\otimes k}) > 0. \quad (2.3.36)$$

Therefore the expectation value of $\mathcal{W}$ is strictly negative, which means that ϱ is entangled. $\square$

2.3.5 Sampling Werner states

The exponential growth of the Hilbert space with the number of subsystems limits our capability to numerically evaluate LU-invariants, which requires using copies of the state in hand. We can surpass this problem using trace polynomial witnesses $W \in \mathbb{C}S_k$ for multipartite Werner states $\varrho \in \mathbb{C}S_n$, which entanglement properties are an ongoing focus of research [153–155]. Their symmetry allows us to sample states and witnesses of nontrivial sizes with a symbolic algebra package, which is computationally cheaper than computing the expectation value of exponentially large matrices.

Sampling states and witnesses symbolically is based on group ring states and can be done by following the lines of Ref. [121]. Let $\sigma \in S_n$ and let $\eta_d(\sigma)$ be its representation permuting the n tensor factors of $(\mathbb{C}^d)^{\otimes n}$,

$$\eta_d(\sigma) |i_1\rangle \otimes \dots \otimes |i_n\rangle = |i_{\sigma^{-1}(1)}\rangle \otimes \dots \otimes |i_{\sigma^{-1}(n)}\rangle. \quad (2.3.37)$$

For each element $a = \sum_{\sigma \in S_n} \alpha_\sigma \sigma \in \mathbb{C}S_n$, we denote its representation in the Hilbert space as $\eta_d(a)$. For each partition of n labeled by λ , define the centrally primitive idempotent

$$\omega_\lambda = \frac{\chi_\lambda(\text{id})}{n!} \sum_{\sigma \in S_n} \chi_\lambda(\sigma) \sigma^{-1} \quad (2.3.38)$$

and its representation in the Hilbert space $P_\lambda = \eta_d(\omega_\lambda)$, where $\chi_\lambda(\text{id})$ is the character of the

identity in the irreducible representation labeled by λ . For example, P_{1^n} is the projector onto the antisymmetric subspace of $(\mathbb{C}^d)^{\otimes n}$.

We now define a trace τ on the elements of S_n as

$$\tau(\sigma) = \begin{cases} n! & \text{if } \sigma = \text{id} \\ 0 & \text{else} \end{cases}, \quad (2.3.39)$$

and the Weingarten operator $\text{Wg}(d, n) = \eta_d(\text{wg}(d, n))$ where [156, 157]

$$\text{wg}(d, n) = \frac{1}{n!} \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \frac{\tau(\omega_\lambda)}{\text{tr}(P_\lambda)} \omega_\lambda. \quad (2.3.40)$$

We denote $J_d = \ker(\eta_d)^\perp$ the subspace of $\mathbb{C}S_n$ with nonzero representation η_d acting on $(\mathbb{C}^d)^{\otimes n}$,

$$J_d = \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \omega_\lambda \mathbb{C}S_n, \quad (2.3.41)$$

which is spanned by ω_λ such that the height of the Young tableau λ is not larger than d . Now we are ready for the following Lemma.

Lemma 2 (Lemma 1 in [121]).

(1) For all $r \in J_d$ and $b \in \mathbb{C}S_n$ it holds that

$$n! \text{tr}(\text{Wg}(d, n) \eta_d(r) \eta_d(b)) = \tau(rb). \quad (2.3.42)$$

(2) Let $r \in J_d$. Then r is a state if and only if $n! \text{Wg}(d, n) \eta_d(r)$ is a state in $\mathcal{B}((\mathbb{C}^d)^{\otimes n})$.

The first part of Lemma 2 allows us to compute inner products of operators which are representations of $\mathbb{C}S_n$ in the group ring. In order to evaluate expressions of the form (2.2.2), one needs to compute tensor products of states and witnesses. For that, we will need the following variation of the first part of Lemma 2.

Lemma 3. For all $a \in J_d$ and $b \in \mathbb{C}S_n$ it holds that

$$\text{tr}(\eta_d(a) \eta_d(b)) = \frac{\tau(\text{wg}(d, n)^{-1} ab)}{n!}. \quad (2.3.43)$$

Proof. First note that the inverse of the Weingarten operator in $\mathbb{C}S_n$ can be written as

$$\text{wg}(d, n)^{-1} = n! \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \frac{\text{tr}(P_\lambda)}{\tau(\omega_\lambda)} \omega_\lambda \quad (2.3.44)$$

since the sum of any subset of isotypic components ω_λ is the identity in the subspace they have support on. Now let $r = \text{wg}(d, n)a$ and assume $a \in J_d$. Lemma 2 applies because $r \in J_d$ as well, and thus Eq. (2.3.43) follows from Eq. (2.3.42). $\square$

Now we are able to compute the expectation value of n tensor products of operators $W_i = \eta_d(\omega_i)$ with respect to k copies of a quantum state $\varrho = n! \text{Wg}(d, n) \eta_d(r)$, using that the representation η_d is linear. We have

$$\begin{aligned} \text{tr}(\varrho^{\otimes k} W_1 \otimes \dots \otimes W_n) &= \text{tr}(n!^k \text{Wg}(d, n)^{\otimes k} \eta_d(r)^{\otimes k} \eta_d(w_1) \otimes \dots \otimes \eta_d(w_n)) \\ &= n!^k \text{tr}(\eta_d(\text{wg}(d, n)^{\times k} r^{\times k}) \eta_d(w_1 \times \dots \times w_n)) \\ &= \frac{n!^k}{(kn)!} \tau(\text{wg}(d, kn)^{-1} \text{wg}(d, n)^{\times k} r^{\times k} w_1 \times \dots \times w_n). \end{aligned} \quad (2.3.45)$$

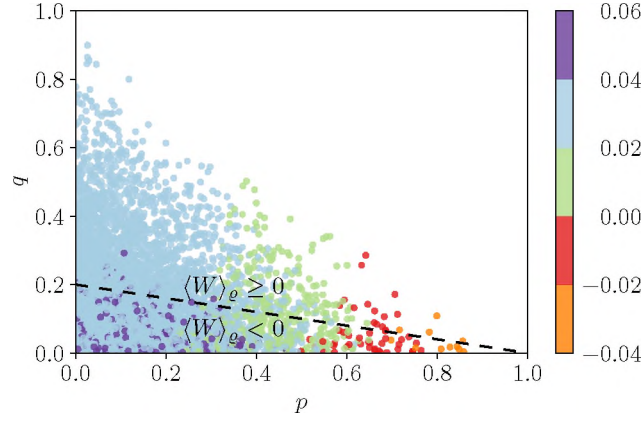


Figure 2.3.1: Detection of 3 qutrit random Werner states using 3 copies. We represent in colored dots a sample of 5000 random Werner states ϱ . The horizontal and vertical axes show their projections onto the antisymmetric and symmetric subspaces respectively, $p = \text{tr}(\varrho P_{111})$ and $q = \text{tr}(\varrho P_3)$, such that $\text{tr}(\varrho P_{21}) = 1 - p - q$. The color gradient represents the expectation value $\langle W \rangle_{\varrho^{\otimes 3}}$ with $W = W^{\otimes 2} \otimes P_{21}$ where $W = P_3 - \frac{P_{21}}{4}$. States below (above) the analytical dashed line are detected (not detected) linearly as $\langle W \rangle_{\varrho} < 0$ ($\langle W \rangle_{\varrho} \geq 0$). Therefore states depicted in red or orange above the dashed line are detected nonlinearly by W , but not linearly by W .

This gives us the following practical recipe to sample Werner states. First we fix an operator W of the form (2.2.2) which is a tensor product of witnesses and positive operators $W_i = \eta_d(w_i)$ with $w_i \in \mathbb{C}S_k$. Then we compute iteratively its expectation values with respect to k -fold copies of n -partite Werner states, according to Eq. (2.3.45). At each iteration, we take $s = \sum_{\sigma} s_{\sigma} \sigma \in \mathbb{C}S_n$ choosing the coefficients $s_{\sigma} \in \mathbb{C}$ randomly, and compute

$$r = \frac{ss^{\dagger}}{\tau(ss^{\dagger})} \quad (2.3.46)$$

with $s^{\dagger} = \sum_{\sigma} s_{\sigma}^* \sigma^{-1}$, such that r is a state in $\mathbb{C}S_n$ and therefore $\varrho = n! Wg(d, n) \eta(r)$ is a quantum state in the Hilbert space $(\mathbb{C}^d)^{\otimes n}$. Finally we compute the expectation value $\langle W \rangle_{\varrho^{\otimes k}}$ via Eq. (2.3.45).

This recipe allows us to compute in $\mathbb{C}S_n$ expectation values of the form (2.2.2) avoiding the exponential growth of the number of parameters with the local dimension d . Let us focus on the system size $k = n = d = 3$. While a desktop computer cannot compute Eq. (2.2.2) in the Hilbert space $(\mathbb{C}^3)^{\otimes 9}$, the fact that the nine-qutrit state factorizes as $\varrho^{\otimes 3}$ makes this computation possible in $(\mathbb{C}S_3)^{\times 3}$ symbolically. For example, in a random sample of 3 qutrit Werner states, there is a full-measure subset of states which are detected nonlinearly by $\langle W \rangle_{\varrho^{\otimes 3}} < 0$ where $W = W^{\otimes 2} \otimes P_{21}$ with $W = P_3 - \frac{P_{21}}{4}$. The detection of random states depends on their support onto the subspace defined by each irreducible representation, as shown in Fig. 2.3.1.

2.4 Evaluating positive maps

Using variations of equation (2.2.2) one can detect nonpositive outcomes of positive maps applied locally. For example one can evaluate the reduction criteria [45] via purity conditions [43] with Eqs. (2.2.14) and (2.2.16), and detect nonpositivity of the partial transpose Γ of a state ϱ via

$$\text{tr}((\phi^+ \otimes V)(\varrho \otimes \sigma)) = \text{tr}(\varrho^{\Gamma} \sigma) \quad (2.4.1)$$

if σ lies in a eigenspace of ϱ^{Γ} corresponding to a negative eigenvalue. In this sense, our method relates to the existence problem of nondecomposable tensor stable maps, which remain positive under tensor powers [158, 159], as follows: there exists a nondecomposable tensor stable map, if

and only if, there exists a bipartite witness W such that $W^{\otimes n}$ cannot detect any pair of n -partite states ϱ and σ . This is because a nondecomposable map Φ_W is tensor stable if and only if

$$\Phi_W^{\otimes n}(\varrho) = \text{tr}_1(W^{\otimes n}(\varrho^T \otimes \mathbb{1}_d)) \geq 0 \quad (2.4.2)$$

for any n -partite state ϱ , where W is a witness for states with positive partial transpose (See Chapter 11 of [12] and references therein); which is true if and only if the expectation value of $\Phi_W^{\otimes n}(\varrho)$ for any other n -partite state σ is nonnegative.

2.5 Conclusions

Our approach shows how entanglement witnesses can be recycled in non-linear fashion, so to succeed even when linear detection fails. In particular, this approach is suitable to randomized measurements [42, 149, 160], trace polynomials being naturally invariant under local unitaries.

Some open questions remain: How can one tailor trace polynomial witnesses to specific states? The Keyl-Werner theorem [151] and subsequent work [152] suggest that it is possible to choose witnesses according to the spectra of the reductions of the state in hand. While Observation 3 serves as a first step in this direction, a systematic approach is yet missing. Lastly, it would be interesting to make use of matrix inequalities involving elementary symmetric polynomials [142], which might be more powerful than the particular case of immanant inequalities.

2.6 Further results: positive maps from immanant inequalities

In recent years, matrix inequalities involving permutations have garnered increasing attention as a tool to detect quantum entanglement. Here we summarize the main results in this area and further strengthen this connection with a new family of positive maps able to detect entanglement, that generalizes and unifies known methods. These generalize known trace polynomial expressions, by introducing extra degrees of freedom. In the simplest case, we find a parametrized family generalizing the reduction map and explore its entanglement detection capabilities in a discretized sampling. In a sample of Ginibre random states, we find that almost all two-qutrit NPT states can be detected with this generalization. Our results suggest a method of enhancing entanglement detection with positive maps, by projecting a local subsystem onto a lower-dimensional subspace.

2.6.1 State of the art: an overview

Entanglement witnesses from matrix inequalities

Recently, a bridge has been found between immanant inequalities, originally studied in mathematical literature, and the detection of entanglement [13, 144, 146]. The key idea is that given a positive matrix G with entries $G_{ij} = \langle v_i | v_j \rangle$, Eq. (2.3.4) can be written as

$$\text{Imm}_\lambda(G) = \frac{n!}{\chi_\lambda(\text{id})} \text{tr}(P_\lambda |v_1\rangle\langle v_1| \otimes \cdots \otimes |v_n\rangle\langle v_n|). \quad (2.6.1)$$

Here

$$P_\lambda = \frac{\chi_\lambda(\text{id})}{n!} \sum_{\sigma \in S_n} \chi_\lambda(\sigma) \eta_d(\sigma^{-1}) \quad (2.6.2)$$

is the Young projector onto the irreducible subspace invariant under the representation λ of the symmetric group, where $\eta_d(\sigma)$ permutes n tensor factors of $(\mathbb{C}^d)^{\otimes n}$,

$$\eta_d(\sigma) |i_1\rangle \otimes \cdots \otimes |i_n\rangle = |i_{\sigma^{-1}(1)}\rangle \otimes \cdots \otimes |i_{\sigma^{-1}(n)}\rangle. \quad (2.6.3)$$

By convexity, it follows from 2.6.1 that for any immanant inequality $\sum_{\lambda} a_{\lambda} \text{Imm}_{\lambda}(X) \geq 0$, there exists associated an operator

$$W = \sum_{\lambda} \frac{n!}{\chi_{\lambda}(\text{id})} a_{\lambda} P_{\lambda} \quad (2.6.4)$$

that has a nonnegative expectation value on positive on all separable states, $\text{tr}(W \varrho_{SEP}) \geq 0$. That is, each immanant inequality can be mapped to an operator with nonnegative expectation value on product states (and separable states, by convexity and linearity).

These operators W are entanglement witnesses if they have some negative eigenvalue, $W \not\geq 0$. However, this is not always the case. For example, the inequality $\det(G) \geq 0$ on 2×2 positive matrices $G \geq 0$ leads to $\text{tr}(P_{1,1} \varrho \otimes \sigma) \geq 0$ for all ϱ and σ , where the antisymmetrizer $P_{1,1} = |\psi^{-}\rangle\langle\psi^{-}|$ with

$$|\psi^{-}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} \quad (2.6.5)$$

is positive semidefinite and thus cannot serve as a witness. Alternatively, from the inequality (2.3.1) one can derive that

$$\text{tr}(\eta_d(12) \varrho \otimes \sigma) \geq 0, \quad (2.6.6)$$

where the SWAP operator acting as $\eta_d(12) |v\rangle |w\rangle = |w\rangle |v\rangle$ is a witness detecting bipartite entanglement.

Positive maps from trace polynomials

Using trace polynomials, the following family of maps was introduced in [145, 161],

$$\Psi_{\lambda}(X_2, \dots, X_n) := \text{tr}_{1, \dots, n \setminus 1}(P_{\lambda}(\mathbb{1} \otimes X_2 \otimes \dots \otimes X_n)). \quad (2.6.7)$$

The maps of this form are multilinear, namely linear in each entry X_i . Using linear algebra and positivity of the Young projectors, $P_{\lambda} \geq 0$, one sees that these maps are completely positive $\Psi_{\lambda}(X_2, \dots, X_n) \geq 0$ on positive semidefinite matrices $X_i \geq 0$ [145].

Moreover, the positive maps $\Psi_{\lambda}(X_2, \dots, X_n)$ were shown in [145] to have several applications in quantum information theory. For example, this family of maps recovers well-known maps like the Cayley-Hamiltonian map [161] and the reduction map [45], through $\Psi_{(1,1)}(X) = \text{tr}(X)\mathbb{1} - X = R(X)$. Thus a straightforward application is the detection of entanglement, since the reduction map is known to be useful for this purpose when applied to one subsystem of a bipartite state. That is, if the following expression is not positive semidefinite,

$$R \otimes \text{id}(\varrho_{AB}) \not\geq 0, \quad (2.6.8)$$

then ϱ_{AB} is entangled.

A further step was done in [146]. It was shown that for any immanant inequality $\sum_{\lambda} a_{\lambda} \text{Imm}_{\lambda}(G) \geq 0$ described by a vector $\vec{a}$, there exists an associated multilinear map

$$\Psi_{\vec{a}}(X_2, \dots, X_n) := \sum_{\lambda} a_{\lambda} \text{tr}_{\{1, \dots, n\} \setminus 1}(P_{\lambda}(\mathbb{1} \otimes X_2 \otimes \dots \otimes X_n)) \quad (2.6.9)$$

which is positive $\Psi_{\vec{a}}(X_2, \dots, X_n) \geq 0$ on positive matrices $X_2, \dots, X_n \geq 0$.

2.6.2 A new family of positive maps

The general case

Here we extend the family of positive maps introduced in [145, 146] by introducing new degrees of freedom given by an additional positive semidefinite matrix X_1 .

Proposition 2. *Let*

$$\sum_{\lambda \vdash n} a_\lambda \text{Imm}_\lambda(G) \geq 0 \quad (2.6.10)$$

be an immanant inequality with coefficients $a_\lambda \in \mathbb{R}$ for any $n \times n$ positive semidefinite matrix G , and let $\{X_1, X_2, \dots, X_n\}$ be an ordered set of n positive semidefinite matrices of size $d \times d$. The following map acting on n positive matrices,

$$\begin{aligned} \Psi_a^{X_1}(X_2, \dots, X_n) := \\ \sum_{\lambda} a_\lambda \frac{n!}{\chi_\lambda(\text{id})} \text{tr}_{\{1, \dots, n\} \setminus 1} (P_\lambda X_1^{\otimes n} (\mathbb{1} \otimes X_2 \otimes \dots \otimes X_n)) \end{aligned} \quad (2.6.11)$$

is positive, $\Psi_a^{X_1}(X_2, \dots, X_n) \geq 0$.

This result is an extension of the derivations with trace polynomials in [145, 146], which are recovered when $X_1 = \mathbb{1}$.

Proof. We will distinguish the first positive matrix, as $X_1 := X$. First of all, let us expand Eq. (2.6.11) into

$$\begin{aligned} \Psi_a^X(X_2, \dots, X_n) = \sum_{\lambda} a_\lambda \sum_{\sigma \in S_n} \chi_\lambda(\sigma) X \prod_{r=2}^R \left(X_{\alpha_r^{(\sigma)}} X \right) \\ \text{tr} \left[\prod_{s=1}^S \left(X_{\beta_s^{(\sigma)}} X \right) \right] \\ \cdots \text{tr} \left[\prod_{t=1}^T \left(X_{\xi_t^{(\sigma)}} X \right) \right], \end{aligned} \quad (2.6.12)$$

where each permutation σ is written in its cycle form $\sigma = (\alpha)(\beta)\dots(\xi) = (\alpha_1 \dots \alpha_R)(\beta_1 \dots \beta_S)\dots(\xi_1 \dots \xi_T)$ with the convention $\alpha_1 = 1$.

Let us now define the inner product $\langle u, w \rangle = \langle u | X | w \rangle$ with $|u\rangle, |w\rangle \in \mathbb{C}^d$ and write G as a Gram matrix,

$$G = \begin{pmatrix} \langle i_1 | X | i_1 \rangle & \cdots & \langle i_1 | X | i_n \rangle \\ \vdots & \ddots & \vdots \\ \langle i_n | X | i_1 \rangle & \cdots & \langle i_n | X | i_n \rangle \end{pmatrix}. \quad (2.6.13)$$

By expanding each permutation $\sigma \in S_n$ into cycles, $\sigma = (\alpha)\dots(\xi)$, a linear combination of Immanants Imm_λ which is positive on any matrix G reads

$$\begin{aligned} \sum_{\lambda} a_\lambda \text{Imm}_\lambda(G) &= \sum_{\lambda} a_\lambda \sum_{\sigma \in S_n} \chi_\lambda(\sigma) \prod_{l=1}^n \langle i_l | X | i_{\sigma(l)} \rangle = \\ &\sum_{\lambda} a_\lambda \sum_{\sigma \in S_n} \chi_\lambda(\sigma) \langle i_{\alpha_1^{(\sigma)}} | X | i_{\alpha_2^{(\sigma)}} \rangle \cdots \langle i_{\alpha_R^{(\sigma)}} | X | i_{\alpha_1^{(\sigma)}} \rangle \\ &\quad \cdots \\ &\quad \langle i_{\xi_1^{(\sigma)}} | X | i_{\xi_2^{(\sigma)}} \rangle \cdots \langle i_{\xi_T^{(\sigma)}} | X | i_{\xi_1^{(\sigma)}} \rangle \geq 0. \end{aligned} \quad (2.6.14)$$

Let us now take a convex combination with nonnegative coefficients $x_i \geq 0$ and identify the positive semidefinite $d \times d$ matrices $\{X_{\alpha_v^{(\sigma)}}\} = \left\{ \sum_{\alpha_r^{(\sigma)}=1}^d x_{i_{\alpha_r^{(\sigma)}}} |i_{\alpha_r^{(\sigma)}}\rangle \langle i_{\alpha_r^{(\sigma)}}| \right\}$ for the first cycle α of the permutation σ , and similarly for the rest of cycles $\beta, \dots, \xi$ of σ .

By assumption, recall that $\alpha_1^{(\sigma)} = 1$ in all terms of the summation. Therefore we can factorize $\langle i_{\alpha_1^{(\sigma)}} |$ to the left hand side and $|i_{\alpha_1^{(\sigma)}}\rangle$ to the right hand side, and take a linear combination of the

whole expression in order to have it in terms of $\{X_{\alpha_r^{(\sigma)}}\}, \dots, \{X_{\xi_t^{(\sigma)}}\}$. It follows that

$$\langle i_{\alpha_1^{(\sigma)}} | \Psi_d^X(X_2, \dots, X_n) | i_{\alpha_1^{(\sigma)}} \rangle$$

is nonnegative for any vector $|i_{\alpha_1^{(\sigma)}}\rangle \in \mathbb{C}^d$ and hence the multilinear map Ψ_d^X is positive, which proves the claim. $\square$

Some examples

We will give some examples considering the inequality $\det(X) \geq 0$, which does not provide a witness using the approach of [144] and [14]. Using Proposition 2 for $n = 2$ in the expanded form of Eq. (2.6.12), we obtain that the map

$$\begin{aligned} \Psi_{(0,1)}^X(Y) &= \text{tr}_2[|\psi^-\rangle\langle\psi^-| (X \otimes X)(\mathbb{1} \otimes Y)] \\ &= X \text{tr}[YX] - XYX \geq 0 \end{aligned} \quad (2.6.15)$$

linear in Y is positive for any positive semidefinite matrices $X, Y \geq 0$. By setting $X = \mathbb{1}$ we recover the reduction map, $\Psi_{\mathbb{1}}(Y) = \mathbb{1} \text{tr}[Y] - Y \geq 0$. Instead, by setting $Y = \mathbb{1}$, we have the so-called *purity inequality* $\Psi_{(0,1)}^X(\mathbb{1}) = X \text{tr}[X] - X^2 \geq 0$.

Let us now consider the case $n = 3$. From positivity of the determinant in 3×3 positive semidefinite matrices, we obtain the positive map

$$\begin{aligned} \Psi_{(0,1)}^X(Y, Z) &= X \text{tr}[YX] \text{tr}[ZX] + XYXZX + XZXYX \\ &\quad - XYX \text{tr}[ZX] - X \text{tr}[YXZX] \\ &\quad - XZX \text{tr}[YX] \geq 0 \end{aligned} \quad (2.6.16)$$

for any three positive semidefinite matrices $X, Y, Z \geq 0$, which is linear in Y and Z . By setting $Z = \mathbb{1}$, we have

$$\begin{aligned} \Psi_{(0,1)}^X(Y, \mathbb{1}) &= X \text{tr}[YX] \text{tr}[X] + XYX^2 + X^2YX \\ &\quad - XYX \text{tr}[X] - X \text{tr}[YX^2] - X^2 \text{tr}[YX] \geq 0. \end{aligned} \quad (2.6.17)$$

By construction, with $X = \mathbb{1}$ we recover the following map defined in [145, 161],

$$\begin{aligned} \Psi_{(0,1)}^{\mathbb{1}}(Y, Z) &= \mathbb{1} \text{tr}[Y] \text{tr}[Z] + YZ + ZY \\ &\quad - Y \text{tr}[Z] - \mathbb{1} \text{tr}[YZ] - Z \text{tr}[Y] \geq 0. \end{aligned} \quad (2.6.18)$$

If in equation (2.6.16) we choose $C = A = \mathbb{1}$ (or similarly $B = A = \mathbb{1}$), we are left with the reduction map up to a rescaling,

$$\Psi_{(0,1)}^{\mathbb{1}}(Y, \mathbb{1}) = (d-2)\mathbb{1} \text{tr}[Y] - (d-2)Y \geq 0. \quad (2.6.19)$$

2.6.3 Entanglement detection

Detecting states from positivity of the determinant

Now we will address the performance of Proposition 2 in detecting bipartite entanglement with a simple example. We shall refer to $\Psi_{(0,1)}^X(Y)$ in Eq. (2.6.15) as $\Psi_{(0,1)}^X$ when no argument Y is specified. Let us recall that $\Psi_{(0,1)}^X$ is a positive map, simply because the determinant is nonnegative on any positive semidefinite matrix. For entanglement detection purposes, let us consider the map

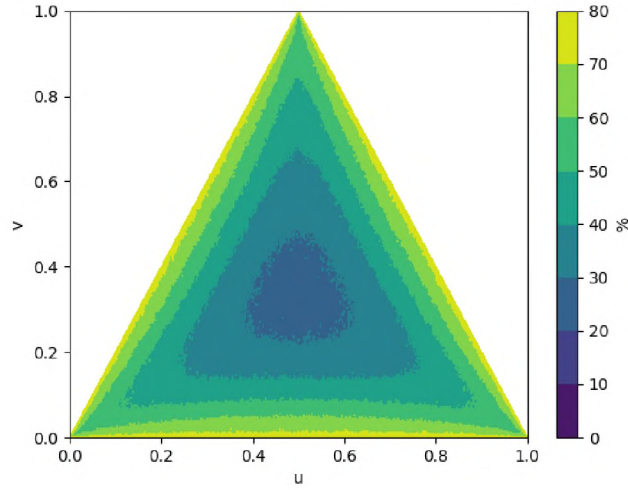


Figure 2.6.1: Detection rate of two-qutrit random states through the generalized reduction map $\Psi_{(0,1)}^X$ with $X = \text{diag}(a, b, c)$. The axes parametrize a , b and c as $u := (a - b + 1)/2$ and $v := c$ in order to plot the results in the probability simplex of size 3, and the color bar indicates the percentage of Ginibre random bipartite states detected through $\Psi_X \otimes \text{id}(\varrho_{AB}) \not\geq 0$ out of a sample of 2000 random states. At the baricenter of the triangle lays the case $X = \text{diag}(1, 1, 1) = \mathbb{1}$, which evaluates the reduction map and detects 27% of the random states approximately. The edges contain the case $X = \text{diag}(1, 1, 0)$, which evaluates the reduction map on a two-dimensional subspace and increases the detection rate to 74%.

$\Psi_{(0,1)}^X$ acting on one side of a bipartite state ϱ_{AB} :

$$\begin{aligned} \Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB}) &= X \otimes \text{tr}_A[(X \otimes \mathbb{1})\varrho] - \\ &\quad (X \otimes \mathbb{1})\varrho(X \otimes \mathbb{1}) \quad \text{and} \\ \text{id} \otimes \Psi_{(0,1)}^X(\varrho_{AB}) &= \text{tr}_B[(\mathbb{1} \otimes X)\varrho] \otimes X - \\ &\quad (\mathbb{1} \otimes X)\varrho(\mathbb{1} \otimes X). \end{aligned} \quad (2.6.20)$$

Recall that the map $\Psi_{(0,1)}^X$ is linear and positive but not completely positive, and therefore $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB})$ can detect entanglement. Without loss of generality, we can choose a basis diagonalizing X , and thus write

$$X = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad (2.6.21)$$

with $a, b, c \geq 0$ and $a + b + c = 1$ for normalization purposes. In Figure 2.6.1 we show the percentage of states with non-positive partial transpose detected by $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB}) \not\geq 0$ with $X = \text{diag}(a, b, c)$, where $a + b + c = 1$. The sample consists of 2000 Ginibre-random states for each choice of a , b and c discretized in a 250×250 lattice. We parameterize the coordinates as

$$\begin{aligned} u &:= \frac{a - b + 1}{2} \\ v &:= c. \end{aligned} \quad (2.6.22)$$

Method of detection	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$d = 6$
NPT	7566	9999	10000	10000	10000
$\Psi_{X_\sigma} \otimes \mathbb{1}$ or $\mathbb{1} \otimes \Psi_{X_\sigma}$	7566	9924	9991	9988	9869
$\Psi_{(0,1)}^X \otimes \mathbb{1}$ or $\mathbb{1} \otimes \Psi_{(0,1)}^X$	7566	9016	8259	6231	3172
$\Psi_{(0,1)}^X \otimes \mathbb{1}$	7566	7406	6078	3874	1732
$\Psi_{\mathbb{1}} \otimes \mathbb{1}$ (Reduction)	7566	2741	10	0	0

Table 2.2: Detection rate of the map $\Psi_{(0,1)}^X$ on bipartite states. A set of 10^4 Ginibre random bipartite states ϱ_{AB} with local dimension $d_A = d_B := d$ is sampled. The first row gives the number of states detected by the negative partial transpose (NPT), in accordance with [12]. Positivity of the map $\Psi_{(0,1)}^X$ is evaluated, choosing $X = \text{diag}(1, 1, 0, \dots, 0)$. We show the number of states detected by the reduction map (fifth row); by $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB}) \not\geq 0$ (fourth row); by either $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB}) \not\geq 0$ or $\text{id} \otimes \Psi_{(0,1)}^X(\varrho_{AB}) \not\geq 0$ (third row); and by the latter, including all permutations σ acting on X , denoted by X_σ (second row). The states detected by the method in each row are also detected by the method in the row above. The results for $d = 2$ hold because the reduction map and the partial transpose detect all entangled two-qubit states.

The reduction map acting on a subspace

Let us now consider the case where X is a projector, $X^2 = X$, on a two-dimensional subspace. Without loss of generality we can use the eigenbasis of X and use $X = \text{diag}(1, 1, 0, \dots, 0)$. Therefore, $\Psi_{(0,1)}^X(Y)$ corresponds to the reduction map acting on the projected matrix XYX , namely

$$\Psi_{(0,1)}^X(Y) = R(XYX). \quad (2.6.23)$$

Physically, $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB})$ detects the entanglement in ϱ_{AB} that remains present after ϱ_{AB} is projected in a lower-dimensional subspace on Alice's side. Table 2.2 summarizes the detection rate of either $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB})$ or $\Psi_{(0,1)}^X \otimes \text{id}(\varrho_{AB})$ on a random sample of Ginibre random states ϱ_{AB} . The following empirical observation from Table 2.2 shows that local projections can enhance the detectability of entanglement with certain methods.

Observation 4. *Entanglement detection with the reduction map can be enhanced locally, as follows.*

- (i) *There are entangled states ϱ_{AB} that are not detected by the Reduction map acting on the subsystem A , namely*

$$R_A \otimes \text{id}_B(\varrho_{AB}) \geq 0; \quad (2.6.24)$$

but can be detected by the reduction map if the subsystem A undergoes a projection P ,

$$R_A \otimes \text{id}_B[(P_A \otimes \mathbb{1}_B)\varrho_{AB}(P_A \otimes \mathbb{1}_B)] \not\geq 0. \quad (2.6.25)$$

- (ii) *In particular, for certain local dimensions $d := d_A = d_B$, most Ginibre random states are not detected by Eq. (2.6.24), but are detected by Eq. (2.6.25).*

2.6.4 Further methods of entanglement detection

Multipartite states

Proposition 2 provides a range of methods to detect entanglement from immanant inequalities, besides evaluation on a party of a bipartite state. An alternative method is given in the following Corollary.

Corollary 1. *Let ϱ be a quantum state shared among $(n - 1)$ parties with local dimension d . Let $\sum_{\lambda \vdash n} a_\lambda \text{Imm}_\lambda(G) \geq 0$ be an immanant inequality for any $d \times d$ positive matrix $G \geq 0$. If ϱ is fully*

separable, then

$$\mathrm{tr}_{1,\dots,n\setminus 1} \left[\sum_{\lambda \vdash n} a_\lambda P_\lambda(X^{\otimes n})(\mathbb{1}_d \otimes \varrho) \right] \geq 0. \quad (2.6.26)$$

Proof. The proof follows from Proposition 2 by decomposing a fully separable state ϱ as

$$\varrho = \sum_i p_i \varrho_1^i \otimes \cdots \otimes \varrho_{n-1}^i \quad (2.6.27)$$

where $p_i \geq 0$, $\sum_i p_i = 1$ and ϱ_l^i are quantum states for all i and l . It is now clear that Eq. (2.6.26) is a convex combination of positive operators. $\square$

As an example for bipartite states, let us choose the inequality $\prod_i G_{ii} - \det(G) \geq 0$ for 3×3 positive semidefinite matrices $G \geq 0$. Corollary 1 dictates that

$$\mathrm{tr}_{BC}((\mathbb{1} - 6P_{1,1,1})(X^{\otimes 3})(\varrho_{AB} \otimes \mathbb{1}_C)) \geq 0 \quad (2.6.28)$$

if the state ϱ_{AB} is separable.

Multiple copies

The formalism introduced in this Chapter extends also to detect entanglement in multiple copies of a state, and in particular recovers the results in [13] as follows. Consider n positive maps acting on $k-1$ positive matrices X_s derived through Proposition 2 from n Immanant inequalities, $\Psi_{\alpha_1}^{\mathbb{1}}(X_2, \dots, X_k), \dots, \Psi_{\alpha_n}^{\mathbb{1}}(X_2, \dots, X_k)$, where we set $X = \mathbb{1}$. To simplify the notation let us denote $\Psi_{\alpha_l}^{\mathbb{1}} := \Psi_l$. Given two n -partite quantum states ϱ and τ , consider the following expression,

$$\theta_\tau(\varrho) := \mathrm{tr}[\Psi_1 \otimes \dots \otimes \Psi_n(\varrho^{\otimes k-1})\tau], \quad (2.6.29)$$

where each map Ψ_l acts on $k-1$ copies of the l 'th party. The following result is a corollary of Proposition 2.

Corollary 2. *If the expression θ_τ in Equation (2.6.29) is negative, $\theta_\tau(\varrho) < 0$, then ϱ is entangled.*

Proof. It follows from Proposition 2 that the following operator is positive semidefinite on separable states, $\varrho = \sum_i p_i \varrho_1^i \otimes \dots \otimes \varrho_n^i$,

$$\Theta := \Psi_1 \otimes \dots \otimes \Psi_n(\varrho^{\otimes k-1}), \quad (2.6.30)$$

since it is a convex combination of tensor products of positive operators of the form

$$\Psi_l(\varrho_l^{i_1} \otimes \dots \otimes \varrho_l^{i_k}). \quad (2.6.31)$$

Equivalently, $\theta_\tau = \mathrm{tr}(\Theta\tau)$ is nonnegative for any positive semidefinite operator τ . $\square$

Note that from Corollary 2, one can obtain entanglement witnesses for multiple copies. This can be seen for the bipartite case by writing θ_τ as follows,

$$\begin{aligned} \theta_\tau &= \mathrm{tr} \left[\mathrm{tr}_{\{1,\dots,n\}\setminus 1} \left[\left(\sum_{\lambda} a_\lambda P_\lambda \right)_{A_1 \dots A_n} \otimes \left(\sum_{\mu} b_\mu P_\mu \right)_{B_1 \dots B_n} \cdot \mathbb{1}_{A_1 B_1} \otimes \varrho_{A_2 B_2} \otimes \dots \otimes \varrho_{A_k B_k} \right] \tau_{A_1 B_1} \right] \\ &= \mathrm{tr} \left[\left(\sum_{\lambda} a_\lambda P_\lambda \right)_{A_1 \dots A_n} \otimes \left(\sum_{\mu} b_\mu P_\mu \right)_{B_1 \dots B_n} \tau_{A_1 B_1} \otimes \varrho_{A_2 B_2} \otimes \dots \otimes \varrho_{A_k B_k} \right]. \end{aligned} \quad (2.6.32)$$

By setting $\tau = \varrho$ we recover the family of witnesses

$$\mathcal{W} = \left(\sum_{\lambda} a_\lambda P_\lambda \right)_{A_1 \dots A_n} \otimes \left(\sum_{\mu} b_\mu P_\mu \right)_{B_1 \dots B_n} \quad (2.6.33)$$

acting on k copies of ϱ introduced in [13], which detect entanglement on ϱ if $\text{tr}(W\varrho^{\otimes k}) < 0$. Naturally, this method accommodates protocols using an ancillary state shared among parties. For example, setting $\tau = |\phi^+\rangle\langle\phi^+|$ and using a witness W and the projector $|\phi^+\rangle\langle\phi^+|$ in d -dimensional homogenous systems, one can evaluate

$$\text{tr}(W_{A_1A_2} \otimes |\phi^+\rangle\langle\phi^+|_{B_1B_2} \cdot \varrho_{A_1B_1} \otimes |\phi^+\rangle\langle\phi^+|_{A_2B_2}) = \frac{1}{d^2} \text{tr}(W \varrho_{A_1A_2}). \quad (2.6.34)$$

Physically, Bob's part of state $\varrho_{A_1B_1}$ in B_1 is teleported to A_2 in Alice's side, and then Alice evaluates the witness W .

2.6.5 Conclusions

We have constructed a family of positive maps from immanant inequalities, that extends existing results and has significant advantage in detecting entanglement over known similar methods. We analyze the simplest example, where the reduction map is generalized to a family of maps able to detect more random states. We extend the applications of this method to detect multipartite states and multiple copies of -bipartite or multipartite- states. In particular instances, the analysis provided in this work suggests that projecting a quantum state on its local subsystems can increase detectability of entanglement.

This work further develops an avenue to entanglement detection which is currently at the beginning of its exploration, and extends the potential of immanants and the symmetric group in relevant tasks in quantum information. Due to the large supply of immanant inequalities in the literature and the multiple usages available of the positive maps introduced, our results suggest that the application of matrix inequalities in entanglement detection needs further exploration. In particular, it remains as an open question whether or not the techniques introduced here can detect bound entanglement.

Unitary invariant semidefinite programs: a user's guide

Invariant semidefinite problems are common in quantum information theory. Here we show how to reduce $U^{\otimes n}$ -invariant problems in practice. This includes the standard symmetry-reduction through the Schur-Weyl duality, reduce the number of variables with the Good Permutations basis, and further simplification by optimizing over expectation values. This approach allows to solve problems for twelve qubits, and is applied to paradigmatic problems such as quantum Max Cut, approximate ground states with translation invariance, and quantum marginal problems for Werner states.

3.1 Introduction

Semidefinite programming (SDP) is a useful tool to solve convex optimization problems with linear constraints [82, 88, 162]. This technique is particularly useful in quantum information, for example when optimizing a function over a convex set or certifying membership to a set [6, 83, 86, 88, 163]. However, one often faces computational limitations when considering nontrivial instances of a certain problem. The Hilbert space grows exponentially with the system size, and therefore optimization needs to be done imposing positivity of large matrices. As a result, further techniques to reduce the size of desired problems are required in specific cases.

It is often the case that the problem in consideration displays symmetries. This means that the constraints and the feasible set are invariant under the action of a certain group G , and the problem is denoted as G -invariant. This invariance results in redundancies in the feasible set to be optimized over, which can be reduced by optimizing over a smaller set. In this spirit, systematic methods to reduce both the number of parameters and the size of the matrix where to impose positivity have been derived [11]. In general, one identifies the symmetry group, block-diagonalizes the space containing the feasible set, and writes an equivalent problem requiring positivity of smaller matrices than in its standard formulation.

In particular, here we focus on problems that are invariant under the conjugate diagonal action of $U(d)$, called $U^{\otimes n}$ -invariant. For this type of problems, we will use Schur-Weyl duality, which states that the Hilbert space can be block-diagonalized in a tensor product of two subspaces: the subspace invariant under the action of $U^{\otimes n}$, and the subspace invariant under the permutation group S_n acting on $(\mathbb{C}^d)^{\otimes n}$. Then positivity is imposed on matrices whose size is the dimension of the irreducible subspaces invariant under copies of unitaries.

This approach allows to access approximations to relevant problems in quantum information theory, such as finding the ground state to a local Hamiltonian. Therefore, symmetry reduction has been used to tackle invariant problems in several works [11, 164, 165]. However, a practical implementation requires further exploration of the reduced size of certain problems and the numerical tools available. Here we provide an analysis on the number of variables needed in real and complex

SDPs, the size of the positivity constraint in symmetry-reduced problems, and a direct decomposition that simplifies the objective function. We compare different computational implementations of a practical approach using numerical tools in symbolic algebra. We discuss the implementation of these methods for the quantum max-cut [164, 165], compatibility of marginals in Werner states [5], and lower-bounding the ground state energy of an infinite spin lattice [166].

3.2 Problem formulation and proof strategy

3.2.1 General problem formulation

In quantum mechanics, a variety of problems can be written as a semidefinite program (SDP),

$$\begin{aligned} \min_X \quad & \text{tr}(AX) \\ \text{such that} \quad & X = X^\dagger \geq 0 \\ & \Xi(X) = C, \end{aligned} \tag{3.2.1}$$

where Ξ is a linear map and $A \in \mathcal{B}(\mathcal{H})$ acts linearly on a complex Hilbert space $\mathcal{H}$. As discussed in Section 1.2.1, the set of matrices X optimizing the objective function and satisfying the constraints is known as the *feasible set*. Although this is the general formulation, it is often the case that symmetries apply. That is, there exists a group G such that the feasible set is invariant under the action of G , $T(g)XT(g^{-1}) = X$ for all $g \in G$ where T is the representation of G on the space $\mathcal{H}$.

This is known as a *symmetry* of the problem (3.2.1). By linearity, if X is a solution, then the group average X^G , written $\sum_{g \in G} \frac{1}{|G|} gXg^{-1}$ for finite groups and $\int gXg^{-1} d\mu g$ for compact groups given a group measure μ , is also a solution. Therefore, the optimum can be found in the G -invariant subspace $\mathcal{H}^G$. In particular, if A and the constraints of the problem are invariant under a group G , then the feasible set is also invariant under G .

In general, two approaches come to mind to use the symmetry to simplify the problem in hand.

1. *Symmetry-adapted basis.* The SDP problem of Eq. (3.2.1) in the complex Hilbert space $\mathcal{H}$ has exponentially many real variables, which are the coefficients x'_i in the expansion $X = \sum_i x'_i X_i$ where $\{X_i\}$ is a basis for the $\mathcal{H}$. However, the spanning basis $\{X_i\}$ can be overcomplete in case of symmetries. This results in linear dependence between the elements $X_i = \sum_g T(g)X_iT(g^{-1})$. Then the variable X can be expanded in a new basis $\{F_i\}$ of the invariant subspace $\mathcal{H}^G$, $X = \sum_i x_i F_i$, and the optimization can be done over a complex vector with new components x_i .
2. *Block diagonalization.* It is well known that any matrix algebra over $\mathbb{C}$ can be block-diagonalized. This is called the Wedderburn decomposition. In the case of semidefinite programs, the positivity constraint in Eq. (3.2.1) is equivalent to a positivity constraint on each block of A [11], which has size smaller than the dimensions of $\mathcal{H}$ (see Theorem 1 below). In practice, finding the basis change for block-diagonalization is challenging. In specific cases with a known symmetry group G , block-diagonalization can be done in irreducible representations of G .

3.2.2 Specific problem formulation

Here we will focus on problems involving n quantum systems with local dimension d , namely $X \in \mathcal{B}((\mathbb{C}^d)^{\otimes n}) = \mathbb{C}^{d^n \times d^n}$ acts on $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$. We will study the particular case of invariant problems with the symmetry $U^{\otimes n}$ with $U \in \mathcal{U}(d)$. That is, the variable X has the property

$$U^{\otimes n} X U^{\otimes n \dagger} = X. \tag{3.2.2}$$

For this invariance, the decomposition through the Schur-Weyl duality is well known,

$$(\mathbb{C}^d)^{\otimes n} \simeq \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathcal{U}_\lambda \otimes \mathcal{S}_\lambda, \quad (3.2.3)$$

where $\mathcal{U}_\lambda$ and $\mathcal{S}_\lambda$ are the invariant subspaces of an irreducible representation labelled by λ of the unitary group $U(d)$ and the symmetric group S_n , respectively. As a result of Eq. (3.2.3), the commutant of $U^{\otimes n}$ is the symmetric group S_n , which acts on $(\mathbb{C}^d)^{\otimes n}$ as

$$\eta_d(\sigma) |v_1\rangle \otimes \cdots \otimes |v_n\rangle = |v_{\sigma^{-1}(1)}\rangle \otimes \cdots \otimes |v_{\sigma^{-1}(n)}\rangle, \quad (3.2.4)$$

where $\eta_d(\sigma)$ is the representation of a permutation $\sigma \in S_n$ on $(\mathbb{C}^d)^{\otimes n}$.

This implies that the invariant subspace of an SDP invariant under $U^{\otimes n}$ is $(\mathbb{C}^{d^n \times d^n})^{\mathcal{U}(d)} = \{\sum_{\sigma \in S_n} a_\sigma \eta_d(\sigma)\}$, and therefore the solution can be expressed as a linear combination of permutations acting on $(\mathbb{C}^d)^{\otimes n}$. For this case, we show how the general approach of Section 3.2.1 can be further fine-tuned using complementary methods.

3.2.3 Methods

1. **Symmetry reduction.** Through the Schur-Weyl duality, the variable X can be block-diagonalized. Each block belongs to the smallest subspace which is invariant under an irreducible representation R_λ of the symmetric group S_n . Thus the positivity constraint can be reduced to imposing that the block of X acting on each irreducible subspace is positive semidefinite. We refer to this as *symmetry reduction*.
2. **Variable reduction.** The permutations $\{\eta_d(\sigma)\}$ form an overcomplete basis for the symmetric group in $(\mathbb{C}^d)^{\otimes n}$. Thus the variable X can be expanded onto other bases with smaller number of elements, like the *good permutation basis* [157]. This can be further reduced for real problems, namely having real variables $x_\sigma \in \mathbb{R}$. Then one has the additional symmetry $x_\sigma = x_{\sigma^{-1}}$, which reduces the number of variables. We will refer to this as *variable reduction*.
3. **Direct decomposition in expectation values.** Certain problems can be written in a more compact form than their general formulation. For example, the energy minimization of a given Hamiltonian can be written such that the variables are the expectation values of the ground state with respect to each element of a reduced basis, like the good permutations basis. While such basis is overcomplete and thus in principle a rescaling is needed for the positivity constraint, it is shown that this can be avoided in invariant SDP problems.

3.2.4 Targeted applications

In particular, we will focus on the following problems, which are a current focus of research [164–167]:

1. *Quantum Max Cut.* Let $\mathbb{G} = \{E, V\}$ be a graph. The problem is to find the ground state of a Heisenberg Hamiltonian $H = \sum_{i,j \in E} \eta_d(ij)$, where $\eta_d(ij)$ permutes the qudits i and j as $\eta_d(ij) |v_i\rangle \otimes |v_j\rangle = |v_j\rangle \otimes |v_i\rangle$. This problem has been recently addressed using variable reduction through Schur-Weyl duality in [164, 165].
2. *The quantum marginal problem for Werner states.* The general marginal problem asks the following question: given a set of prescribed k_i -partite marginal states ϱ_i , does there exist a global n -partite state whose partial traces coincide with the marginals ϱ_i ? Here we focus on the particular case of Werner states, namely whether given a set of marginals $\varrho_i = \sum_{\sigma \in S_{k_i}} a_\sigma^i \eta_d(\sigma)$ acting on Hilbert subspaces $\mathcal{H}_i = (\mathbb{C}^d)^{\otimes k_i}$, there exists a global state

$\varrho \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ whose marginals are ϱ_i . In a different approach, compatibility relations between global symmetric states and their marginals has been recently addressed in [167].

3. *Ground states of translation invariant Hamiltonians.* Consider an infinite lattice $\mathcal{L}$ of spins. The problem consists in finding the ground state of the nearest-neighbour Hamiltonian that is invariant under translations. That is, the Hamiltonian reads

$$H = \sum_{j \in \mathcal{L}} \tau_j(h), \quad (3.2.5)$$

where h is the nearest-neighbour hamiltonian of a subset of m sites and τ_j places the nearest-neighbour Hamiltonian term h at a site $j \in \mathcal{L}$ [166]. The lowest energy of translationally invariant Hamiltonians can be lower bounded with Anderson bounds [10], by considering the ground state of the Hamiltonian corresponding to a finite sublattice. An improvement for these bounds has been recently shown in [166] by adding marginal constraints to the ground state of each sublattice.

3.2.5 Notation

We will use a similar notation as in Section 1.2.2: here $\sigma \in S_n$ is a permutation and $\eta_d(\sigma)$ is its representation acting on $(\mathbb{C}^d)^{\otimes n}$, according to Eq. (3.2.4). Given the symmetric group S_n of order n , one obtains the symmetric group algebra $\mathbb{C}S_n$ by taking formal sums $\sum_{\sigma \in S_n} a_\sigma \sigma$ where $a_\sigma \in \mathbb{C}$ and $\sigma \in S_n$. Multiplication of two elements $a, b \in \mathbb{C}S_n$ is done via

$$a \cdot b = \sum_{\sigma, \pi \in S_n} a_\sigma b_\pi \sigma \pi = \sum_{\tau \in S_n} c_\tau \tau, \quad (3.2.6)$$

where

$$c_\tau = \sum_{\substack{\sigma, \pi \text{ s.t.} \\ \sigma \pi = \tau}} a_\sigma b_\pi \quad (3.2.7)$$

For each element $a = \sum_{\sigma \in S_n} a_\sigma \sigma \in \mathbb{C}S_n$ we define its adjoint, $a^\dagger = \sum_{\sigma \in S_n} a_\sigma^* \sigma^{-1}$. Here $*$ denotes complex conjugation and σ^{-1} is the group inverse of σ such that $\sigma \sigma^{-1} = \text{id}$. Given a natural number n , we denote each partition by λ . This defines an irreducible representation of S_n , which we denote as R_λ . By default we will assume R_λ to be an orthogonal representation. We denote the representation of $\mathbb{C}S_n$ in the Hilbert space of n qu-dits as $\Sigma_n(\mathbb{C}^d)$, following [157].

3.3 Symmetry reduction

3.3.1 General approach

We will follow the lines of [11, 168] to explain the general methods used for invariant SDP problems where the solution lies in $\mathcal{B}((\mathbb{C}^d)^{\otimes n})$ and is symmetric under tensor factors of unitaries, $U^{\otimes n}$. Consider the semidefinite program of Eq. (3.2.1),

$$\begin{aligned} \min_{X} \quad & \text{tr}(AX) \\ \text{such that} \quad & X = X^\dagger \geq 0, \\ & \Xi(X) = C, \end{aligned} \quad (3.3.1)$$

where $X \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ is invariant under the action of the group $G = \{U^{\otimes n}\}$, namely $gAg^{-1} = A$ for all $g \in G$, where Ξ is a linear map with dual Ξ^* . This has two implications:

(1) *Restriction to invariant subspace.* If X is a solution to (3.3.1), then the group average X^G is also a solution [88, 168]. In this case the average reads

$$X^G = \int U^{\otimes n} X U^{\otimes n} d_\mu U, \quad (3.3.2)$$

where μ is the Haar measure over unitaries $U \in \mathcal{U}(d)$. Therefore, the solution X can be found in the G -invariant subspace $\mathcal{B}((\mathbb{C}^d)^{\otimes n})^G$. This can be imposed by adding the extra constraint

$$X^G = X \quad (3.3.3)$$

to Eq. (3.3.1).

Since $X \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})^G$, the variable X expands as $X = \sum_i x_i F_i$ in a basis $\{F_i\}$ of $\mathcal{B}((\mathbb{C}^d)^{\otimes n})^G$ and is effectively a vector with components $\{x_i\}$ instead of over a complex matrix of size $d^n \times d^n$.

(2) *Block diagonalization.* A key observation is that the invariant subspace $(\mathbb{C}^{d^n \times d^n})^G$ forms a matrix $*$ -algebra. That is, $(\mathbb{C}^{d^n \times d^n})^G$ is closed under matrix multiplication, since $gXYg^{-1} = gXg^{-1}gYg^{-1} = XY$ for all $X, Y \in (\mathbb{C}^{d^n \times d^n})^G$; $(\mathbb{C}^{d^n \times d^n})^G$ is closed under conjugate transpose, since $gX^\dagger g^{-1} = (gXg^{-1})^\dagger = X^\dagger$; and $(\mathbb{C}^{d^n \times d^n})^G$ is closed under addition and scalar multiplication, since the action of G is linear. Therefore, the following theorem applies:

Theorem 1 (Theorem 1.2 in [11]). *Let $\mathcal{A}$ be a matrix $*$ -algebra. There are numbers k , and $m_1, \dots, m_k$ such that there exists a $*$ -isomorphism between $\mathcal{A}$ and a direct sum of complete matrix algebras,*

$$\Psi : \mathcal{A} \rightarrow \bigoplus_{r=1}^k \mathbb{C}^{m_r \times m_r}. \quad (3.3.4)$$

Since Ψ is an isomorphism, it preserves the scalar product and in particular positive semidefiniteness. Therefore, the positivity condition $X = \sum_i x_i F_i \geq 0$ is equivalent to $\Psi(X) = \sum_i x_i \Psi(F_i) \geq 0$, where all $\Psi(F_i)$ have the form of k blocks of size m_r each. This reduces the size of the positivity constraint in (3.2.1).

With this, Equation (3.2.1) can be written as

$$\begin{aligned} \min_{x_i \in \mathbb{C}} \quad & \sum_i x_i \operatorname{tr}(AF_i) \\ \text{such that} \quad & \sum_i x_i \Psi(F_i) = \sum_i x_i^* \Psi(F_i)^\dagger \geq 0, \\ & \sum_i x_i \Xi(F_i) = C. \end{aligned} \quad (3.3.5)$$

3.3.2 Schur-Weyl duality

In general finding the $*$ -isomorphism Ψ that block-diagonalizes an invariant semidefinite program is nontrivial. However, for problems which are invariant under $U^{\otimes n}$ with $U \in \mathcal{U}(d)$, a practical guideline is given by Schur-Weyl duality. This states that there exists an isomorphism between the n -qudit Hilbert space and a block-diagonal space where each block has a tensor product structure,

$$(\mathbb{C}^d)^{\otimes n} \simeq \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathcal{U}_\lambda \otimes \mathcal{S}_\lambda, \quad (3.3.6)$$

where $\mathcal{U}_\lambda$ and $\mathcal{S}_\lambda$ are the invariant subspaces of an irreducible representation labelled by λ of the unitary group $U(d)$ and the symmetric group S_n respectively, and $|\lambda|$ is the number of nonzero components of the partition λ .

If $A \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ is symmetric under $U^{\otimes n}$, then Eq. (3.3.6) implies that A can be block-diagonalized in some basis as

$$A = \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} Q_\lambda(u) \otimes R_\lambda(a), \quad (3.3.7)$$

where $Q_\lambda(u)$ and $R_\lambda(a)$ are irreducible representations of $u \in \mathbb{C}U(d)$ and $a \in \mathbb{C}S_n$. It follows from Eq. (3.3.6) that the group actions of $U(d)$ and S_n in the Hilbert space are the unique commutant of each other. In particular, A in (3.2.1) commutes with $U^{\otimes n}$ and therefore the irrep $Q_\lambda(u)$ in (3.3.7) can only be the identity. Thus we have

$$A = \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathbb{1}_{s_\lambda} \otimes R_\lambda(a), \quad (3.3.8)$$

where $a = \sum_{\sigma \in S_n} a_\sigma \sigma \in \mathbb{C}S_n$ and the dimension s_λ of the identity corresponds to the multiplicity of R_λ and is given by the Schur polynomial evaluated on $|\lambda|$ ones,

$$s_\lambda := s_\lambda(1, \dots, 1) = \prod_{1 \leq i < j \leq |\lambda|} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \quad (3.3.9)$$

3.3.3 Objective function, constraints and positivity

Similarly as in Eq. (3.3.8), by Theorem 1 the variable X can be diagonalized in blocks as follows,

$$X = \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathbb{1}_{s_\lambda} \otimes R_\lambda(x), \quad (3.3.10)$$

where $x = \sum_{\sigma \in S_n} x_\sigma \sigma \in \mathbb{C}S_n$. Therefore, the objective function in Eq. (3.2.1) can be written in terms of the irreducible representations R_λ as

$$\text{tr}(AX) = \text{tr}(\eta_d(ax)) = \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \text{tr}(R_\lambda(a) R_\lambda(x)) s_\lambda. \quad (3.3.11)$$

The Hermiticity condition $X = X^\dagger$ reads

$$\sum_{\sigma \in S_n} x_\sigma \sigma = \sum_{\sigma \in S_n} x_\sigma^* \sigma^{-1}, \quad (3.3.12)$$

where we take the adjoint in the group algebra.

Since η_d is a linear representation and Ξ a linear map, there exists a linear function ξ on $\mathbb{C}S_n$ satisfying $\Xi(X) = \eta_d(\xi(x))$. Thus the constraints in (3.2.1) can be written in several ways:

$$\begin{aligned} \xi(x) &= c, \\ R_\lambda(\xi(x)) &= R_\lambda(c) \quad \text{or} \\ \Xi(X) &= \eta_d(\xi(x)) = \eta_d(c). \end{aligned} \quad (3.3.13)$$

Since the Schur-Weyl duality establishes an isomorphism, the right hand side of Eq. (3.3.6) is positive semidefinite if and only if the left hand side is positive semidefinite. As a result, the positivity condition $X \geq 0$ in (3.2.1) is satisfied if and only if the matrix irreducible representation $R_\lambda(x)$ is positive semidefinite for all λ appearing in Eq. (3.3.10), namely satisfying $|\lambda| \leq d$.

A particularly relevant linear constraint is requiring that the partial trace of the solution coincides with a prescribed reduction, including the standard normalization condition $\text{tr}(X) = 1$ for

quantum states. Making use of the freedom given by Eq. (3.3.13), we define the following linear map.

Definition 1 (Partial trace in $\mathbb{C}S_n$). *Let $\sigma = (\alpha_1 \dots \alpha_l) \dots (\alpha_{n-l} \dots \alpha_n) \in S_n$ be a permutation with $c(\sigma)$ cycles, and let $\{\alpha_i\} \subset \{1, \dots, n\}$ be k subsystems to be traced out. Let $\sigma_{\setminus \{\alpha_i\}} \in S_k$ be the result of omitting the parties α_i to the permutation σ , with $c(\sigma_{\setminus \{\alpha_i\}})$ cycles. We define the partial trace of σ as*

$$\tau_{\{\alpha_i\}}(\sigma) = d^{c(\sigma) - c(\sigma_{\setminus \{\alpha_i\}})} \sigma_{\setminus \{\alpha_i\}}. \quad (3.3.14)$$

One verifies that the partial trace in $\mathbb{C}S_n$ and in $\eta_d(\mathbb{C}S_n)$ are related as follows,

$$\text{tr}_{\{\alpha_i\}}(\eta_d(\sigma)) = \eta_d(\tau_{\{\alpha_i\}}(\sigma)). \quad (3.3.15)$$

This is because (1) the full trace of a cycle equals the local dimension, e.g.

$$\text{tr}(\eta_d(\alpha_1 \dots \alpha_l)) = d. \quad (3.3.16)$$

And (2) the partial trace of a cycle results a shorter cycle where the traced subsystems are omitted, e.g.

$$\text{tr}_{1,2,4}(\eta_d(\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5)) = \eta_d(\alpha_3 \alpha_5). \quad (3.3.17)$$

3.3.4 Main SDP symmetry reduction

At this point we recover a particular case of Theorem (1) where the block-diagonalization is done in irreducible representations. The semidefinite programming (3.2.1) can be written as follows,

$$\begin{aligned} \min_{\{x_\sigma \in \mathbb{C}\}} \quad & \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} s_\lambda \sum_{\sigma, \pi \in S_n} x_\sigma a_\pi \text{tr}(R_\lambda(\sigma) R_\lambda(\pi)) \\ \text{such that} \quad & \sum_{\sigma \in S_n} x_\sigma R_\lambda(\sigma) = \sum_{\sigma \in S_n} x_\sigma^* R_\lambda(\sigma^{-1}) \geq 0 \quad \text{for all } \lambda \\ & \Xi(X) = C, \end{aligned} \quad (3.3.18)$$

where the linear constraint $\Xi(X) = C$ is problem specific.

3.4 Variable reduction

3.4.1 Good permutations basis

Although the method described can be useful to reduce the dimensionality of certain symmetric semidefinite programs, a general solution

$$X = \sum_{\sigma \in S_n} x_\sigma \eta_d(\sigma) \quad (3.4.1)$$

is composed of $|S_n| = n!$ complex parameters x_σ , whose scaling is comparable to an exponential scaling. The key fact to reduce the number of variables is that the permutations $\sigma \in S_n$ form an overcomplete basis for the subspace spanned by $\eta_d(\sigma)$ with $\sigma \in S_n$. Some redundancy can be taken away by choosing a restricted basis, consisting of so-called *good permutations* [157], defined as follows.

Definition 2 (Definition 3.1 in [157]). *A permutation $\sigma \in S_n$ is called d -bad if it has a descending subsequence of length d , namely, if there exists a sequence $1 \leq i_1 < i_2 < \dots < i_d < n$ such that $\sigma(i_1) > \sigma(i_2) > \dots > \sigma(i_d)$. Otherwise, σ is called d -good.*

n	3	4	5	6	7	8	9	10	11	12
$ GP(2)_n $	5	14	42	132	429	1430	4862	16796	58786	208012
$ F_n $	3	9	25	75	231	763	2619	9495	35695	140151
$n!$	6	24	120	720	5040	40320	362880	3628800	39916800	479001600

Table 3.1: Variable reduction in an n -qubit $U^{\otimes n}$ -invariant SDP. The first row displays the number $|GP(2)_n|$ of $(2+1)$ -good permutations $GP(2)_n$ for each number of elements n , given by the Catalan number, which is the number of variables of a complex semidefinite programming problem involving n qubits. The second row shows the number $|F_n|$ of permutations which are disjoint products of swaps, given by Eq. (3.4.2), which is the number of variables involved in a real semidefinite programming. These compare to the third row, which contains the number of variables without variable reduction determined by the Factorial function $n!$.

Theorem 2 (Theorem 3.3 in [157]). *The $d+1$ -good permutations form a basis of the invariant subspace $\Sigma_n(\mathbb{C}^d) = \{\sum_{\sigma \in S_n} a_\sigma \eta_d(\sigma) : a_\sigma \in \mathbb{C}, \sigma \in S_n\}$.*

We shall denote the set of $(d+1)$ -good permutations of n elements as $GP(d)_n$, and its cardinality as $|GP(d)_n|$.

Consider for example three qubits, which define the algebra $\Sigma_3(\mathbb{C}^2)$. The only descending sequence of $2+1=3$ elements is $1 < 2 < 3$. Thus a permutation $\sigma \in S_3$ is 3-bad if $\sigma(1) > \sigma(2) > \sigma(3)$, and 3-good otherwise. Since the only 3-bad permutation is $\sigma = (13)$, by Theorem 2 the algebra $\Sigma_3(\mathbb{C}^2)$ is spanned by 5 permutations: $\eta_2(\text{id})$, $\eta_2(12)$, $\eta_2(23)$, $\eta_2(123)$ and $\eta_2(132)$.

3.4.2 Real problems

In some cases, the solution to Eq. (3.2.1) is real, $X \in (\mathbb{R}^d)^{\otimes n}$. In this case the variables have no complex component, $x_\sigma \in \mathbb{R}$ in (3.3.18). In practice, this reduces the number of variables needed. For example, the Hermiticity condition in (3.3.18) imposes $x_\sigma = x_{\sigma^{-1}}^* = x_{\sigma^{-1}}$. Permutations satisfying this property are called *of order 2* or *involutions*, and correspond to disjoint products of swaps. We will denote the set of disjoint products of swaps in S_n as F_n , and its cardinality as $|F_n|$. One can reduce the number of variables for real n -qudit SDP problems using the following result.

Theorem 3 (Theorem 1.2-(3) in [169]). *The space of real and symmetric elements has a basis over $\mathbb{R}$ formed by involutions.*

This reduces the size of real problems with respect to complex ones, where the good permutations basis is needed. More precisely, the number of involutions in S_n is known to be [170]

$$|F_n| = \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{n!}{(n-2k)!2^k k!}. \quad (3.4.2)$$

Instances of the reduction with respect to the good permutation basis are given in Table 3.2.

3.5 Scaling

Here we will focus on the size of the matrices where the positivity condition is imposed. In the standard SDP, one has a complex square matrix of size with d^n . This means $d^n(d^n+1)/2$ real parameters and $(d^n(d^n-1))/2$ imaginary parameters, which sum to $(d^n)^2$ real parameters.

System $(d; n)$	Vars naive	Vars SR (GP)	Real	Blocks	Max. Part.	Size
$(2; 2)$	$(2^2)^2 = 16$	2×2	2	2	$(1, 1)$	1
$(3; 2)$	$(3^2)^2 = 81$	2×5	4	2	$(1, 1)$	1
$(2; 3)$	$(2^3)^2 = 64$	2×2	2	2	$(2, 1)$	2
$(3; 3)$	$(3^3)^2 = 729$	2×6	5	3	$(2; 1)$	2
$(2; 10)$	1048576	2×16796	8524	5	$(6, 4)$	90
$(10; 10)$	10^{20}	$2 \times 10!$	$10!$	42	$(4, 3, 2, 1)$	768

Table 3.2: Size of the positivity constraint in different SDP formulations. Shown in columns are the size parameters of the naive SDP and its symmetry reduced version. *System:* Size of the system, composed of n d -dimensional local subsystems. *Vars naive:* Real parameters of the variable X , imposing hermiticity (necessary for positivity) and unit trace. There are $(d^n(d^n + 1))/2$ real parameters and $(d^n(d^n - 1))/2$ imaginary parameters, which sum to d^n real parameters. *Vars SR:* Total number of parameters of the symmetry reduced SDP. This number coincides with both $\sum_{\lambda \vdash n, |\lambda| \leq d} \dim(\mathcal{S}_\lambda)^2$, where $\dim(\mathcal{S}_\lambda)$ is the dimension of the irrep R_λ , and the number of good permutations (GP). *Real:* Number of variables when the problem admits a real solution. *Blocks:* Total number of blocks in the diagonalization. *Max. Part.:* Partition corresponding to the largest block. *Size:* Size of the largest block.

In the reduced SDP, one only asks for positivity on complex linear combinations of irreps. This means that the total number of parameters is

$$\theta = \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \dim(\mathcal{S}_\lambda)^2 \quad (3.5.1)$$

where $\dim(\mathcal{S}_\lambda)$ is the dimension of the irrep R_λ . The comparison between the two methods is illustrated in Table 3.2.

Example 1. Consider an SDP with symmetries over $\mathcal{U}(d)^{\otimes n}$ with three qubits, namely $d = 2$ and $n = 3$. Its standard version has the form of Eq. (3.2.1), where we enforce positivity of a matrix of size $2^3 = 8$, and thus has $(2^3)^2 - 1 = 63$ real parameters assuming that the solution to the problem X is a normalized state with trace 1.

Now consider its symmetry reduction in Eq. (3.3.18). By Schur-Weyl duality we only need to consider positivity of the matrices

$$A = s_{(3)} \sum_{\sigma \in S_3} x_\sigma R_{(3)}(\sigma) \quad \text{and} \quad (3.5.2)$$

$$B = s_{(2,1)} \sum_{\sigma \in S_{(2,1)}} x_\sigma R_{(2,1)}(\sigma). \quad (3.5.3)$$

The irreducible representations $R_{(3)}$ and $R_{(2,1)}$ act on subspaces of dimensions 1 and 2 respectively. Therefore, we need to impose positivity of one matrix of size 1 (A) and one matrix of size 2 (B). Since by both A and B are hermitian, the problem has $1 + 2^2 - 1 = 4$ real variables, assuming that the solution X to the problem has trace 1.

Remark. In practice, solving an SDP problem can be done using different computational tools. The irreducible representation of each good permutation needs to be computed and saved for further use. This requires computational effort, and it is thus worth to compare different methods for the sake of efficiency. In Table 3.3 we compare the module “SymmetricGroupRepresentation” [171] from the “Combinat” package in SageMath with the module “Representation” from “NumericalRepresentationTheory” [172, 173] in Julia (open source), and find that the latter

n	3	4	5	6	7	8
Time with SGR (seconds)	0.006	0.031	0.238	2.442	66.963	1973.236
Time with NRT (seconds)	0.002	0.021	0.083	0.331	4.113	13.798

Table 3.3: Comparison between numerical methods to generate the irreducible representations of the $n!$ permutations in the symmetric group S_n . Shown is the time in seconds spent in a single run by each method: “SymmetricGroupRepresentation” (SGR) with Sage [171] and “NumericalRepresentationTheory” (NRT) with Julia [172].

is faster in computing the irreducible representations.

3.6 Direct decomposition

The semidefinite programs (SDP) treated in this chapter admit a further simplification. Recall that by Schur-Weyl duality, the representation $\eta_d(\sigma)$ of the permutations on $(\mathbb{C}^d)^{\otimes n}$ spans the subspace invariant under $U^{\otimes n}$. Therefore, both the variable X and the fixed operator A in Eq. (3.2.1) can be expanded in the basis $\{\eta_d(\sigma)\}$, where a rescaling is needed since the basis is overcomplete. The rescaling relies on the *Weingarten operator*

$$\text{Wg}(d, n) = \frac{1}{n!^2} \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \frac{\chi_\lambda(\text{id})^2}{s_\lambda} \sum_{\sigma \in S_n} \chi_\lambda(\sigma^{-1}) \sigma \quad (3.6.1)$$

and the *Weingarten function*

$$\text{Wg}(d, \sigma) = \frac{1}{n!^2} \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \frac{\chi_\lambda(\text{id})^2}{s_\lambda} \chi_\lambda(\sigma). \quad (3.6.2)$$

Let us expand the operators A and X as

$$A = \sum_{\sigma \in S_n} h_\sigma \eta_d(\sigma) \quad (3.6.3)$$

with $h_\sigma = \sum_{\tau \in S_n} \text{tr}(\eta_d(\tau^{-1})A) \text{Wg}(\eta_d(\tau)^{-1}\sigma, d)$ and

$$X = \sum_{\pi \in S_n} x_\pi \eta_d(\pi) \quad (3.6.4)$$

with $x_\pi = \sum_{\tau \in S_n} \text{tr}(\eta_d(\tau^{-1})X) \text{Wg}(\eta_d(\tau)^{-1}\pi, d)$, where the coefficients h_σ and x_π define the operators A and X respectively.

3.6.1 Objective function and positivity constraint

Using that the Weingarten operator $\text{Wg}(d, n)$ commutes with all permutations, we can write the objective function as

$$\begin{aligned} \text{tr}(AX) &= \sum_{\pi, \sigma \in S_n} \text{tr}(\eta_d(\sigma)^{-1}A) \text{tr}(\eta_d(\sigma)\eta_d(\pi) \text{Wg}(d, n)^2) \text{tr}(\eta_d(\pi^{-1})X) \\ &= \sum_{\pi \in S_n} \tilde{h}_\pi \langle \pi \rangle, \end{aligned} \quad (3.6.5)$$

by defining

$$\begin{aligned}\tilde{h}_\pi &= \sum_{\sigma \in S_n} \text{tr}(\eta_d(\sigma^{-1})A) \text{tr}(\eta_d(\sigma)\eta_d(\pi) \text{Wg}(d, n)^2) \quad \text{and} \\ \langle \pi \rangle &= \text{tr}(\eta_d(\pi)X),\end{aligned}\tag{3.6.6}$$

where $\langle \pi \rangle$ can be understood as the expectation value of the permutation π on the variable X . To impose the positivity constraint $X \geq 0$, we will use the following result of [174].

Proposition 3. [174]. *The operator X is positive semidefinite, $X \geq 0$, if and only if*

$$\sum_{\pi \in S_n} \langle \pi^{-1} \rangle R_\lambda(\pi) \geq 0 \tag{3.6.7}$$

for all partitions $\lambda \vdash n$ with $|\lambda| \leq d$.

Proof. Let us expand X as

$$X = \sum_{\pi \in S_n} \text{tr}(\eta_d(\pi^{-1})X) \eta_d(\pi) \text{Wg}(d, n) \tag{3.6.8}$$

$$= \frac{1}{n!} \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \frac{\chi_\lambda(\text{id})}{s_\lambda} \eta_d(p_\lambda) \sum_{\pi \in S_n} \text{tr}(\eta_d(\pi^{-1})X) \eta_d(\pi), \tag{3.6.9}$$

where $\eta_d(p_\lambda) = P_\lambda$ are the Young projectors and we used that $\text{Wg}(d, n)$ commutes with all permutations. Recall that the representation of an element on of the symmetric group algebra $a \in \mathbb{C}S_n$ on $(\mathbb{C}^d)^{\otimes n}$ is positive semidefinite, $\eta_d(a) \geq 0$, if and only if all its irreducible representations are positive semidefinite, namely $R_\mu(a) \geq 0$ for all partitions $\mu \vdash n$ with $|\mu| \leq d$. Since $n!$, $\chi_\lambda(\text{id})$ and s_λ are nonnegative, we have that

$$X \geq 0 \iff R_\mu\left(p_\lambda \sum_{\pi \in S_n} \langle x_\pi \rangle \pi\right) \geq 0. \tag{3.6.10}$$

Since p_λ acts as the identity onto the subspace invariant under R_λ and as the zero matrix onto other subspaces, we have

$$R_\mu(p_\lambda) = \mathbb{1} \delta_{\lambda, \mu}. \tag{3.6.11}$$

Therefore we conclude that $X \geq 0$ if and only if $\sum_{\pi \in S_n} \langle x_\pi \rangle R_\mu(\pi) \geq 0$ for all $\mu \vdash n$ with $|\mu| \leq d$. $\square$

Proposition 3 allows us to write a semidefinite program in (3.2.1) as follows,

$$\begin{aligned}\max \quad & \sum_{\sigma \in S_n} \tilde{h}_\sigma \langle \sigma \rangle \\ \text{such that} \quad & R_\lambda\left(p_\lambda \sum_{\sigma \in S_n} \langle x_\sigma \rangle \sigma\right) \geq 0 \quad \forall \lambda \vdash n, |\lambda| \leq d \\ & \sum_{\sigma \in S_n} \Phi(\sigma) \langle \sigma \rangle = C,\end{aligned}\tag{3.6.12}$$

where the linear map Φ is problem specific. This formulation simplifies the objective function with respect to Eq. (3.3.18), where a summation with a larger number of terms than in Eq. (3.6.12) is required.

3.7 Quantum Max Cut

The quantum max-cut problem consists of finding the minimal energy of a Hamiltonian of a spin lattice. Given a graph $G = (V, E)$ with vertices V and edges E , the Hamiltonian reads [164, 165]

$$H = \sum_{i,j \in E} \eta_d(ij) \quad (3.7.1)$$

and the ground state ϱ minimizes $\text{tr}(\varrho H)$. This problem can be written as a semidefinite program, as follows:

$$\begin{aligned} \min_X \quad & \text{tr}(HX) \\ \text{such that} \quad & X \geq 0 \\ & \text{tr}(X) = 1 \end{aligned} \quad (3.7.2)$$

The symmetry reduction techniques derived in this chapter allow to formulate this SDP in the following form,

$$\begin{aligned} \min_{\{x_\sigma \in \mathbb{C}, x = x^\dagger\}} \quad & \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} s_\lambda \sum_{\sigma \in P_n} \sum_{i,j \in E} x_\sigma \text{tr}(R_\lambda(\sigma) R_\lambda(i, j)) \\ \text{such that} \quad & \sum_{\sigma \in P_n} x_\sigma R_\lambda(\sigma) \geq 0 \quad \text{for all } \lambda: |\lambda| \leq d \\ & \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \sum_{\sigma \in P_n} x_\sigma \text{tr}(R_\lambda(r)) s_\lambda = 1, \end{aligned} \quad (3.7.3)$$

where the permutations are sampled over the set P_n of disjoint products of swaps since the problem admits a real solution.

A further simplification can yet be done by direct decomposition, as follows:

$$\begin{aligned} \min_{\{\langle \sigma \rangle: \sigma \in P_n\}} \quad & \sum_{\sigma \in P_n} \langle \sigma^{-1} \rangle h_\sigma \\ \text{so that} \quad & \sum_{\pi \in S_n} \langle \pi^{-1} \rangle R_\lambda(\pi) \geq 0 \\ & \langle \text{id} \rangle = 1. \end{aligned} \quad (3.7.4)$$

The condition $\langle (\text{id}) \rangle = 1$ imposes normalization of X since

$$\langle \text{id} \rangle = \text{tr}(\eta_d(\text{id})X) = \text{tr}(\mathbb{1}X). \quad (3.7.5)$$

3.8 Marginal problem for Werner states

Another application of the symmetry reduction of SDP's via Schur-Weyl duality is to asses compatibility of marginals of n -partite states which are invariant under $\mathcal{U}(d)^{\otimes n}$. This is called the *marginal problem* [175], and SDP techniques have been recently shown to be useful to asses similar compatibility problems for marginal spectra [5] and highly symmetric states [167]. The marginal problem is in general described by the following SDP,

$$\begin{aligned} \min_X \quad & \text{find } X \\ \text{such that} \quad & X = X^\dagger \geq 0 \\ & \text{tr}_{\mathcal{A}^c}(X) = X_{\mathcal{A}}, \end{aligned} \quad (3.8.1)$$

given prescribed marginals $X_{\mathcal{A}}$ of subsystems $\mathcal{A}$ with complementary $\mathcal{A}^c$. By imposing the symmetry $U^{\otimes n} X U^{\dagger \otimes n} = X$, this problem can be restated as

$$\begin{aligned}
& \text{find } \{x_\sigma\} \\
& \text{such that } \sum_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \sum_{\sigma \in GF(d)_n} x_\sigma \text{tr}(R_\lambda(r)) s_\lambda = 1 \\
& \sum_{\sigma \in GF(d)_n} x_\sigma R_\lambda(\sigma) = \sum_{\sigma \in GF(d)_n} x_\sigma^* R_\lambda(\sigma^{-1}) \geq 0 \quad \text{for all } \lambda \\
& \sum_{\sigma \in GF(d)_n} x_\sigma R_\lambda(\tau_{\mathcal{A}^c}(\sigma)) = \sum_{\sigma \in GF(d)_n} c_\sigma R_\lambda(\sigma) \quad \text{for all } \lambda \text{ and } \mathcal{A}
\end{aligned} \tag{3.8.2}$$

with Definition 1, where both the marginals and the solution are Werner states. Here the set of good permutations $GF(d)_n$ is needed, since in general the solution can be complex.

3.9 Infinite spin lattice with translation invariance

Another problem is that of finding the ground state of the Ising Hamiltonian in an infinite lattice, such that the reduced state at each cell (of say, 4 nodes) is invariant under translations. This problem cannot be tackled with exact diagonalization. Instead, upper approximations can be found with matrix product states (MPS) [176] and lower approximations can be found with SDP [164, 165]. Here we follow the relaxation proposed in [166], which can improve the Anderson bound [10].

Given a local translation invariant Hamiltonian $H_n = \sum_{j=1}^n \eta_d(\tau_j(12))$, where $\tau_j(12) = (j, j+1)$ is a translation on the Swap, we define three main ingredients:

- The Hamiltonian $h_m = \eta_d(12) + \dots + \eta_d(m-1, m)$ on a cell with m qudits,
- The Hamiltonian $h = \eta_d(m, 1)$ between the first and last nodes of the cell, and
- A set $B = \{m + j \bmod m : j = -s + 1, -s + 2, \dots, s\}$ for a chosen integer $s \leq m/2$.

With this we consider the following semidefinite program.

$$\begin{aligned}
& \min_{X, S} \quad \text{tr}(X h_m) + \text{tr}(S(\mathbf{1} \otimes h)) \\
& \text{such that} \quad \text{tr}_{\setminus B}(X) = S \\
& \quad \quad \quad X, S \geq 0, \text{tr}(X) = \text{tr}(S) = 1.
\end{aligned} \tag{3.9.1}$$

As an example, consider the parameters $m = 4$, $s = 2$. We have $h_4 = \eta_d(12) + \eta_d(23) + \eta_d(34)$ and $h = \eta_d(41)$. The set $B = \{3, 4, 1, 2\}$ defines the whole lattice and therefore the partial trace condition in Eq. (3.9.1) enforces $X = S$. Thus the optimization can be done with a single variable X with respect to the 4-nodes closed-graph Hamiltonian

$$h_{\Delta, 4} := h_4 + h = \eta_d(12) + \eta_d(23) + \eta_d(34) + \eta_d(41), \tag{3.9.2}$$

as in the quantum max-cut problem:

$$\begin{aligned}
& \min_X \text{tr}(X h_{\Delta, 4}) \\
& \text{such that} \quad X \geq 0, \text{tr}(X) = 1.
\end{aligned} \tag{3.9.3}$$

As another example, set $m = 4$ and $n = 1$. Now one has $B = \{4, 1\}$ and thus the Eq. (3.9.1)

reads as follows,

$$\begin{aligned} & \min_{X,S} \operatorname{tr}(Xh_4) + \operatorname{tr}(S(h)) \\ \text{such that } & \operatorname{tr}_{2,3}(X) = S \\ & X, S \geq 0, \operatorname{tr}(X) = \operatorname{tr}(S) = 1. \end{aligned} \tag{3.9.4}$$

3.10 Conclusions

We have studied semidefinite programs symmetric under the conjugate diagonal action of the unitary group, namely $U^{\otimes n}$ -invariant. We reviewed standard symmetry reduction techniques starting from Schur-Weyl duality [11, 164, 165] and further reduced the size of the problem using variable reduction: complex problems admit usage of the good-permutation basis, and real problems can be solved using a basis of permutations which are disjoint products of swaps. We have studied the implementation of these symmetry reduction techniques and compared different practical methods of their use. Our analysis suggests significant advantage of the package “Numerical Representation Theory” in Julia [172] over the package “Symmetric Group Representation” in SageMath [171].

Our approach is complemented with a direct decomposition of the objective function, which allows to write problems like ground state minimization in terms of the expectation values of the Hamiltonian alone. We have explicitly written symmetry-reduced versions of relevant problems in quantum information theory: the quantum max-cut, the marginal problem for Werner states, and the ground state minimization of an infinite spin lattice.

4

Discrete dynamics in the set of quantum measurements

A quantum measurement, often referred to as positive operator-valued measurement (POVM), is a set of semipositive operators $P_j = P_j^\dagger \geq 0$ summing to identity, $\sum_j P_j = \mathbf{1}$. This can be seen as a generalization of a probability distribution of positive real numbers summing to unity, whose evolution is given by a stochastic matrix. In this chapter we describe discrete transformations in the set of quantum measurements by *blockwise stochastic matrices*, composed of semipositive blocks that sum columnwise to identity, and the notion of *sequential product* of matrices. We show that such transformations correspond to a sequence of quantum measurements. Imposing additionally the dual condition that the sum of blocks in each row is equal to identity we arrive at blockwise bistochastic matrices (also called *quantum magic squares*) and study their properties. Analyzing the dynamics they induce we formulate a quantum analog of the Ostrowski description of the classical Birkhoff polytope and introduce the notion of majorization between quantum measurements. Our framework provides a dynamical characterization of the set of blockwise bistochastic matrices and establishes a resource theory in this set.

4.1 Introduction

Quantum theory can be regarded as a generalization of classical probabilistic theory, where classical phenomena emerge as a special case of quantum phenomena under certain constraints. Understanding the similarities and differences between quantum and classical theory is of both fundamental and applied interest. Specifically, quantum measurements are based on mathematical postulates that yield outcome probability distributions [177], revealing certain properties of a quantum system [21].

Although extensive research has been devoted to exploring the relationship between quantum measurements and probabilistic experiments [178], the comparison of the dynamical properties of discrete quantum measurements and discrete probability distributions requires further studies. This work aims to identify the conditions under which standard properties of the evolution of discrete probability distributions apply analogously to the evolution of discrete quantum measurements.

In classical probability theory, a discrete probability distribution is a column vector, $p = (p_1, \dots, p_n)^T$, with nonnegative entries p_i which sum to unity. The set of probability vectors Δ_n is known as the *probability simplex*. Discrete evolution of a probability vector p is described by multiplication by a stochastic matrix S with nonnegative entries s_{ij} and columns s_j summing to unity. Physically, this can be understood as follows: first, an experimental measurement with outcome distribution p (e.g. rolling dices) is performed; depending on the outcome j , a second measurement is performed with outcome probability distribution s_j , where each outcome i occurs with conditional probability s_{ij} . The marginal probability distribution after both measurements is $q = (q_1, \dots, q_n)^T$ with

$$q_i = \sum_{j=1}^n s_{ij} p_j. \quad (4.1.1)$$

Discrete probabilistic systems whose evolution is given by a stochastic matrix via Eq. (4.1.1) are known as *Markov chains*, and have the property that the state of the system at each time step depends only on the state of the system at the previous time step [179].

A particular dynamics in the space of classical probability distributions is induced by bistochastic matrices, with nonnegative entries which sum to unity for both rows and columns. The set of bistochastic matrices of a certain size n has been characterized in two equivalent ways:

1. *Birkhoff-von Neumann characterization* [180]: The set of bistochastic matrices of size n , also known as the Birkhoff polytope, is the convex hull of all permutation matrices of size n .
2. *Ostrowski characterization* [181]: A square matrix of size n with nonnegative entries is bistochastic, if and only if, one has the majorization relation $p \succ Bp$ for any probability vector $p \in \Delta_n$.

The majorization relation $p \succ q$ between two vectors $p, q \in \Delta_n$ means that

$$\sum_{j=1}^k p_j \geq \sum_{j=1}^k q_j \quad (4.1.2)$$

for all $k = 1, 2, \dots, n$, where the entries p_j are sorted in non-increasing order [182, 183]. Due to the characterization of Ostrowski [181, 182] summarized in item (ii), bistochastic matrices induce majorization relations between probability distributions and a corresponding resource theory [184].

Similar ideas to the ones describing the probability simplex can be used in quantum theory, in various ways. The most common approach is to consider density matrices (positive semidefinite normalized operators $\varrho = \varrho^\dagger \geq 0$) as coherent generalizations of probability distributions. Discrete dynamics in the set Ω_d of density matrices of order d is described by quantum operations $\mathcal{E}$

(completely positive and trace preserving) represented by Kraus operators K_i in the form of

$$\mathcal{E}(\varrho) = \sum_i K_i \varrho K_i^\dagger, \quad (4.1.3)$$

where the *effects* $K_i^\dagger K_i$ sum to identity, $\sum_i K_i^\dagger K_i = \mathbb{1}$. In this work we use an alternative approach in the spirit of Gudder [93, 107, 185], in which probability vectors are generalized to vectors whose components are effects, $P_i := K_i^\dagger K_i$, summing to identity [92, 186].

The goal of this paper is to understand whether certain dynamical properties of the probability simplex Δ_n hold analogously for the set of quantum measurements, when the agents of the evolution are elements of their own set. We start from a vector description of quantum measurements [92] and define a product between these objects, which is based on matrix convex combination of effects [107, 187]. This product can be understood as a concatenation of quantum measurements, where transformations of measurement are induced by objects *within* their own set. Therefore, transformations correspond to sequential Lüders measurements, in which quantum instruments [188] (or more generally, quantum supermaps [28, 64, 189, 190]) are reduced to measurement effects. In particular, bistochastic matrices have been generalized over an operator algebra in several contexts [92, 102, 186, 191–196], as blockwise square matrices where positive semidefinite entries sum to identity for rows and columns. These objects are called *blockwise bistochastic matrices* [191] or *quantum magic squares* [186], and constitute a very special case of blockwise stochastic matrices inducing dynamics in our framework.

This paper is organized as follows. In section 1.2.3 we introduce the set of quantum measurements as blockwise probability vectors, and discuss its geometry and generalized convex hulls. In section 4.2 we introduce blockwise stochastic matrices and define a product between quantum measurements based on a generalized convex combination of effects. We discuss its physical meaning in terms of conditional measurements and measurement compatibility. In section 4.3 we introduce a generalized notion of majorization between quantum measurements and use it to establish a dynamical characterization of blockwise bistochastic matrices analogous to the Ostrowski characterization of the Birkhoff polytope [180, 181]. In section 4.4 we show that a resource theory of quantum measurements arises from our results, as a quantum analog of the majorization resource theory in the probability simplex.

4.2 Discrete dynamics induced by blockwise stochastic matrices

Consider a stochastic matrix $S = (s_{ij})$ of size n , so that $s_{ij} \geq 0$ and $\sum_i s_{ij} = 1$. Such a matrix can be considered as a collection of n independent probability vectors $s_j = (s_{1j}, \dots, s_{nj})^T$,

$$S = (s_1 \quad \dots \quad s_n) \in \Delta_n^{\times n}, \quad (4.2.1)$$

where $s_{ij} = p(i|j)$ is the probability that an outcome i is obtained given that a classical measurement j was performed. As sketched in Section 4.1, S determines the discrete evolution from $p \in \Delta_n$ to $q \in \Delta_n$ via Eq. (4.1.1), where each convex combination q_i is the outcome probability after concatenating two conditional measurements. Here we will consider the quantum analog of these notions.

4.2.1 Blockwise stochastic matrices and their blockwise product

A collection of n independent measurements with n outcomes can be described by a Cartesian product of n blockwise probability vectors S_j with n effects $\{S_{1j}, \dots, S_{nj}\}$. This suggests considering the following quantum version of a stochastic matrix.

Definition 3 (Blockwise stochastic). *Let S be a square matrix of size $dn \times dn$ composed of n^2 blocks $\{S_{ij}\}$ of size $d \times d$ each. We call S blockwise stochastic if*

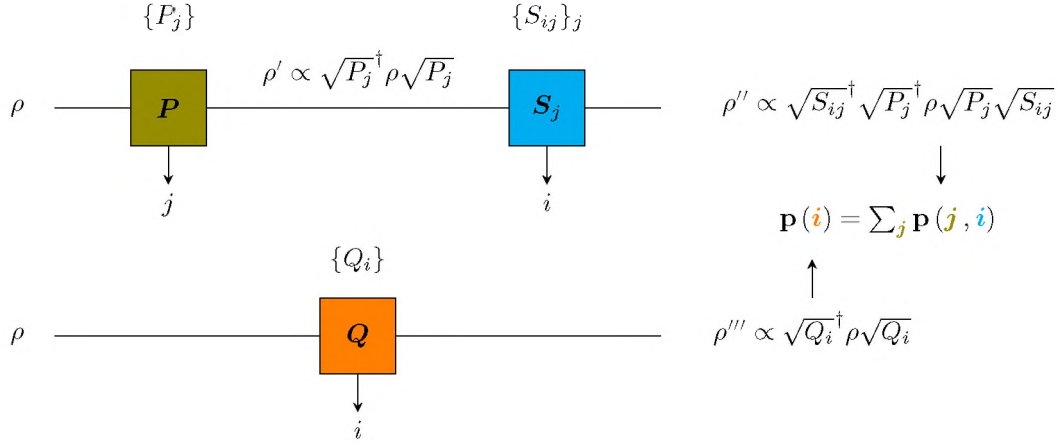


Figure 4.2.1: Circuit description of sequential measurements and measurement interconversion. *Top:* The system is initialized in a quantum state ϱ . A first measurement $\mathbf{P}$ with effects P_j is performed (in green), and depending on the outcome j , a second measurement $\mathbf{S}_j$ with effects S_{ij} is performed (in blue). *Bottom:* A single measurement $\mathbf{Q} = \mathbf{S} * \mathbf{P}$ with effects $Q_i = \sum_j \sqrt{P_j} S_{ij} \sqrt{P_j}$ is performed (in orange) on an initial quantum state ϱ . Both scenarios give equivalent marginal probability distributions of an outcome i .

1. Its entries $\{S_{ij}\}$ are hermitian, positive semi-definite matrices, $S_{ij} \geq 0$, and
2. Its n block-columns of size d sum to the identity, $\sum_i S_{ij} = \mathbb{1}$.

Since a blockwise stochastic matrix $\mathbf{S}$ is a Cartesian product of n blockwise probability vectors, $\mathbf{S} = (\mathbf{S}_1, \dots, \mathbf{S}_n)$, their set can be written as $\Delta_{n,d}^{\times n}$. Similarly as for blockwise probability vectors, the standard notion of a stochastic matrix is recovered by setting $d = 1$, when taking the expectation value $\text{tr}(S_{ij}\varrho) = p(i|j)$, which gives the probability of obtaining an outcome i provided a measurement $\mathbf{S}_j$ was performed.

To establish interconversion rules in the set of quantum measurements, we will consider the quantum version of Eq. (4.1.1), sketched in Figure 4.2.1. First a quantum measurement described by a vector $\mathbf{P} \in \Delta_{n,d}$ is performed. If the outcome is j (i.e. the effect P_j has been applied), then a second measurement $\mathbf{S}_j$ is performed. While in the most general case P_j is an instrument acting on S_{ij} , here we will restrict ourselves to the Lüders assumption of a non-filtering setup [23]. Here the state after measuring $\mathbf{P}$ on an initial state ϱ reads $\varrho' = \sqrt{P_j}\varrho\sqrt{P_j}/\text{tr}(P_j\varrho)$, where positive semidefinite square root $\sqrt{P_j} = P_j^{1/2}$ is the Kraus operator associated to the effect $P_j = \sqrt{P_j}\sqrt{P_j}$ corresponding to an outcome j ; and similarly, the state after the second measurement reads $\varrho'' = \sqrt{S_{ij}}\varrho'\sqrt{S_{ij}}/\text{tr}(\varrho'S_{ij})$ in the event of obtaining an outcome i . Then the probability distribution after this sequence of two measurements can be obtained by an effective quantum measurement, whose effects $\{Q_1, \dots, Q_n\}$ are given by matrix convex combination [187] of $S_{i1}, \dots, S_{in}$ over $P_1, \dots, P_n$, via $Q_i = \sum_{j=1}^n \sqrt{P_j} S_{ij} \sqrt{P_j}$ – see Eq. (4.2.4) below. That is, the probability of obtaining the outcome i reads

$$p(i) = \text{tr}(Q_i\varrho) = \sum_{j=1}^n \text{tr}(\sqrt{S_{ij}}\sqrt{P_j}\varrho\sqrt{P_j}\sqrt{S_{ij}}) = \sum_{j=1}^n p(j, i), \quad (4.2.2)$$

where $p(j, i)$ is the probability of obtaining j in the first measurement and i in the second one in the setup described above.

In the outcome effects Q_i , each term $\sqrt{P_j} S_{ij} \sqrt{P_j}$ can be seen as the *sequential product* [107, 185] or *star product* [197, 198] between the effects P_j and S_{ij} . One verifies that indeed $Q_i = Q_i^\dagger \geq 0$ and $\sum_{i=1}^n Q_i = \mathbb{1}_d$, so that $\mathbf{Q}$ defines a quantum measurement. Therefore, the discrete evolution of quantum measurements introduced in this work carries the following generalization of the matrix

product.

Definition 4 (Blockwise product). Let $\mathbf{A}$ and $\mathbf{B}$ be two matrices $\mathbf{A} = (A_{ij}) \in \mathbb{C}_{dn} \times \mathbb{C}_{dn'}$ and $\mathbf{B} = (B_{ij}) \in \mathbb{C}_{dn'} \times \mathbb{C}_{dn''}$ composed of $n \times n'$ and $n' \times n''$ positive semidefinite blocks of size d , $A_{ij}, B_{ij} \in \mathbb{C}_d \times \mathbb{C}_d$. Let each block B_{kl} decompose as $B_{kl} = \sqrt{B_{kl}}\sqrt{B_{kl}}$. We define the blockwise product,

$$(\mathbf{A} * \mathbf{B})_{ik} = \left(\sum_j \sqrt{B_{jk}} A_{ij} \sqrt{B_{jk}} \right) \in \mathbb{C}_{dn} \times \mathbb{C}_{dn''}. \quad (4.2.3)$$

One verifies that the blockwise product preserves positivity at each block $(\mathbf{A} * \mathbf{B})_{ij}$. Further properties of this product are presented in 4.6.1. In natural analogy to the evolution of classical measurements, the transformed effects Q_i discussed above are obtained with the blockwise product of $\mathbf{S}$ and $\mathbf{P}$,

$$\mathbf{S} * \mathbf{P} = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} * \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \sqrt{P_j} S_{1j} \sqrt{P_j} \\ \vdots \\ \sum_{j=1}^n \sqrt{P_j} S_{nj} \sqrt{P_j} \end{pmatrix} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \end{pmatrix} = \mathbf{Q}, \quad (4.2.4)$$

where it is straightforward to note that $\sum_{i=1}^n Q_i = \mathbf{1}$ as $\mathbf{S}$ is blockwise stochastic.

Example 2. Consider the Pauli projective measurements characterized by blockwise probability vectors

$$\mathbf{P}_Z = \begin{pmatrix} P_{Z+} \\ P_{Z-} \end{pmatrix}, \quad \mathbf{P}_X = \begin{pmatrix} P_{X+} \\ P_{X-} \end{pmatrix}, \quad \mathbf{P}_Y = \begin{pmatrix} P_{Y+} \\ P_{Y-} \end{pmatrix} \quad \text{and} \quad \mathbf{P}_1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \quad (4.2.5)$$

where each $P_{i\pm}$ is the rank-1 projector with eigenvalue ± 1 on the Pauli matrix σ_i , e.g.

$$P_{X\pm} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}. \quad (4.2.6)$$

Three possible measurements can be obtained up to changes of the basis, depending on the alignment between the first and second measurement:

$$(\mathbf{P}_Z, \mathbf{P}_Z) * \mathbf{P}_Z = \mathbf{P}_Z, \quad (4.2.7)$$

$$(\mathbf{P}_Z, \mathbf{P}_Z) * \mathbf{P}_X = (\mathbf{P}_Z, \mathbf{P}_X) * \mathbf{P}_Y = \mathbf{P}_1 \quad \text{or} \quad (4.2.8)$$

$$(\mathbf{P}_Z, \mathbf{P}_X) * \mathbf{P}_X = \begin{pmatrix} P_{X+}/2 \\ \mathbf{1} - P_{X+}/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} P_{X+} \\ P_{X-} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ \mathbf{1} \end{pmatrix}. \quad (4.2.9)$$

4.2.2 Dynamics in the space of quantum measurements

Similarly as for stochastic matrices, the set of blockwise stochastic matrices is closed under the blockwise product $*$ and has a unit element $\mathbf{1}$. This can be formulated as follows.

Proposition 4. The set of blockwise stochastic matrices equipped with the blockwise product, $\{\Delta_{n,d}^{\times n}, *\}$, forms an unital magma, i.e. a closed set with binary operation and neutral element.

Proof. The neutral element is the identity $\mathbf{1}_{dn}$, since $\mathbf{1} * \mathbf{S} = \mathbf{S} * \mathbf{1} = \mathbf{S}$ for any blockwise stochastic matrix $\mathbf{S}$. Moreover, given two blockwise stochastic matrices $\mathbf{S}, \mathbf{T} \in \Delta_{n,d}^{\times n}$, their blockwise product $\mathbf{R} = \mathbf{S} * \mathbf{T}$ is also blockwise stochastic. This is because the product $*$ preserves the positivity of blocks, and the sum over all elements of the k -th column of $\mathbf{R}$ reads

$$\sum_i R_{ik} = \sum_j \sqrt{T_{jk}} \left(\sum_i S_{ij} \right) \sqrt{T_{jk}} = \mathbf{1}. \quad (4.2.10)$$

□

In the classical case, one verifies that given two probability vectors $p, q \in \Delta_n$, there exists a stochastic matrix $A_{p \rightarrow q}$ such that $q = A_{p \rightarrow q} \cdot p$ [183]. One can for example choose $A_{p \rightarrow q} = (q, \dots, q)$ such that all columns are equal to q . Although for a full analysis one needs to consider the blockwise product $*$ and its dual version $*^\dagger$ – see 4.6.1, below we show that this property does not hold for POVM interconversion via Definition 4.

Proposition 5. *Not all transformations between two quantum measurements given by $\mathbf{P} \in \Delta_{n,d}$ and $\mathbf{Q} \in \Delta_{n,d}$ are possible with a blockwise stochastic matrix $\mathbf{S} \in \Delta_{n,d}^{\times n}$ via $\mathbf{Q} = \mathbf{S} * \mathbf{P}$.*

Proof. We will construct a counterexample from the blockwise probability vectors $\mathbf{P} = (P_{Z+}, P_{Z-})^\dagger$ and $\mathbf{Q} = (P_{X+}, P_{X-})^\dagger$, where $P_{X\pm}$ are given by Eq. (4.2.6) and $P_{Z\pm}$ are diagonal projectors. For any blockwise stochastic matrix

$$\mathbf{S} = \begin{pmatrix} \mathbf{S} & \mathbf{T} \\ \mathbf{1} - \mathbf{S} & \mathbf{1} - \mathbf{T} \end{pmatrix}, \quad (4.2.11)$$

one has

$$\mathbf{S} * \mathbf{P} = \begin{pmatrix} P_{Z+} \mathbf{S} P_{Z+} + P_{Z-} \mathbf{T} P_{Z-} \\ P_{Z+} (\mathbf{1} - \mathbf{S}) P_{Z+} + P_{Z-} (\mathbf{1} - \mathbf{T}) P_{Z-} \end{pmatrix}, \quad (4.2.12)$$

where all blocks are diagonal. Therefore one cannot obtain a blockwise probability vector with effects that are not diagonal, e.g. there exists no blockwise stochastic matrix $\mathbf{S}$ such that $\mathbf{S} * \mathbf{P} = (P_{X+}, P_{X-})^\dagger$. $\square$

4.2.3 Regions accessible by the blockwise product

Proposition 8 shows that not all transformations between quantum measurements can be done by conditional concatenation of measurements which simulate the output probabilities. This raises the following question: *What is the set of quantum measurements which can be obtained from an arbitrary blockwise probability vector $\mathbf{P}$ by blockwise product with a blockwise stochastic matrix $\mathbf{S}$?* To answer this question, first note that the set of quantum measurements is convex and characterized by its extremal points, which contain projective measurements $\mathbf{P} = (P_1, \dots, P_n)^\dagger \in \Delta_{n,d}$ with $P_j^2 = P_j$. Based on this fact, the following proposition bounds the set of blockwise probability vectors which can be obtained by a fixed measurement described by $\mathbf{P} \in \Delta_{n,d}$.

Proposition 6. *Let $\mathbf{P} \in \Delta_{n,d}$ be a blockwise probability vector. The following interconversion rules hold.*

1. *The set of measurements $\mathbf{Q}$ which can be obtained by blockwise product $*$ with a blockwise bistochastic matrix is convex.*
2. *The extremal points of this set are obtained by blockwise stochastic matrices, the columns of which form extremal measurements.*

Proof. The proof relies on the fact that the blockwise product $*$ is partially distributive,

$$(\mathbf{S} + \mathbf{S}') * \mathbf{P} = \mathbf{S} * \mathbf{P} + \mathbf{S}' * \mathbf{P}. \quad (4.2.13)$$

For (i) we need to show that if $\mathbf{Q}$ and $\mathbf{Q}'$ can be attained from $\mathbf{P}$ via $\mathbf{S}$ and $\mathbf{S}'$, then $\alpha \mathbf{Q} + (1 - \alpha) \mathbf{Q}'$ with $0 \leq \alpha \leq 1$ can also be attained from $\mathbf{P}$ with a blockwise stochastic matrix. One has by assumption $\mathbf{Q} = \mathbf{S} * \mathbf{P}$ and $\mathbf{Q}' = \mathbf{S}' * \mathbf{P}$, and therefore by Eq. (4.2.13) we have

$$\alpha \mathbf{Q} + (1 - \alpha) \mathbf{Q}' = (\alpha \mathbf{S} + (1 - \alpha) \mathbf{S}') * \mathbf{P} \quad (4.2.14)$$

where $\mathbf{T} := \alpha \mathbf{S} + (1 - \alpha) \mathbf{S}' \in \Delta_{n,d}^{\times n}$ is blockwise stochastic, since the set of blockwise stochastic matrices is convex.

For (ii) we need to show that the extremal points of the set of vectors attainable from a blockwise probability vector $\mathbf{P}$ via $\mathbf{Q} = \mathbf{S} * \mathbf{P}$ are obtainable from a matrix $\mathbf{S} \in \Delta_{n,d}^{\times n}$ such that all its

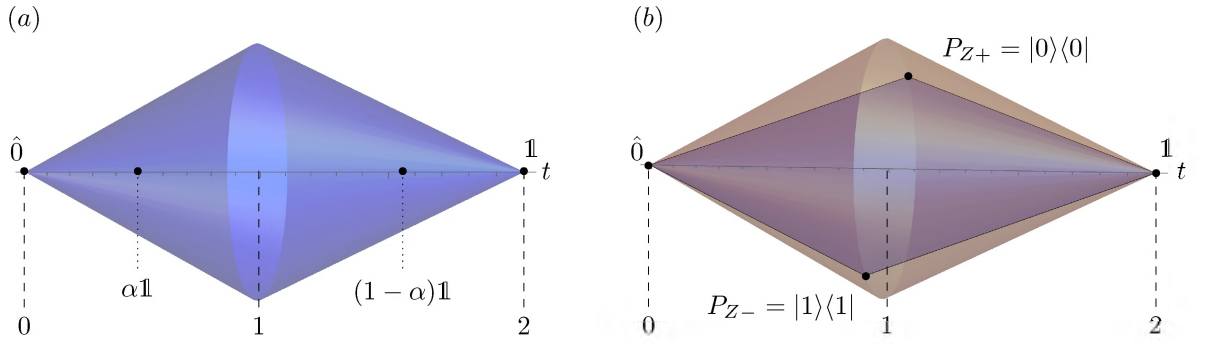


Figure 4.2.2: Regions of the set of measurements $\Delta_{2,2}$ accessible with the blockwise product. As in Fig. 1.2.1, the set of vectors $\mathbf{P} = (P, \mathbb{1} - P)$ defines two anti-parallel cones along the coordinate $t = \text{tr}(P)$. (a) All quantum measurements are accessible from $\mathbf{P}_\alpha = (\alpha\mathbb{1}, (1-\alpha)\mathbb{1})^\dagger$ with a blockwise stochastic matrix. (b) The subset of measurements accessible from $\mathbf{P}_Z = (P_{Z+}, P_{Z-})^\dagger = (|0\rangle\langle 0|, |1\rangle\langle 1|)^\dagger$ is a fragment of a plane obtained by the convex hull of the effects $|0\rangle\langle 0|$, $|1\rangle\langle 1|$, 0 and $\mathbb{1}$.

columns define extremal points in $\Delta_{n,d}$. To see this, it is enough to show that if $\mathbf{Q}$ does not describe an extremal measurement, namely $\mathbf{Q} = \alpha\mathbf{Q}' + (1-\alpha)\mathbf{Q}''$ with $\mathbf{Q}', \mathbf{Q}'' \in \Delta_{n,d}$ and $0 \leq \alpha \leq 1$, then $\mathbf{S} = \alpha\mathbf{S}' + (1-\alpha)\mathbf{S}''$. Indeed, this holds due to Eq. (4.2.13). $\square$

Example 3. Consider the case $n = d = 2$. A trivial case is the set of measurements that can be obtained from $\mathbf{P}_\alpha = (\alpha\mathbb{1}, (1-\alpha)\mathbb{1})^\dagger$. One verifies that by using a blockwise stochastic matrix with equal columns, any measurement is accessible, as shown in Fig. 4.2.2a.

Let us now analyze the set of accessible measurements from $\mathbf{P} = (P, \mathbb{1} - P)^\dagger$ with nonzero projectors $P = P^2$. Writing $\mathbf{P}_Z = (|0\rangle\langle 0|, |1\rangle\langle 1|)^\dagger$ in the eigenbasis of P , we see that one can only obtain measurements with diagonal effects. Therefore, the obtainable set is a fragment of a plane defined by the convex hull of 4 extremal points, depicted in Fig. 4.2.2b.

$$\{\mathbf{S} * \mathbf{P}_Z\}_{\mathbf{S} \in \Delta_{2,2} \times 2} = \text{conv} \left\{ \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix}, \begin{pmatrix} |0\rangle\langle 0| \\ |1\rangle\langle 1| \end{pmatrix}, \begin{pmatrix} |1\rangle\langle 1| \\ |0\rangle\langle 0| \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix} \right\}. \quad (4.2.15)$$

To obtain the full set of measurements, unitary operations $U \in \mathcal{U}(d)$ in the form of $(UPU^\dagger, \mathbb{1} - UPU^\dagger)^\dagger$ are needed in addition to the transformations induced by blockwise product.

Now we will show that the question of what quantum measurements can be obtained from an original quantum measurement $\mathbf{P}$ via the blockwise product of Definition 4, is in fact related to so-called joint measurability of two POVMs, defined as follows.

Definition 5 (Jointly measurable). A quantum measurement with n effects $\{P_j\}$ is compatible or jointly measurable with a quantum measurement with m effects $\{Q_i\}$, if they are both marginals of a mother quantum measurement with $n \times m$ effects $\{M_{ij}\}$, namely

$$P_j = \sum_{i=1}^m M_{ij} \quad \text{and} \quad (4.2.16)$$

$$Q_i = \sum_{j=1}^n M_{ij}. \quad (4.2.17)$$

In the special case of a von Neumann measurement, $\{P_j\}$ and $\{Q_i\}$ are projectors and the corresponding measurements are compatible if and only if the effects commute, $P_j Q_i = Q_i P_j$. Thus

the notion of compatibility assesses when two measurements can be implemented simultaneously (see [199] for a recent review).

The problem of deciding measurement compatibility [200–205] can be related to that of post-processing: In the most general case, it was shown in [206] that two quantum instruments are compatible if and only if one can be obtained by conditional composition with other instruments from each other. The following relation to dynamics through the blockwise product provides a fine-grained version of this result for the case of Lüders effects:

Proposition 7. *Let $\mathbf{P}, \mathbf{Q} \in \Delta_{n,d}$ define two quantum measurements with effects $\{P_j\}$ and $\{Q_i\}$. These two measurements are compatible if and only if there exists a blockwise stochastic matrix $\mathbf{S} \in \Delta_{n,d}^{\times n}$ mapping $\mathbf{P}$ to $\mathbf{Q} = \mathbf{S} * \mathbf{P}$.*

Proof. We will show first the *only if* part, namely that compatibility implies convertibility. Assume $\mathbf{P}$ and $\mathbf{Q}$ define compatible measurements on a system with dimension d , and let $\{M_{ij}\}$ be their mother measurement. Define $S_{ij} := \sqrt{P_j}^{-1} M_{ij} \sqrt{P_j}^{-1}$ where $\sqrt{P_j}$ is the Hermitian square root of P_j and the inverse is taken on the support of $\sqrt{P_j}$. Eq. (4.2.16) implies that

$$\sum_{i=1}^n S_{ij} = \sqrt{P_j}^{-1} \sum_{i=1}^n M_{ij} \sqrt{P_j}^{-1} = \sqrt{P_j}^{-1} P_j \sqrt{P_j}^{-1} = \mathbb{1}, \quad (4.2.18)$$

and therefore $\mathbf{S}$ is blockwise stochastic. On the other hand, Eq. (4.2.17) implies that

$$Q_i = \sum_{j=1}^n M_{ij} = \sum_{j=1}^n \sqrt{P_j} S_{ij} \sqrt{P_j}, \quad (4.2.19)$$

and hence $\mathbf{Q}$ can be obtained from $\mathbf{P}$ via $\mathbf{Q} = \mathbf{S} * \mathbf{P}$.

The *if* part becomes now straightforward: let $\mathbf{S} \in \Delta_{n,d}^{\times n}$ be blockwise stochastic such that $\mathbf{S} * \mathbf{P} = \mathbf{Q}$. Then, the blockwise matrix with entries $M_{ij} = \sqrt{P_j} S_{ij} \sqrt{P_j}$ satisfies Eq. (4.2.16) and (4.2.17) and thus it defines a mother measurement of $\mathbf{P}$ and $\mathbf{Q}$. $\square$

Since compatibility is a symmetric relationship, this implies that all possible transformations between two POVMs $\mathbf{P} \in \Delta_{n,d}$ and $\mathbf{Q} \in \Delta_{n,d}$ can be done in both directions, with two blockwise stochastic matrices $\mathbf{S} \in \Delta_{n,d}^{\times n}$ and $\mathbf{T} \in \Delta_{n,d}^{\times n}$ such that $\mathbf{Q} = \mathbf{S} * \mathbf{P}$ and $\mathbf{P} = \mathbf{T} * \mathbf{Q}$.

4.3 Discrete dynamics induced by blockwise bistochastic matrices

Here we will consider a particular case of the dynamics discussed in the previous section. In the classical case, a stochastic matrix is called bistochastic if both its columns and rows sum to 1. These induce a very special dynamics in the probability simplex and play an important role in physical scenarios [207–209], as they induce growth of entropy in probability vectors.

Here we study the discrete dynamics of blockwise probability vectors induced by quantum versions of bistochastic matrices [92, 186, 191], extending some of the results which are known for $d = 1$ to the case $d \geq 2$.

Definition 6 (Blockwise bistochastic [191]). *Let $\mathcal{B}$ be a square matrix of size $dn \times dn$ composed of n^2 blocks $\{B_{ij}\}$ of size $d \times d$ each. We call $\mathcal{B}$ blockwise bistochastic if*

1. *its entries $\{B_{ij}\}$ are positive semidefinite matrices of size $d \geq 2$*
2. *its blockwise columns and rows sum to identity,*

$$\sum_i B_{ij} = \sum_j B_{ij} = \mathbb{1}_d. \quad (4.3.1)$$

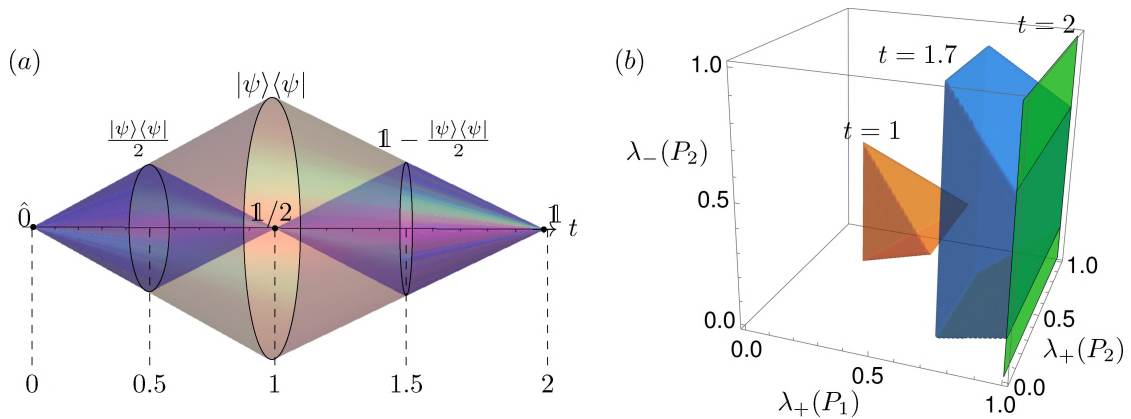


Figure 4.3.1: Visualization of the set of sortable blockwise probability vectors corresponding to quantum measurements of small system sizes. (a) Three-dimensional dual cone containing the real blockwise vectors $\mathbf{P} = (P, \mathbb{1} - P)^T \in \Delta_{2,2}$ (the imaginary component is omitted) with axial coordinate $t = \text{tr}(P)$, shown in brown. The blue region contains the set of sortable vectors. (b) Three-dimensional cube of vectors $(P_1, P_2, \mathbb{1} - P_1 - P_2)^\dagger \in \Delta_{3,2}$ with commutative entries, $P_1 P_2 = P_2 P_1$ with a fixed parameter $\text{tr}(P_1) = 1$ and a variable parameter $t = \text{tr}(P_2)$. The axes are the maximum eigenvalue of P_1 , $\lambda_+(P_1)$, and the maximum and minimum eigenvalues of P_2 , $\lambda_+(P_2)$ and $\lambda_-(P_2)$. The orange, blue and green regions contain the sets of sortable vectors for $t = 1$, $t = 1.7$ and $t = 2$ respectively.

Here we will denote the set of blockwise bistochastic matrices composed of n^2 square blocks of size d as $\mathcal{B}_{n,d}$. Similarly as in the definitions of blockwise probability vector and blockwise stochastic matrix, we recover the standard notion of *bistochastic matrix* if $\{B_{ij}\}$ are positive real numbers (setting $d = 1$) by selecting any state $\varrho \in \Omega_d$ and considering entry-wise expectation values $\text{tr}(B_{ij}\varrho)$. The simplest non-classical case is $n = 2$,

$$B = \begin{pmatrix} B & \mathbb{1} - B \\ \mathbb{1} - B & B \end{pmatrix}, \quad (4.3.2)$$

where B is any Hermitian matrix of size $d \times d$ satisfying $0 \leq B \leq \mathbb{1}$, and hence we have $\mathcal{B}_{2,d} \Delta_{2,d}$. In particular, for $d = n = 2$, the set of blockwise bistochastic matrices is inherently described by Figure 1.2.1.

4.3.1 Bistochastic dynamics and partial order between blockwise probability vectors

Let us first recall the notion of extremal blockwise probability vector $\mathbf{V}_j = v_j \otimes \mathbb{1}$ with a single nonzero component, as defined in equations (1.2.22) and (1.2.26). In the probability simplex, any probability vector can be obtained by matrix multiplication of a bistochastic matrix times an extremal probability vector v_i . In the set of blockwise probability vectors a similar result holds, as shown below.

Observation 1. *Let $\mathbf{P} \in \Delta_{n,d}$. For any $1 \leq j \leq n$, there exists a blockwise bistochastic matrix $B \in \mathcal{B}_{n,d}$ such that*

$$B * \mathbf{V}_j = \mathbf{P}. \quad (4.3.3)$$

Proof. Any blockwise probability vector $\mathbf{P} = (P_1, \dots, P_n)^\dagger$ can be obtained from $\mathbf{V}_1 = (\mathbb{1}, 0, \dots, 0)^\dagger$

as follows,

$$\begin{pmatrix} P_1 & P_2 & \cdots & P_d \\ P_2 & P_3 & \cdots & P_1 \\ \vdots & \vdots & & \vdots \\ P_d & P_1 & \cdots & P_{d-1} \end{pmatrix} * \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_d \end{pmatrix}. \quad (4.3.4)$$

A similar construction does the job for any other vector $\vec{V}_j$. $\square$

Since a blockwise bistochastic matrix is also blockwise stochastic, the blockwise product by a blockwise bistochastic matrix also preserves the set of measurements $\Delta_{n,d}$. Therefore the product of two blockwise bistochastic matrices is blockwise stochastic, although in general it is not blockwise bistochastic. This is the case only for $n = 2$ and arbitrary d , and if all entries to be multiplied via the sequential product [107] in the form of $\sqrt{XY}\sqrt{X}$ commute. Unlike the standard matrix multiplication, the blockwise product $*$ is not associative, and thus

$$\mathcal{B} * (\mathcal{B}' * P) \neq (\mathcal{B} * \mathcal{B}') * P, \quad (4.3.5)$$

see 4.6.1. This means that the discrete dynamics of a quantum measurement $P \in \Delta_{n,d}$ by blockwise product by subsequent blockwise bistochastic matrices $\mathcal{B}, \mathcal{B}' \in \mathcal{B}_{n,d}$ in general cannot be obtained by a global matrix $\mathcal{B} * \mathcal{B}'$. Due to this fact, the fixed points under blockwise bistochastic dynamics are non-unique (see 4.6.2).

Standard bistochastic matrices induce special dynamics in the classical probability simplex. If B is bistochastic and p is a probability vector, then the probability vector $q = B \cdot p$ is majorized by p , written $p \succ q$ (see Theorem II.1.9 in [183]), which means that

$$\sum_{j=1}^k p_j^\downarrow \geq \sum_{j=1}^k q_j^\downarrow \quad \text{for } k = 1, \dots, n, \quad (4.3.6)$$

where the superindex $\downarrow$ denotes non-increasing order of the components. This induces a partial order in the probability simplex Δ_n which plays a key role in quantum information tasks, such as entanglement transformations of pure states [207] or mixed state interconversion under incoherent operations [210, 211].

Although standard probabilities can always be sorted nonincreasingly, this is not the case for effects of quantum measurements, which are positive semidefinite operators in dimension $d \geq 2$. Therefore, to compare quantum measurements in an analogous way as in the classical case, the following definition needs to be specified.

Definition 7 (Sortable quantum measurement). *A quantum measurement of n effects $P_1, \dots, P_n$ acting on a d -dimensional system is called sortable if its effects can be sorted in the Löwner order [212],*

$$1_d \geq P_1 \geq P_2 \geq \cdots \geq P_n \geq 0. \quad (4.3.7)$$

In Fig. 4.3.1 we plot the set of blockwise probability vectors which can be ordered in such way for small system sizes. Sortability of a quantum measurement is physically determined by its output probability distribution with respect to any state of the system, as follows.

Observation 2. *A quantum measurement is sortable if and only if regardless of the state $\varrho \in \Omega_d$ of the system, the probability of obtaining an outcome j is larger than the probability of obtaining an outcome $j + 1$.*

This becomes evident since each difference between effects $P_j - P_{j+1}$ is positive semidefinite if and only if for any state ϱ we have

$$0 \leq \text{tr}((P_j - P_{j+1})\varrho) = \text{tr}(P_j\varrho) - \text{tr}(P_{j+1}\varrho). \quad (4.3.8)$$

4.3.2 Majorization relation between quantum measurements

The Birkhoff - von Neuman theorem [180] establishes that a matrix is bistochastic, if and only if it is a convex combination of permutation matrices, and thus characterizes the set of bistochastic matrices. Looking for a quantum generalization for blockwise bistochastic matrices, it was shown by De les Coves and Netzer [186] that this is not the case, namely the quantum set of blockwise bistochastic matrices is strictly larger. However, a well-known equivalent formulation of the Birkhoff - Von Neumann theorem, originally stated by Ostrowski [181] and captured in standard books [182, 183], characterizes bistochastic matrices in the following way:

Theorem 4 ([181]). *A stochastic matrix B is bistochastic if and only if any probability vector $p \in \Delta_{n,1}$ majorizes its image, $p \succ Bp$.*

This formulation allows us to characterize blockwise bistochastic matrices as follows, in a quantum analogy to Theorem 4.

Theorem 5 (State-independent operator majorization). *Discrete dynamics in the set of blockwise probability vectors satisfies the following equivalence.*

1. *Let $\mathcal{B} \in \mathcal{B}_{n,d}$ be a blockwise bistochastic matrix. Let $\mathbf{P} \in \Delta_{n,d}$ describe a sortable quantum measurement with image $\mathbf{Q} = \mathcal{B} * \mathbf{P} \in \Delta_{n,d}$. Then the vectors $\mathbf{P}$ and $\mathbf{Q}$ satisfy the majorization relation $\mathbf{P} \succ \mathbf{Q}$, written*

$$\sum_{j=1}^k P_j \geq \sum_{j=1}^k Q_j \quad (4.3.9)$$

in the Löwner order, for all $1 \leq k \leq n$, with equality for $k = n$.

2. *Conversely, a matrix $\mathcal{B} \in \Delta_{n,d}^{\times n}$ is blockwise bistochastic if Eq. (4.3.9) holds for any sortable quantum measurement $\mathbf{P} \in \Delta_{n,d}$ and for any ordering of the effects of $\mathbf{Q} = \mathcal{B} * \mathbf{P} \in \Delta_{n,d}$.*

Note that in the classical case $d = 1$, the above statement reduces to Theorem 4. In the quantum case, the set of sortable quantum measurements plays a crucial role in characterizing blockwise bistochastic matrices, although its volume shrinks significantly when the number of effects and dimensions increase.

Proof. We will follow a similar structure as in the proof of Theorem II.1.9 in [183].

(i) Let us first prove that if $\mathcal{B} \in \mathcal{B}_{n,d}$, then Eq. (4.3.9) holds. For convenience, let us define

$$\Theta = \sum_{j=1}^k P_j - Q_j \quad (4.3.10)$$

and $B_j = \sum_{i=1}^k B_{ij}$, which satisfies

$$\sum_{j=1}^n B_j - k\mathbf{1} = 0. \quad (4.3.11)$$

We want to show that the operator Θ is positive semidefinite. For that, we introduce Eq. (4.3.11) and obtain

$$\begin{aligned} \Theta &= \sum_{j=1}^k P_j - \sum_{j=1}^n \sqrt{P_j} B_j \sqrt{P_j} + \sqrt{P_k} \left(\sum_{j=1}^n B_j - k\mathbf{1} \right) \sqrt{P_k} \\ &= \sum_{j=1}^k \sqrt{P_j} (\mathbf{1} - B_j) \sqrt{P_j} - \sqrt{P_k} (\mathbf{1} - B_j) \sqrt{P_k} \\ &\quad + \sum_{j=k+1}^n \sqrt{P_k} B_j \sqrt{P_k} - \sqrt{P_j} B_j \sqrt{P_j}. \end{aligned} \quad (4.3.12)$$

We wish to show that the minimal eigenvalue of Θ is nonnegative. On one hand, as a special case of Weyl's inequalities [213], we have that given two Hermitian matrices A and B , the minimal eigenvalue of their sum satisfies the bound

$$\lambda_{\min}(A + B) \geq \lambda_{\min}(A) + \lambda_{\min}(B). \quad (4.3.13)$$

On the other hand, recall the sequential product $A \circ B = \sqrt{A}B\sqrt{A}$ defined in [107]. Since by assumption $P_a - P_{b \geq a} \geq 0$, due to Observation 3 the spectrum of $(P_a - P_{b \geq a}) \circ (\mathbb{1} - B_c)$ is the same as the spectrum of $(\mathbb{1} - B_c) \circ (P_a - P_{b \geq a})$. Combining these two facts, we have

$$\lambda_{\min}(\Theta) \geq \sum_{j=1}^k \lambda_{\min}((\mathbb{1} - B_j) \circ (P_j - P_k)) + \sum_{j=k+1}^n \lambda_{\min}(B_j \circ (P_k - P_j)). \quad (4.3.14)$$

Since the sequential product preserves positivity, all terms above are nonnegative and we have $\lambda_{\min}(\Theta) \geq 0$, which implies that Θ is positive semidefinite. This proves item (I).

(ii) Now we will show the converse direction, namely that if Eq. (4.3.9) holds for all sortable vectors $\mathbf{P} \in \Delta_{n,d}$ satisfying (4.3.7), then $\mathcal{B}$ is blockwise bistochastic. For that we recall that Eq. (4.3.9) needs to be fulfilled in particular for $\mathbf{V}_j = (0, \dots, 0, \mathbb{1}, 0, \dots, 0)$ with $\mathbb{1}$ at the j -th position and 0 elsewhere. This implies that

$$\sum_{i=1}^k B_{ij} \leq \mathbb{1} \quad (4.3.15)$$

with equality for $k = n$, which implies that $\mathcal{B}$ is blockwise stochastic. Eq. (4.3.9) needs to be fulfilled also for $\mathbf{P}_u = (\mathbb{1}, \dots, \mathbb{1})/n$. This implies that

$$\frac{1}{n} \sum_{i=1}^k \left(\sum_{j=1}^n B_{ij} \right) \leq \frac{k}{n} \mathbb{1}, \quad (4.3.16)$$

with equality for $k = n$. Now assume that for some k , Eq. (4.3.16) holds with a strict inequality. Then it cannot hold with equality for $k = n$, which is a contradiction. Therefore, the blockwise rows of $\mathcal{B}$ sum exactly to identity and thus $\mathcal{B}$ is blockwise bistochastic. $\square$

The theorem above can physically be interpreted as in the following Corollary.

Corollary 3. *Suppose a quantum system is prepared in a state ϱ . A sortable measurement $\mathbf{P}$ is performed with outcome probability distribution p , and then a second measurement $\mathbf{S}_j$ is performed depending on the outcome j , with outcome probability distribution q (see Fig. 4.2.1). The following statements are equivalent.*

1. *For any quantum state ϱ , the output probability vector q can be obtained with classical post-processing from the vector p as*

$$q = Bp, \quad (4.3.17)$$

where B is a standard bistochastic matrix of order n .

2. *For any initial state ϱ , the first and second probability distributions p and q satisfy the majorization relation $p \succ q$.*
3. *The collection of the i -th effect of each of the possible second measurements $\mathbf{S}_j$, $\{S_{i1}, \dots, S_{in}\}$, defines a quantum measurement.*

Proof. (i) $\iff$ (ii) It is a standard result in classical probability theory that there exists a bistochastic matrix B such that $q = Bp$, if and only if the majorization relation $p \succ q$ holds [183].

(ii) $\iff$ (iii) The majorization relation $p \succ q$ can be written as

$$\sum_{j=1}^k \text{tr}(P_j \varrho) \geq \sum_{j=1}^k \text{tr}(Q_j \varrho), \quad (4.3.18)$$

which holds for any state ϱ if and only if the majorization relation $\mathbf{P} \succ \mathbf{Q}$ of Eq. (4.3.9) holds, where $\mathbf{P}$ is sortable by assumption. By Theorem 5, this holds for any state ϱ and measurement $\mathbf{P}$ if and only if $\mathbf{S} = (\mathbf{S}_1, \dots, \mathbf{S}_n)$ is blockwise bistochastic, which by definition means that the sets $\{S_{1j}, \dots, S_{nj}\}$ and $\{S_{i1}, \dots, S_{in}\}$ are quantum measurements for all i and j . Whereas $\{S_{1j}, \dots, S_{nj}\}$ is a quantum measurement for all j by the stochasticity assumption, the fact that $\{S_{i1}, \dots, S_{in}\}$ is a measurement for all i is exactly the additional condition for bistochasticity. $\square$

By arguing in an analogous way as in Theorem 5, one can establish a state-dependent majorization relation under blockwise bistochastic dynamics.

Theorem 6 (State-dependent operator majorization). *Let $\mathbf{P} \in \Delta_{n,d}$, $\mathcal{B} \in \mathcal{B}_{n,d}$ and $\mathbf{Q} = \mathcal{B} * \mathbf{P}$. Given a quantum state $\varrho \in \Omega_d$, assume that $\mathbf{P}$ can be ordered such that for any $s \geq r$ the minimal eigenvalues $\lambda_{\min}$ satisfy the inequality*

$$\lambda_{\min}(\sqrt{P_r} \varrho \sqrt{P_r}) \geq \lambda_{\min}(\sqrt{P_s} \varrho \sqrt{P_s}) \geq 0. \quad (4.3.19)$$

Then the classical probability vectors with components $p_j = \text{tr}(P_j \varrho)$ and $q_i = \text{tr}(Q_i \varrho)$ satisfy the standard majorization relation $p \succ q$.

Note that Theorem 6 holds in particular if $\sqrt{P_i} \varrho \sqrt{P_i} \geq \sqrt{P_{j \leq i}} \varrho \sqrt{P_{j \leq i}}$, which is a state-dependent analogue to the condition for Theorem 5.

Proof. Analogously as in Theorem 5, we want to show that the real number

$$\theta = \sum_{j=1}^k \text{tr}(P_j \varrho) - \text{tr}(Q_j \varrho) \quad (4.3.20)$$

is nonnegative. In analogy to Eq. (4.3.12) concerning the operator Θ , here we obtain the following expression for the scalar θ ,

$$\begin{aligned} \theta = \text{tr} \left(\sum_{j=1}^k (\mathbb{1} - B_j) \left(\sqrt{P_j} \varrho \sqrt{P_j} - \sqrt{P_k} \varrho \sqrt{P_k} \right) \right. \\ \left. + \sum_{j=k+1}^n B_j \left(\sqrt{P_k} \varrho \sqrt{P_k} - \sqrt{P_j} \varrho \sqrt{P_j} \right) \right), \end{aligned} \quad (4.3.21)$$

where $B_j = \sum_{i=1}^k B_{ij}$. To proceed we need the property that given an Hermitian matrix X and a positive semidefinite matrix Y , we have

$$\text{tr}(XY) = \text{tr}(\sqrt{Y} X \sqrt{Y}) \geq \text{tr}(\sqrt{Y} \mathbb{1} \lambda_{\min}(X) \sqrt{Y}) = \text{tr}(Y) \lambda_{\min}(X). \quad (4.3.22)$$

This property can be applied as follows. Since $\mathcal{B}$ is blockwise bistochastic, $(\mathbb{1} - B_j)$ is positive semidefinite. Since $\sqrt{P_i} \varrho \sqrt{P_i}$ is positive semidefinite for any i , the operator $\sqrt{P_k} \varrho \sqrt{P_k} - \sqrt{P_j} \varrho \sqrt{P_j}$

is Hermitian and therefore Eq. (4.3.22) applies. Thus we have

$$\begin{aligned} \theta \geq \sum_{j=1}^k \operatorname{tr}(\mathbb{1} - B_j) \lambda_{\min} \left(\sqrt{P_j} \varrho \sqrt{P_j} - \sqrt{P_k} \varrho \sqrt{P_k} \right) \\ + \sum_{j=k+1}^n \operatorname{tr}(B_j) \lambda_{\min} \left(\sqrt{P_k} \varrho \sqrt{P_k} - \sqrt{P_j} \varrho \sqrt{P_j} \right). \end{aligned} \quad (4.3.23)$$

Using Eq. (4.3.13) again we can further bound θ from below by splitting the terms as

$$\begin{aligned} \theta \geq \sum_{j=1}^k \operatorname{tr}(\mathbb{1} - B_j) \left(\lambda_{\min} \left(\sqrt{P_j} \varrho \sqrt{P_j} \right) - \lambda_{\min} \left(\sqrt{P_k} \varrho \sqrt{P_k} \right) \right) \\ + \sum_{j=k+1}^n \operatorname{tr}(B_j) \left(\lambda_{\min} \left(\sqrt{P_k} \varrho \sqrt{P_k} \right) - \lambda_{\min} \left(\sqrt{P_j} \varrho \sqrt{P_j} \right) \right). \end{aligned} \quad (4.3.24)$$

By assumption $\mathbf{P}$ is ordered according to Eq. (4.3.19) and therefore all factors are nonnegative. $\square$

4.4 Resource theory of quantum measurements

Resource theories are abstract frameworks which characterize the possible transformations within physical scenarios. Any resource theory has three main ingredients [184, 214]: (i) a set of free elements, (ii) a set of transformations that leave invariant the free elements, and (iii) a set of resourceful elements. It is often useful to introduce two extra ingredients: (iv) the subset of maximally resourceful elements, from which any other element can be obtained with free operations, and (v) a monotone, which is a quantity associated to each element that cannot increase under free operations, and thus it quantifies how resourceful a given element is.

Recently, resource theories became a useful tool to analyze the set of POVMs [103, 215–217]. Here we show that the framework introduced in this work allows to establish an alternative resource theory, by making use of generalized notions of entropy and majorization. Operationally, the standard resource theory in the probability simplex Δ_n is obtained by considering the expectation value of quantum effects with respect to the state of the quantum system in hand.

To describe the discrete dynamics within Δ_n , which elements are column probability vectors, the resource theory of majorization is well established. The free elements (i) are the uniform vectors with equal components, the free operations (ii) are bistochastic matrices equipped with the matrix product on probability vectors, the set of resourceful elements (iii) are the probability vectors which are not uniform, the extremal vectors $\{v_i\}$ defined in (1.2.22) (iv) are maximally resourceful, and the monotones (v) are the so-called *cumulative probability distributions* of a probability vector sorted in non-increasing order, $\mu_k = \sum_{j=1}^{k \leq n} p_j^\downarrow$.

The quantum case has a more involved structure, in several senses: the extremal points lie in a continuous high-dimensional hypersurface (see [100, 108]); the elements which remain invariant by a given free operation are not straightforward to study (see 4.6.2); and several functionals might be chosen to describe dynamics (see Fig. 4.6.1 and 4.6.3). Nonetheless, a resource theory can be established naturally as follows.

- (i-ii) The set of blockwise bistochastic matrices $\mathcal{B}_{n,d}$ describes the free operations and the vector $\mathbf{P}_u = (\mathbb{1}_d/n, \dots, \mathbb{1}_d/n)^\dagger \in \Delta_{n,d}$ describes the free element for each n and d , since for any $\mathbf{B} \in \mathcal{B}_{n,d}$ it holds that

$$\mathbf{B} * \mathbf{P}_u = \mathbf{P}_u, \quad (4.4.1)$$

and therefore with $\mathcal{B}_{n,d}$ one cannot reach any point outside of $\{\mathbf{P}_u\}$.

- (iii) From any point $\mathbf{P} \in \Delta_{n,d}$ one can reach a free element $\mathbf{P}_u$ by choosing $\mathbf{B}_u \in \mathcal{B}_{n,d}$ defined component wise by $(B_u)_{ij} = \mathbb{1}_d/n$ as

$$\mathbf{B}_u * \mathbf{P} = \mathbf{P}_u. \quad (4.4.2)$$

Thus the set of resourceful states can be identified with $\Delta_{n,d} \setminus \{\mathbf{P}_u\}$.

- (iv) The maximally resourceful elements with respect to the blockwise product are those blockwise probability vectors $\mathbf{P} \in \Delta_{n,d}$ with the property that for any $\mathbf{Q} \in \Delta_{n,d}$, there exists $\mathcal{B} \in \mathcal{B}_{n,d}$ such that

$$\mathbf{Q} = \mathcal{B} * \mathbf{P}. \quad (4.4.3)$$

According to Observation 1, the vectors $\{\mathbf{V}^{(i)}\}$ defined in Eq. (1.2.22) are maximally resourceful. This stands in contraposition to the resource theory introduced in [103], where $\{\mathbf{V}^{(i)}\}$ are precisely the free elements.

- (v) Theorem 5 provides us with a natural state-dependent monotone E_ϱ , for sortable quantum measurements, which is defined as

$$E_\varrho(\mathbf{P}) = \sum_{j=1}^n \log(\text{tr}(P_j \varrho)) + n \log n, \quad (4.4.4)$$

and a natural state-independent monotone,

$$E = \min_{\varrho \in \Omega_d} (E_\varrho(\mathbf{P})). \quad (4.4.5)$$

This is shown in the following lemma.

Lemma 4. *Let $\mathbf{P} \in \Delta_{n,d}$ be a sortable blockwise probability vector and let $\mathcal{B} \in \mathcal{B}_{n,d}$ be blockwise bistochastic. Let $\mathbf{Q} = \mathcal{B} * \mathbf{P}$. For any quantum state $\varrho \in \Omega_d$, the entropy E_ϱ satisfies*

$$E_\varrho(\mathbf{P}) \geq E_\varrho(\mathbf{Q}). \quad (4.4.6)$$

Moreover, one has $E_\varrho(\mathbf{P}_u) = 0$ for any quantum state ϱ .

Proof. By Theorem 5 and Corollary 3, the vector $p = (\text{tr}(P_1 \varrho), \dots, \text{tr}(P_n \varrho))$ majorizes the vector $q = (\text{tr}(Q_1 \varrho), \dots, \text{tr}(Q_n \varrho))$, written $p \succ q$ – see Eq. (4.3.6). This implies [183] that the standard entropy of p is smaller than the entropy of q , which implies that $E_\varrho(\mathbf{P}) \geq E_\varrho(\mathbf{Q})$ for any state ϱ . Therefore, the quantity E_ϱ defined in Eq. (4.4.4) cannot increase under blockwise product by a blockwise bistochastic matrix.

It is straightforward to see that $E_\varrho(\mathbf{P}_u) = -n \log n + n \log n = 0$ for the (sortable) free measurement $\mathbf{P}_u = (\mathbb{1}, \dots, \mathbb{1})^T/n$ defined in (1.2.25). Since the results hold for any state ϱ , they also hold for the tightest case of Eq. (4.4.6) obtained by minimizing the monotone over ϱ . $\square$

A possible state-independent majorization relation for non-sortable quantum measurements based on non-linear quantities is analyzed in 4.6.3.

4.5 Concluding remarks

We introduced a framework to analyze dynamics in the set $\Delta_{n,d}$ of quantum measurements, such that the outcome probabilities can be simulated by concatenation with a conditional measurement. For that we consider a blockwise product with blockwise stochastic matrices, which define collections of quantum measurements. This approach, inspired by a work of Gudder [93], can be seen as a generalization of discrete dynamics of probability distributions induced by stochastic matrices.

Unlike in the classical case, transformations induced in this way are possible only when the input and output measurements are compatible.

Special dynamics is obtained if considering blockwise bistochastic matrices, where all columns and rows define a quantum measurement. We have seen that, as in the classical case, these induce an operator majorization relation between quantum measurements. This generalization of central results in standard majorization theory characterizes the set of blockwise bistochastic matrices [186, 191] and can be understood in terms of classical postprocessing of a sequence of quantum measurements. The proposed approach allows to establish a resource theory of quantum measurements, which can be seen as a noncommutative version of the theory of majorization in the probability simplex. These results are summarized in Table 4.1.

Several questions and possible applications remain open to further study. What extra operations are needed to fully describe transformations between incompatible measurements, is at the moment unknown. Moreover, further studies are needed to understand the structure of the set of blockwise probability vectors $\Delta_{n,d}$ and relevant subsets, as well as the structure of the set of blockwise bistochastic matrices $\mathcal{B}_{n,d}$. On the fundamental side, the systematic framework for measurement interconversion introduced here is suitable to study contextuality in sequential measurements [218–220]. For quantum protocols, one could engineer an optimal algorithmic sequence of conditional measurements to obtain particularly interesting measurements, such as informationally complete ones, with certain precision.

4.6 Further results

When the research described in this chapter was undertaken, several additional results were found. These are not strictly needed for the chapter itself, but provide a further comprehension and complement the main results therein. This additional work is summarized in the present section, which includes a study of the properties of the $*$ product and its dual version, an analytical and numerical analysis of quantum measurements that are invariant under the action of blockwise bistochastic matrices, and a conjectured monotone for the blockwise bistochastic dynamics of quantum measurements. The conjecture is analytically proven for measurements with two outcomes and numerically supported for more general cases.

4.6.1 Properties of the blockwise product and its dual version

Here we will discuss the properties of the blockwise product $*$ introduced in Definition 4, which mathematically has a natural dual version $*^\dagger$ upon conjugate transposition.

Unitary invariance of the sequential product

A key element of the blockwise product of Definition 4 is the sequential product between two positive semidefinite matrices $A \circ B = \sqrt{A}B\sqrt{A}$ introduced in [107], where $\sqrt{A}$ is the unique positive square root of A . This product satisfies the following property, which we use to prove Theorem 5.

Observation 3. *The sequential products $A \circ B$ and $B \circ A$ have the same spectrum.*

This should not be confused with the fact that $A \geq A \circ B$ but $A \not\geq B \circ A$, pointed out in [107].

Proof. Defining $X = \sqrt{A}\sqrt{B}$, we have $A \circ B = XX^\dagger$ and $B \circ A = X^\dagger X$. By computing the singular value decomposition $X = UDV$ where U and V are unitary and D is diagonal, we have

$$XX^\dagger = UDVV^\dagger D^*U^\dagger = UDD^*U^\dagger \quad (4.6.1)$$

and

$$X^\dagger X = V^\dagger D^*U^\dagger UDV = V^\dagger D^*DV. \quad (4.6.2)$$

n -point probability simplex Δ_n	Set of quantum measurements $\Delta_{n,d}$
Probability vector $p = (p_1, \dots, p_n)^T \in \Delta_n$ $p_j \geq 0, \quad \sum_{j=1}^n p_j = 1$	Blockwise probability vector (POVM) [92] $\mathbf{P} = (P_1, \dots, P_n)^\dagger \in \Delta_{n,d}$ $P_j \geq 0, \quad \sum_{j=1}^n P_j = \mathbb{1}_d$
Stochastic matrix $S = (s_1, \dots, s_n) \in \Delta_n^{\times n}$ $s_j = (s_{1j}, \dots, s_{nj}) \in \Delta_n$	Blockwise stochastic matrix $\mathbf{S} = (\mathbf{S}_1, \dots, \mathbf{S}_n) \in \Delta_{n,d}^{\times n}$ $\mathbf{S}_j = (S_{1j}, \dots, S_{nj}) \in \Delta_{n,d}$
Transformations within Δ_n $Sp = q$ $q_i = \sum_{j=1}^n s_{ij} p_j$	Transformations within $\Delta_{n,d}$ $\mathbf{S} * \mathbf{P} = \mathbf{Q}$ $Q_i = \sum_{j=1}^n \sqrt{P_j} S_{ij} \sqrt{P_j}$
Allowed transformations $p \xrightarrow{S} q \in \Delta_n$ always possible.	Allowed transformations $\mathbf{P} \xrightarrow{\mathbf{S}} \mathbf{Q} \in \Delta_{n,d} \iff$ Jointly measurable
Bistochastic $B = (b_{ij})$ $b_{ij} \geq 0; \quad \sum_i b_{ij} = \sum_j b_{ij} = 1$	Blockwise bistochastic [186, 191] $\mathbf{B} = (B_{ij})$ $B_{ij} \geq 0; \quad \sum_{i=1}^n B_{ij} = \sum_{j=1}^n B_{ij} = \mathbb{1}$
Sortability Nonincreasing order, $1 \geq p_1 \geq \dots \geq p_n \geq 0$	Sortability Sortable subset, $\mathbb{1} \geq P_1 \geq \dots \geq P_n \geq 0$
Majorization for all $p \in \Delta_n$ $p \succ q = Bp$: $\sum_{j=1}^k p_j \geq \sum_{j=1}^k q_j$	Majorization for $\mathbf{P} \in \text{sortable} \subset \Delta_{n,d}$ $\mathbf{P} \succ \mathbf{Q} = \mathbf{B} * \mathbf{P}$: $\sum_{j=1}^k P_j \geq \sum_{j=1}^k Q_j$

Table 4.1: Comparison of the main elements of classical probability theory and their generalization for the set of quantum measurements. Since in the quantum case one considers positive operators instead of nonnegative real numbers, some of the connections established in the classical case need additional structure to hold in the quantum case.

Therefore, we have $A \circ B = UV(B \circ A)V^\dagger U^\dagger$. Since UV is unitary and $D^*D = DD^*$ (as D is diagonal), both products have the same spectra. $\square$

The blockwise product and its dual definition

Here we will analyze a dual version of the blockwise product $*$, which in the classical case cannot be identified. To this end, we will use the fact that given two blockwise probability vectors $\mathbf{P}, \mathbf{Q} \in \Delta_{n,d}$, their adjoints $\mathbf{P}^\dagger$ and $\mathbf{Q}^\dagger$ are row vectors. We recall that this is because the blocks P_j and Q_j are Hermitian and thus the adjoint transposes the position the block components, while leaving the blocks invariant. Similarly, the adjoint $\mathbf{S}^\dagger$ defines a blockwise matrix with positive semidefinite entries which sum to identity row-wise.

Let us now take a closer look at the blockwise product $*$, in comparison to its classical version. While in the classical case one obtains the same result choosing $q = Sp$ or $q^T = p^T S^T$ up to transposition, in the quantum regime one obtains different results with $\mathbf{Q} = \mathbf{S} * \mathbf{P}$ and $\mathbf{Q}^\dagger = \mathbf{P}^\dagger * \mathbf{S}^\dagger$. This is illustrated in Proposition 8 below, which is a dual analog to Propositions 4 and 5. For completeness, we define the following dual version of the blockwise product introduced in Definition 4,

$$\mathbf{A} *^\dagger \mathbf{B} := (\mathbf{B}^\dagger * \mathbf{A}^\dagger)^\dagger, \quad (4.6.3)$$

which acts as a sequential product in the form of $\sqrt{A_{ij}}B_{jk}\sqrt{A_{ij}}$ as demonstrated below in Eq. (4.6.7).

Proposition 8. *The following properties hold in the set of measurements $\Delta_{n,d}$ for the dual blockwise product $*^\dagger$:*

1. *Given a blockwise probability vector $\mathbf{P} \in \Delta_{n,d}$ and a blockwise stochastic matrix $\mathbf{S} \in \Delta_{n,d}^{\times n}$, let $\mathbf{Q}$ be defined as $\mathbf{Q} = \mathbf{S} *^\dagger \mathbf{P}$. Then, $\mathbf{Q}$ is not necessarily a blockwise probability vector.*
2. *Let $\mathbf{P}$ and $\mathbf{Q}$ be two blockwise probability vectors, $\mathbf{P}, \mathbf{Q} \in \Delta_{n,d}$. There exists a blockwise stochastic matrix $\mathbf{S} \in \Delta_{n,d}^{\times n}$ such that $\mathbf{Q} = \mathbf{S} *^\dagger \mathbf{P}$.*

Proof. A counterexample for (i) can be found by defining the blockwise stochastic matrix $\mathbf{S}$ and blockwise probability vector $\mathbf{P}$ as

$$\mathbf{S} = \begin{pmatrix} P_{Z+} & P_{X+} \\ P_{Z-} & P_{X-} \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} \alpha P_{X+} + \beta P_{X-} \\ (1-\alpha)P_{X+} + (1-\beta)P_{X-} \end{pmatrix}, \quad (4.6.4)$$

where the rank-1 projectors $P_{Z+} = |0\rangle\langle 0|$, $P_{Z-} = |1\rangle\langle 1|$, $P_{X+} = |+\rangle\langle +|$ and $P_{X-} = |-\rangle\langle -|$ diagonalize the Pauli matrices σ_Z and σ_X with eigenvalue $+1$ and -1 respectively, and $0 \leq \alpha, \beta \leq 1$. Given $\mathbf{Q} = \mathbf{S} *^\dagger \mathbf{P}$, the sum of the components of $\mathbf{Q}$ reads in general

$$\frac{\alpha + \beta}{2} \mathbf{1} + (1-\alpha)P_{X+} + (1-\beta)P_{X-} \neq \mathbf{1} \quad (4.6.5)$$

for $0 < |\alpha - \beta| < 1$.

To show item (ii), similarly to the classical case one can choose $\mathbf{S}$ to be

$$\mathbf{S} = (\mathbf{Q} \quad \dots \quad \mathbf{Q}) = \begin{pmatrix} Q_1 & \dots & Q_1 \\ \vdots & \ddots & \vdots \\ Q_n & \dots & Q_n \end{pmatrix}. \quad (4.6.6)$$

Then one has

$$\mathbf{S} *^\dagger \mathbf{P} = \begin{pmatrix} \sum_j \sqrt{Q_1} P_j \sqrt{Q_1} \\ \vdots \\ \sum_j \sqrt{Q_n} P_j \sqrt{Q_n} \end{pmatrix} = \mathbf{Q}, \quad (4.6.7)$$

as the sum over j factorizes.

□

We have seen that the product $*$ preserves the set $\Delta_{n,d}$, whereas the product $*^\dagger$ does not. Moreover, the product $*$ can be interpreted causally in the sense that the measurement which is performed first determines the second measurement, whereas the analogous interpretation of the product $*^\dagger$ is not causal: even though the outcome j of a measurement $\mathbf{P}$ determines a second measurement $\mathbf{S}_j$, the measurement $\mathbf{S}_j$ is performed before $\mathbf{P}$ in the sense that all terms inside $\mathbf{S} *^\dagger \mathbf{P}$ are of the form $\sqrt{S_{ij}} P_j \sqrt{S_{ij}}$. Although from a mathematical perspective one can choose both the product $*$ of Eq. (4.2.3) and its dual $*^\dagger$ in (4.6.3), the first option has a direct physical interpretation as it describes discrete dynamics in the set $\Delta_{n,d}$ induced by concatenation of quantum measurements. This is the reason why we consider this option in the main text of this work.

Algebraic properties of $*$ and $*^\dagger$

The blockwise product $*$ defined in Eq. (4.2.3) and its dual $*^\dagger$ defined in Eq. (4.6.3) have the following properties, where $\neq$ denotes that equality does not hold in general.

1. Non-commutativity,

$$\begin{aligned} \mathbf{A} *^\dagger \mathbf{B} &\neq \mathbf{B} *^\dagger \mathbf{A}; \\ \mathbf{A} * \mathbf{B} &\neq \mathbf{B} * \mathbf{A} \end{aligned} \quad (4.6.8)$$

2. Partial distributivity, in the sense that

$$\begin{aligned} \mathbf{A} *^\dagger (\mathbf{B} + \mathbf{C}) &= \mathbf{A} *^\dagger \mathbf{B} + \mathbf{A} *^\dagger \mathbf{C} \quad \text{but} \\ (\mathbf{A} + \mathbf{B}) *^\dagger \mathbf{C} &\neq \mathbf{A} *^\dagger \mathbf{C} + \mathbf{B} *^\dagger \mathbf{C}; \\ (\mathbf{A} + \mathbf{B}) * \mathbf{C} &= \mathbf{A} * \mathbf{C} + \mathbf{B} * \mathbf{C} \quad \text{but} \\ \mathbf{A} * (\mathbf{B} + \mathbf{C}) &\neq \mathbf{A} * \mathbf{B} + \mathbf{A} * \mathbf{C}. \end{aligned} \quad (4.6.9)$$

3. Non-associativity,

$$\begin{aligned} (\mathbf{A} *^\dagger \mathbf{B}) *^\dagger \mathbf{C} &\neq \mathbf{A} *^\dagger (\mathbf{B} *^\dagger \mathbf{C}); \\ (\mathbf{A} * \mathbf{B}) * \mathbf{C} &\neq \mathbf{A} * (\mathbf{B} * \mathbf{C}). \end{aligned} \quad (4.6.10)$$

4. Unlike the standard product of matrices behaves under the Hermitian conjugate operation $\dagger$ as $(\mathbf{A} \cdot \mathbf{B})^\dagger = \mathbf{B}^\dagger \cdot \mathbf{A}^\dagger$, now we have that

$$\begin{aligned} (\mathbf{A} *^\dagger \mathbf{B})^\dagger &\neq \mathbf{B}^\dagger *^\dagger \mathbf{A}^\dagger; \\ (\mathbf{A} * \mathbf{B})^\dagger &\neq \mathbf{B}^\dagger * \mathbf{A}^\dagger. \end{aligned} \quad (4.6.11)$$

4.6.2 Fixed points under blockwise bistochastic dynamics

Given a blockwise bistochastic matrix $\mathbf{B} \in \mathcal{B}_{n,d}$, here we study which quantum measurements in $\Delta_{n,d}$ remain invariant under the dynamics induced by $\mathbf{B}$. We define these measurements as follows.

Definition 8 (Fixed points). *Given $\mathbf{B} \in \mathcal{B}_{n,d}$ and $\mathbf{P} \in \Delta_{n,d}$, $\mathbf{P}$ is a fixed point of $\mathbf{B}$ with respect to the operation $*$ ($*^\dagger$) if*

$$\begin{aligned} \mathbf{B} * \mathbf{P} &= \mathbf{P} \\ (\mathbf{B} *^\dagger \mathbf{P} &= \mathbf{P}). \end{aligned} \quad (4.6.12)$$

If no specification about $*$ or $*^\dagger$ is given, this denotes that $\mathbf{P}$ is a fixed point of $\mathcal{B}$ with respect to both products.

Let us focus on the simplest case, which is $n = 2$ and arbitrary d . It will be convenient to consider the first entries B and P of a blockwise bistochastic matrix and a blockwise probability

vector of the form

$$\mathcal{B} = \begin{pmatrix} B & \mathbb{1} - B \\ \mathbb{1} - B & B \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} P \\ \mathbb{1} - P \end{pmatrix}. \quad (4.6.13)$$

Observation 4. Let $\mathcal{B} \in \mathcal{B}_{2,d}$ be a blockwise bistochastic matrix defined by an entry B and let $\mathbf{P} \in \Delta_{2,d}$ be a blockwise probability vector defined by an entry P . Let B and P have the following form,

$$B = \mathbb{1} + X \quad \text{and} \quad P = \frac{\mathbb{1} + Y}{2}, \quad (4.6.14)$$

where X and Y have support in mutually orthogonal subspaces, $X \cdot Y = 0$. Then P is a fixed point of B with respect to both $*$ and $*^\dagger$.

Proof. We can write equation (4.6.12) in terms of commutators $[R, S] = R \cdot S - S \cdot R$ as

$$\begin{aligned} (\mathbb{1} - 2P)(\mathbb{1} - B) &= \sqrt{\mathbb{1} - B} [P, \sqrt{\mathbb{1} - B}] - \sqrt{B} [P, \sqrt{B}] \quad \text{and} \\ (\mathbb{1} - 2P)(\mathbb{1} - B) &= \sqrt{\mathbb{1} - P} [B, \sqrt{\mathbb{1} - P}] - \sqrt{P} [B, \sqrt{P}] \end{aligned} \quad (4.6.15)$$

for the products $*^\dagger$ and $*$ respectively. Note that in both cases, if $[B, P] = 0$ the right hand side vanishes and one has

$$\mathbb{1} - 2P - B + 2P \cdot B = 0. \quad (4.6.16)$$

In particular, by inspection we see that if B and P take the form of equation (4.6.14), then they commute and moreover equation (4.6.16) is fulfilled. $\square$

For example, from Observation 4 we know that if

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 - a \end{pmatrix} \quad \text{and} \quad P = \frac{1}{2} \begin{pmatrix} 1 + b & 0 \\ 0 & 1 \end{pmatrix}, \quad (4.6.17)$$

then $\mathbf{P} = (P, \mathbb{1} - P)^\dagger$ is a fixed point. Note that this holds up to global changes of basis, namely for $\{UBU^\dagger, UPU^\dagger\}$ where U is unitary. For example, $P = \mathbb{1}/2 + b|+\rangle\langle+|$ defines a fixed point with respect to $A = \mathbb{1} - a|-\rangle\langle-|$, where we denote $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$.

We have found fixed points for the commutative case. To tackle the noncommutative case, we will restrict to $n = d = 2$ and define a linear map Ξ denoting the following matrix convex combination [187, 221],

$$\Xi(B, P) := \sqrt{BP}\sqrt{B} + \sqrt{\mathbb{1} - B}(\mathbb{1} - P)\sqrt{\mathbb{1} - B}, \quad (4.6.18)$$

which is obtained in the first entry of $\mathbf{Q} = \mathcal{B} *^\dagger \mathbf{P}$ where $\mathcal{B}$ and $\mathbf{P}$ are given in (4.6.13). We will say that P is a fixed point of B if $\mathbf{P}$ is a fixed point of $\mathcal{B}$, which is equivalent to $\Xi(B, P) = P$.

Observation 5. Let $\mathbf{P} \in \Delta_{2,2}$ and $\mathcal{B} \in \mathcal{B}_{2,2}$. If P is a fixed point of B , then P commutes with B .

Proof. Without loss of generality, let us work on the eigenbasis of B . Defining

$$B = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \quad (4.6.19)$$

algebraic manipulation shows that the equation $\Xi(B, P) = P$ can be restated as

$$\begin{pmatrix} 2p_{11}(a-1) & p_{12}(\sqrt{1-a}\sqrt{1-b} + \sqrt{ab}-1) \\ p_{21}(\sqrt{1-a}\sqrt{1-b} + \sqrt{ab}-1) & 2p_{22}(b-1) \end{pmatrix} = \begin{pmatrix} a-1 & 0 \\ 0 & b-1 \end{pmatrix}. \quad (4.6.20)$$

Now suppose that P is nondiagonal. Then, we have

$$\sqrt{1-a}\sqrt{1-b} + \sqrt{ab} - 1 = 0 \implies 1 + ab - a - b = \pm(1 + ab - 2\sqrt{ab}), \quad (4.6.21)$$

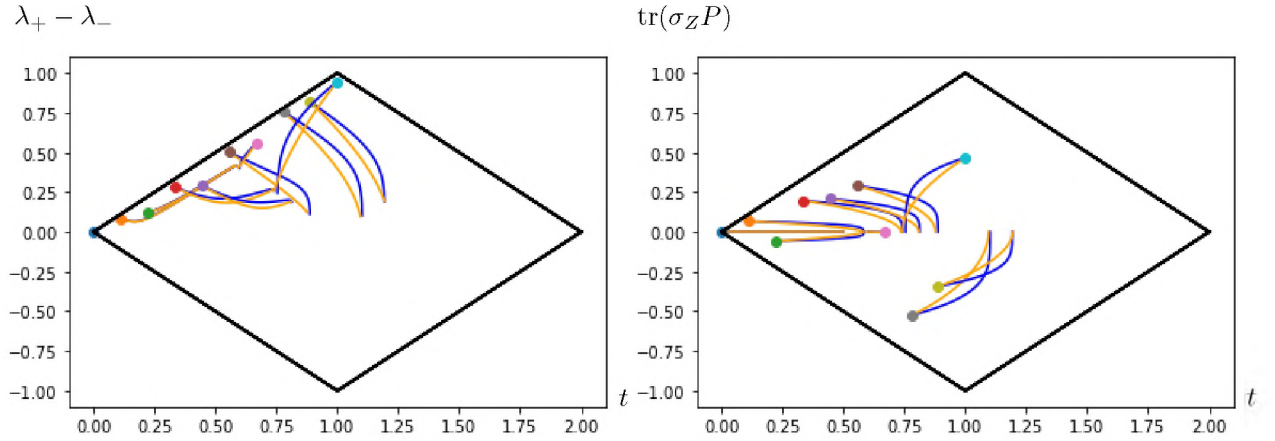


Figure 4.6.1: Blockwise bistochastic dynamics of quantum measurements in $\Delta_{2,2}$, under small noise. Evolution of quantum measurements starting at ten random initial points P (coloured dots) with different traces $0 \leq \text{tr}(P) \leq 1$, induced by a blockwise bistochastic matrix with diagonal entry ϵ -close to identity, $B = \mathbb{1} - \epsilon|-\rangle\langle-|$, choosing $\epsilon = 0.01$. Trajectories induced by the operation $*$ ($*^\dagger$) are depicted in blue (orange). We show along the coordinate t , the difference of eigenvalues $\lambda_+(P) - \lambda_-(P)$ (left) and the projection $\text{tr}(\sigma_Z P)$ onto the Pauli operator σ_Z (right).

which has only the trivial solution $B = \mathbb{1}$ for the $-$ sign and $B \propto \mathbb{1}$ for the $+$ sign. Similar results can be shown for the product $*$, by working in the eigenbasis of P . Thus, we conclude that B and P must commute. $\square$

To finish this subsection we will study slow dynamics where B is the identity with a real noise ϵ , namely

$$B = \begin{pmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{pmatrix} = \mathbb{1} - 2\epsilon|-\rangle\langle-| \quad (4.6.22)$$

with $0 \leq \epsilon \ll 1$. Now we know that in this case, any P of the form $P = \mathbb{1}/2 + \lambda|+\rangle\langle+|$, with $-1/2 \leq \lambda \leq 1/2$ to ensure positivity of both P and $\mathbb{1} - P$, defines a fixed point. For instances of the form of Eq. (4.6.22), numerical evidence suggests that all fixed points are of this form. The random matrices which generate the sample are defined in Eq. (4.6.25). This is demonstrated in Figure 4.6.1.

Observation 6. *For a large sample (of the order of 10^6) of Ginibre-random matrices P (see 4.6.3) and matrices B of the form (4.6.22), in all instances the dynamics converge to fixed points of the form of Eq. (4.6.14) for both products $*$ and $*^\dagger$.*

4.6.3 Nonlinear monotone

Theorems 5 and 6 give a family of linear monotones for the blockwise bistochastic dynamics of sortable quantum measurements. Here we address the majorization relations beyond the linear case and ask if a nonlinear monotone can be found outside of the sortable set.

A conjecture for a nonlinear monotone

To address this question, note that p majorizes q if and only if $(p_1 - \frac{1}{n}, \dots, p_n - \frac{1}{n})$ majorizes $q = (q_1 - \frac{1}{n}, \dots, q_n - \frac{1}{n})$, since the non-increasing order makes all cumulative sums be nonnegative. On the other hand, note that if p majorizes q , then the cumulative sums of p in nonincreasing order are larger or equal than the cumulative sums of q for any possible order of q . We are now ready for the following conjecture.

Conjecture 1. Let $\mathbf{P} \in \Delta_{n,d}$ and $\mathbf{Q} \in \Delta_{n,d}$ be blockwise probability vectors. If there exists a blockwise bistochastic matrix $\mathbf{B} \in \mathcal{B}_{n,d}$ such that $\mathbf{Q} = \mathbf{B} * \mathbf{P}$, then for any ordering of $\{Q_i\}$ there exists an ordering of $\{P_i\}$ such that

$$\left\| \sum_{i=1}^k \left(P_i - \frac{1}{n} \mathbf{1} \right) \right\|_2 \geq \left\| \sum_{i=1}^k \left(Q_i - \frac{1}{n} \mathbf{1} \right) \right\|_2 \quad (4.6.23)$$

for all $1 \leq k \leq n$, where $\|A\|_2 = \sqrt{\text{tr}[A^\dagger A]}$ is the 2-norm.

Physically, this conjectured monotone for $k = 1$ can be understood in terms of the generalized Bloch vector τ introduced in subsection 1.2.3, with origin at $\mathbf{1}/n$ and pointing to P_1 . One has

$$\left\| P_1 - \frac{1}{n} \mathbf{1} \right\|_2 = \sqrt{d \text{tr} \left[\left(P_1 - \frac{1}{n} \mathbf{1} \right)^2 \right]} = \sqrt{\left(t - \frac{d}{n} \right)^2 + \tau^2}, \quad (4.6.24)$$

which gives the Hilbert-Schmidt distance between P_1 and $\mathbf{1}/n$, and thus it contains information about the position of $\mathbf{P}$ with respect to the trivial POVM described by the uniform vector $\mathbf{P}_u = (\frac{1}{n}, \dots, \frac{1}{n})^\dagger$.

Conjecture 1 is supported by a sample of random blockwise probability vectors and blockwise bistochastic matrices of different sizes. This sample consists of combining different ways of choosing random blockwise probability vectors and random blockwise bistochastic matrices. Random blockwise probability vectors from $\Delta_{n,d}$ have been chosen in the following way.

1. Choose a square random matrix X of order d from the *Ginibre ensemble* [222, 223], with entries of the form $a + ib$, where i is the complex unit and a (b) is chosen randomly from a Gaussian distribution of real numbers with mean 0 and variance 1. Assign to the first component the operator

$$P_1 = t_1 \frac{X^\dagger X}{\text{tr}[X^\dagger X]}, \quad (4.6.25)$$

where t_1 is a random number drawn uniformly from $[0, 1]$. In this way a random positive operator P_1 is generated according to the Hilbert-Schmidt measure in the set of quantum states Ω_d rescaled by the factor t_1 [223].

The rest of the components of $\mathbf{P}$ are computed using a compatibility optimizer in a semidefinite program, with the constraints that all effects P_i are positive semidefinite and sum to $\mathbf{1}$.

2. Take $\mathbf{P}$ chosen as previously and multiply every component by $\epsilon \ll 1$. Add to the first component $(1 - \epsilon)\mathbf{1}$. This creates a blockwise probability vector which is ϵ -close to $(\mathbf{1}, \dots, 0)^T$.
3. Take $\mathbf{P}$ chosen as in the first case and multiply every component by $\epsilon \ll 1$. Add to each component $\frac{1 - \epsilon}{n}\mathbf{1}$. This creates a blockwise probability vector which is ϵ -close to $(\mathbf{1}, \dots, \mathbf{1})^T/n$.

In a similar way, we sample blockwise bistochastic matrices constructed as follows.

1. Choose B_{11} according to Eq. (4.6.25) with X a random Ginibre matrix and t_1 a random number drawn uniformly from $[0, 1]$. The rest of the components B_{ij} of $\mathbf{B}$ are computed using a compatibility optimizer in a semidefinite program, with the constraints $B_{ij} \geq 0$ and $\sum_i B_{ij} = \sum_j B_{ij} = \mathbf{1}$ according to Definition 6.
2. Take $\mathbf{B}$ chosen as previously and multiply every component by $\epsilon \ll 1$. Add to each diagonal component $(1 - \epsilon)\mathbf{1}$. This creates a blockwise probability vector which is ϵ -close to $\mathbf{1}_{nd}$, which leaves invariant any blockwise probability vector.

3. Take $\mathbf{B}$ chosen as in the first case and multiply every component by $\epsilon \ll 1$. Add to each component $\frac{1-\epsilon}{n}\mathbf{1}$. This creates a blockwise probability vector which is ϵ -close to $B_{ij} = \mathbf{1}/n$, which brings any blockwise probability vector to $(\mathbf{1}, \dots, \mathbf{1})/n$.
4. Choose the first row of $\mathbf{B}$ to be a blockwise probability vector constructed as in step (i) above. Choose the rest of the entries of $\mathbf{B}$ as $B_{i,j} = B_{1,j+i}$ where the sum is taken modulo n . Let us mention that this is a generalization of so-called *bistochastic circulant matrices*, which are defined in this way for $d = 1$.

Analytical proof for the two-effect case in arbitrary dimensions

We shall see that Conjecture 1 holds for the case $n = 2$. Before that, let us comment on three properties of this special case. First, any blockwise probability vector $\mathbf{P} = (P, \mathbf{1} - P)$ is fully determined by its first P . Second, one has

$$\|\Xi(B, P)\|_2 = \|\Xi(B, \mathbf{1} - P)\|_2, \quad (4.6.26)$$

from which we recover the fact that in Δ_2 both components of any vector are equally distant from $1/2$. Third, let us recall that for $n = 2$ both products from the right and from the left preserve the set of measurements $\Delta_{2,d}$. Accordingly, in the following Lemma we shall see that Conjecture 1 holds for both products in the case $n = 2$.

Lemma 5. *After the blockwise products $*$ and $*^\dagger$ by a blockwise bistochastic matrix $\mathcal{B} \in \mathcal{B}_{2,d}$ on a blockwise probability vector $\mathbf{P} = (P, \mathbf{1} - P)^\dagger$, the vector position of P in the dual cone $\|P\|_2$ shrinks,*

$$\|P\|_2 \geq \|\Xi(B, P)\|_2. \quad (4.6.27)$$

Proof. We will first prove the Lemma for the product $*^\dagger$, using the map $\Xi(B, P)$. By the cyclic property of the trace, the result extends to the product $*$. Starting from Eq. (4.6.27), by using Eq. (4.6.24) and squaring both sides of the equality, we arrive at the equivalent condition

$$\text{tr}(\Xi(B, P)^2) - \text{tr}(\Xi(B, P)) \leq \text{tr}(P^2) - \text{tr}(P). \quad (4.6.28)$$

By developing this expression and canceling terms out, we arrive at

$$2 \text{tr}[PBPB] - 4 \text{tr}[PB^2] - 2 \text{tr}[BP^2] - 4 \text{tr}[PB] + \text{tr}[B^2] - \text{tr}[B] - 2 \text{tr}[(P\sqrt{B - B^2})^2] \leq 0. \quad (4.6.29)$$

By positivity of $\mathbf{1} - B$, the eigenvalues of B must be smaller or equal than 1, and therefore we have

$$\text{tr}[B^2] - \text{tr}[B] \leq 0. \quad (4.6.30)$$

It will be convenient to invert the sign of the last term and express what is left to prove as

$$0 \geq 2 \text{tr}[PBPB] - 4 \text{tr}[PB^2] - 2 \text{tr}[BP^2] - 4 \text{tr}[PB] + 2 \text{tr}[(P\sqrt{B^2 - B})^2] := F, \quad (4.6.31)$$

where we have defined the expression above as F to shorten the notation below. Using that

$$\text{tr}[(XY)^2] \leq \text{tr}[X^2Y^2] \quad (4.6.32)$$

for X, Y Hermitian matrices [224], we have that

$$\begin{aligned} F &\leq 2 \text{tr}[PBPB] - 4 \text{tr}[PB^2] - 4 \text{tr}[BP^2] - 4 \text{tr}[PB] + 2 \text{tr}[P^2B^2] \\ &\leq 4(\text{tr}[P^2B^2] - \text{tr}[PB] - \text{tr}[PB^2] + \text{tr}[P^2B]) = \text{tr}[BP(PB - \mathbf{1} - B + P)] := F'. \end{aligned} \quad (4.6.33)$$

Using that

$$\text{tr}[XY] \leq \text{tr}[X] \text{tr}[Y] \quad (4.6.34)$$

for X, Y positive semidefinite matrices and reordering terms, we see that

$$F' \leq \operatorname{tr}[BP](\operatorname{tr}[PB - B] + \operatorname{tr}[P - \mathbf{1}]) \leq -\operatorname{tr}[BP] \operatorname{tr}[B + \mathbf{1}] \operatorname{tr}[\mathbf{1} - P] \leq 0 \quad (4.6.35)$$

as P , B and $\mathbf{1} - P$ are positive semidefinite by assumption. $\square$

Note that for $\Delta_{2,1}$, namely if P and B are real numbers p and b contained between 0 and 1, then the Lemma above holds trivially, as the product of any two real numbers between 0 and 1 is always smaller or equal than each of them. This implies that $|1/2 - p|$ is a monotone under bistochastic dynamics, which is necessary and sufficient for the standard majorization condition $p \succ Bp$.

Certifying nonlocal properties of noisy quantum operations

This chapter provides a unified framework to certify nonlocal properties of quantum channels from the correlations obtained in measurement protocols. The approach introduced gathers and extends fully- and semi-device independent methods for this purpose. We study the effect of different models of dephasing noise, some of which are shown to generate nonlocality and entanglement in special cases. In the extreme case of complete dephasing, the measurement protocols discussed yield particularly simple tests to certify nonlocality. These are based on the relations between bipartite quantum channels and their classical analogs: bipartite stochastic matrices defining conditional distributions.

5.1 Introduction

Quantum theory predicts phenomena that cannot be explained with local hidden variable models [51, 53]. This property, known as nonlocality, is a resource for certain protocols in quantum information and computation [225–227]. These are implemented by using quantum states describing physical systems, which are processed and manipulated through quantum operations. Information is obtained by performing measurements, which provide a probability distribution in terms of relative frequencies of outcomes.

Therefore, it is of primal importance to identify nonlocal properties of both quantum states [54, 228] and the quantum operations acting on them [229–231], from the nonlocal behaviours in the outcome probability distributions [232–234]. Bell games are linear tests for measurement protocols that identify nonlocality in quantum states, by analyzing the correlations obtained through classical inputs and outputs [51–53]. By using quantum inputs, these can be modified to detect nonlocality of all entangled states [56]. Although these methods have been employed to study nonlocal properties of quantum states since several decades ago [51–53, 56], it was more recently that similar schemes were proposed to identify nonlocality in quantum channels [7, 229–231, 235, 236].

However, quantum operations in realistic setups suffer from the phenomenon of *decoherence*: due to imperfect isolation, internal degrees of freedom of the system are coupled with the environment [21, 237]. This typically leads to a full or partial loss of quantum properties [238], and therefore identifying nonlocality in decoherent setups becomes a challenge [69]. In particular, quantum systems are affected by *dephasing* noise [58, 239], which acts on a quantum state by damping the off-diagonal terms in the density matrix while leaving intact the diagonal terms in a fixed basis [240–242].

In this paper we provide a framework to certify nonlocal properties of quantum channels, in the presence of dephasing noise. To this end, we use measurement protocols depicted in Fig 5.3.1 to formulate (semi-)device independent criteria to certify nonlocality of quantum channels. This approach includes and extends existing methods [7, 229] and allows us to identify the role of dephas-

ing noise in the nonlocal properties of quantum channels: while local dephasing cannot generate nonlocality, there are cases in which a global dephasing operation can generate both entanglement and nonlocality. Our analysis in the case of complete decoherence yields simple criteria to detect nonlocality in quantum channels through a geometric insight into the transition from quantum to classical operations.

5.2 Preliminaries

5.2.1 Quantum and classical states and channels

The set of quantum states ϱ acting on a Hilbert space of dimension d is represented by $\mathcal{M}_d$. The set of classical states, namely the d -dimensional probability simplex, is denoted by Δ_d . A classical channel transforming probability distributions is a *stochastic matrix*, that is, a matrix with non-negative entries that sum up to one in each column. The set of stochastic matrices with the dimension of inputs and outputs equal to d is represented by $\mathcal{S}_d$. Analogously, completely positive (CP) and trace preserving (TP) maps, $\mathcal{E}_{A_0 \rightarrow A_1}$, also known as quantum operations or *quantum channels*, transform quantum states. Their inputs and outputs are density matrices acting on a Hilbert space $\mathcal{H}_{A_0}$ and $\mathcal{H}_{A_1}$ with dimensions d_{A_0} and d_{A_1} , respectively. When the input and output systems are held by the same party, we will denote by A the joint system $A_0 A_1$ and use the notation $\mathcal{E}_A$, or simply $\mathcal{E}$, if no confusion arises.

We denote by $J_{A_0 A_1}^{\mathcal{E}}$ (or simply $J^{\mathcal{E}}$) the $d_{A_0} d_{A_1} \times d_{A_0} d_{A_1}$ *Choi matrix* acting on $\mathcal{H}_{A_0} \otimes \mathcal{H}_{A_1}$, which defines a quantum channel $\mathcal{E}_A$ through the Choi-Jamiołkowski isomorphism [26, 27],

$$\mathcal{E}_A(\varrho_{A_0}) = \text{tr}_{A_0}(J_{A_0 A_1}(\varrho_{A_0}^T \otimes \text{id}_{A_1})). \quad (5.2.1)$$

The complete positivity and trace preserving (CPTP) conditions require that $J_{A_0 A_1} \geq 0$ and

$$J_{A_0} := \text{tr}_{A_1}(J_{A_0 A_1}) = \text{id}_{A_0}, \quad (5.2.2)$$

where we denote the partial trace by omitting the traced subsystems.

Following [28, 64], we denote a *quantum superchannel* as a bipartite channel Ξ_{AB} with the special property that it transforms Choi matrices to Choi matrices. That is, a quantum channel $\mathcal{E}_{A_1 \rightarrow B_0}$ is transformed according to

$$\Xi_{A_0 B_0 \rightarrow A_1 B_1}(\mathcal{E}_{A_1 \rightarrow B_0}) = \mathcal{E}'_{A_0 \rightarrow B_1}. \quad (5.2.3)$$

For the above to hold, the Choi matrix of Ξ should satisfy [28]

$$J^{\Xi} \geq 0 \quad (5.2.4a)$$

$$J_{A_0 B_0}^{\Xi} = \text{id}_{A_0 B_0} \quad (5.2.4b)$$

$$J_{A B_0}^{\Xi} = J_{A_0 A_1}^{\Xi} \otimes \text{id}_{B_0} / d_{B_0}. \quad (5.2.4c)$$

Similarly, a classical superchannel Γ [243], transforming stochastic matrices of size $d_{B_0} \times d_{A_1}$ to those of size $d_{B_1} \times d_{A_0}$, is isomorphic to a bipartite stochastic matrix of dimension $d_{A_1} d_{B_1} \times d_{A_0} d_{B_0}$. To be a valid superchannel, it has to additionally satisfy

$$\sum_{b_1} \Gamma_{a_1 b_1, a_0 b_0} = \sum_{b_1} \Gamma_{a_1 b_1, a_0 b'_0}. \quad (5.2.5)$$

The set of quantum and classical superchannels acting on d -dimensional channels are respectively denoted by Θ_d and $\mathcal{T}_d$.

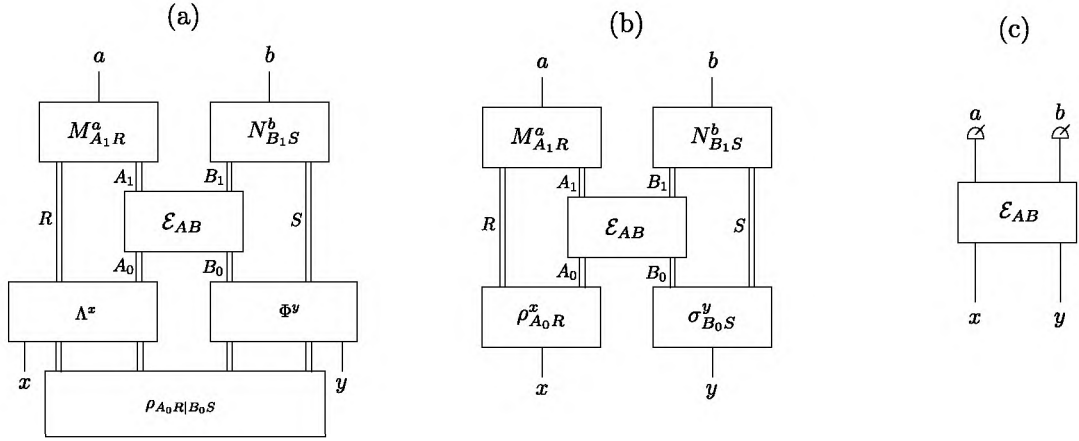


Figure 5.3.1: Schemes to obtain bipartite conditional probability distributions (bipartite stochastic matrices) from bipartite channels: (a) A shared entangled state is processed with local operations (with ancillary systems) to produce the inputs which undergo the channel $\mathcal{E}$ to be analyzed, and then local measurements are performed to obtain the probabilities of Eq. (5.3.1). (b) Product states prepared locally undergo a bipartite channel and then local generalized measurements are performed to obtain the distribution in Eq. (5.3.7). This case can be seen as bipartite distributions being obtained after local superchannels are performed in the form of $\Xi_A \otimes \Xi_B[\mathcal{E}_{AB}]$. (c) A bipartite distribution in Eq. (5.4.8) is obtained with inputs in product states spanning a preferred basis, which undergo a bipartite channel and are then projective-measured in the same preferred basis.

5.3 Channel nonlocality from local measurements

5.3.1 Local measurement protocols with shared entanglement

Bipartite distributions can be obtained from bipartite quantum channels by choosing input states and measuring their conditional outputs [229, 231, 235, 244], as shown in Fig. 5.3.1. In this spirit, consider the probability distribution obtained by the measurement protocol of Fig. 5.3.1 (a),

$$p(a, b|x, y) = \text{tr} \left(\left(M^a_{RA_1} \otimes N^b_{SB_1} \right) \cdot \mathcal{E}_{AB} \otimes \text{id}_{RS} [\Lambda^x_{A_0R} \otimes \Phi^y_{B_0S} (\varrho_{ABRS})] \right). \quad (5.3.1)$$

Proposition 9. *Let $p(a, b|x, y)$ be a probability distribution obtained from a bipartite quantum channel $\mathcal{E}_{AB}$ through Eq. (5.3.1). For any input channels Λ^x and Φ^y and for any measurements M^a and N^b , the following holds:*

- (i) *If $\mathcal{E}_{AB}$ is in LOSR, then for any input state $\varrho_{AR|BS}$ which is local in the bipartition $AR|BS$, $p(a, b|x, y)$ is in L .*
- (ii) *If $\mathcal{E}_{AB}$ is in LOSE, then for any shared state ϱ , $p(a, b|x, y)$ is in Q .*
- (iii) *If $\mathcal{E}_{AB}$ is in qNS, then for any shared state ϱ , $p(a, b|x, y)$ is in NS .*

Proof. Proof of (i). For any channel $\mathcal{E}$, we denote its adjoint by $\mathcal{E}^\dagger$, which satisfies $\text{tr}(\mathcal{E}[A_0]A_1) = \text{tr}(\mathcal{E}^\dagger[A_1]A_0)$ for any A_0 and A_1 . It is enough to consider an extremal LOSR operation $\mathcal{E}_{AB} = \mathcal{E}_A \otimes \mathcal{E}_B$, since the rest of the proof follows by convex combinations. Define the operators

$$\tilde{M}^{a|x} := \Lambda^x_{AR}{}^\dagger \circ (\mathcal{E}_A^\dagger \otimes \text{id}_R)[M^a_{RA_1}], \quad (5.3.2a)$$

$$\tilde{N}^{b|y} := \Phi^y_{BS}{}^\dagger \circ (\mathcal{E}_B^\dagger \otimes \text{id}_S)[N^b_{SB_1}]. \quad (5.3.2b)$$

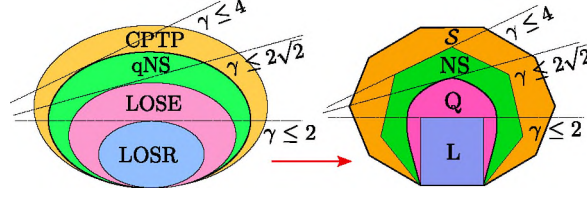


Figure 5.3.2: Subsets of distributions obtained from bipartite operations. *Left.* Inclusions of the following subsets of quantum channels (CPTP): quantum nonsignaling (qNS), local operations assisted by the shared resources of entanglement (LOSE) and randomness (LOSR). *Right.* Inclusions of the subsets of the polytope of bipartite stochastic matrices ($\mathcal{S}$): nonsignaling (NS), quantum (Q) and local (L) correlations. Bell functionals γ bounding each subset of bipartite distributions bound the corresponding subsets of bipartite channels, through Prop. 9, Cor. 4 and Prop. 12. Here we depict the linear cuts delimited by CHSH functional.

Since the adjoint $\mathcal{E}^\dagger$ of a CPTP map $\mathcal{E}$ is completely positive and unital, i.e., it satisfies $\mathcal{E}^\dagger(\text{id}) = \text{id}$ [12], the operators $\tilde{M}^{a|x}$ ($\tilde{N}^{b|y}$) are positive semidefinite and their summation over a (b) is the identity for any x (y). Thus, we can rewrite Eq. (5.3.1) as

$$p(a, b|x, y) = \text{tr} \left(\varrho_{ABRS} \cdot \tilde{M}_{RA_1}^{a|x} \otimes \tilde{N}_{SB_1}^{b|y} \right), \quad (5.3.3)$$

which is a standard Bell game. Now it becomes evident by definition, that if the state ϱ_{ABRS} is local in the bipartition $AR|BS$, then a local distribution is obtained.

Proof of (ii). Consider an LOSE channel, which factorizes as in Eq. (1.2.33). Similarly as in the proof of item (i), define the operators

$$\tilde{M}^{a|x} := (\Lambda_{AR}^{x\dagger} \otimes \text{id}_E) \circ \quad (5.3.4a)$$

$$(\mathcal{E}_{AE}^\dagger \otimes \text{id}_R)[M_{RA_1}^a \otimes \text{id}_{E_1}],$$

$$\tilde{N}^{b|y} := (\Phi_{BS}^{y\dagger} \otimes \text{id}_F) \circ \quad (5.3.4b)$$

$$(\mathcal{E}_{BF}^\dagger \otimes \text{id}_S)[N_{SB_1}^b \otimes \text{id}_{F_1}].$$

Using Eqs. (5.3.4), one verifies that Eq. (5.3.1) reads

$$p(a, b|x, y) = \text{tr} \left(\tilde{M}_{ARE}^{a|x} \otimes \tilde{N}_{BSF}^{b|y} \cdot \varrho_{ARBS} \otimes \tau_{EF} \right). \quad (5.3.5)$$

This proves that by applying an LOSE channel one can only recover quantum correlations.

Proof of (iii). By imposing in Eq. (5.3.1) that $\mathcal{E}$ satisfies $J_{A_0 B_0 B_1}^\mathcal{E} = \text{id}_{A_0} \otimes J_{B_0 B_1}/d_{A_0}$, we obtain

$$\sum_a p(a, b|x, y) = \sum_a \text{tr} \left((\text{id}_{B_0} \otimes N_{SB_1}^b) \cdot (J_{B_0 B_1} \otimes \text{id}_S)(\varrho_{B_0 S}^y \otimes \text{id}_{B_1}) \right) \quad (5.3.6)$$

where $\varrho_{B_0 S}^y := \Phi^y(\text{tr}_{A_0 R}(\varrho_{ARBS}))$, which does not depend on x . Similarly, by imposing $J_{A_0 A_1 B_0}^\mathcal{E} = J_{A_0 A_1} \otimes \text{id}_{B_0}/d_{B_0}$ we obtain that $\sum_b p(a, b|x, y)$ does not depend on y and therefore $p(a, b|x, y)$ is a nonsignaling distribution. $\square$

Note that Proposition 9 provides a criterion to detect quantum channels outside of the sets LOSR, LOSE and qNS which is fully device independent, in the sense that none of the quantum devices (the shared state ϱ , the input channels Λ^x and Φ^y , and the measurements M^a and N^b) is trusted.

5.3.2 Local measurement protocols without entanglement

Consider the protocol in Fig. 5.3.1 (b),

$$p(a, b|x, y) = \text{tr} \left((M_{RA_1}^a \otimes N_{SB_1}^b) \cdot \mathcal{E}_{AB} \otimes \text{id}_{RS}(\varrho_{A_0R}^x \otimes \sigma_{B_0S}^y) \right), \quad (5.3.7)$$

which was introduced in [229]. This can be obtained from Eq. (5.3.1) by considering a product state $\varrho_{ARBS} = |00\rangle\langle 00|_{AR} \otimes |00\rangle\langle 00|_{BS}$, so that the channels Λ_{AR}^x and Φ_{BS}^y act as state preparation channels producing input states ϱ_{AR}^x and σ_{BS}^y . LOSR channels produce local probability distributions from Eq. (5.3.7), as shown in [229]. This result can be extended for other classes of channels.

Corollary 4. *Let $p(a, b|x, y)$ be a probability distribution obtained from a bipartite quantum channel $\mathcal{E}_{AB}$ through Eq. (5.3.7). The following holds, for any choice of input states and measurements:*

- (i) *If $\mathcal{E}_{AB}$ is LOSR, then $p(a, b|x, y)$ is in L .*
- (ii) *If $\mathcal{E}_{AB}$ is LOSE, then $p(a, b|x, y)$ is in Q .*
- (iii) *If $\mathcal{E}_{AB}$ is qNS, then $p(a, b|x, y)$ is in NS .*

Proof. Proof of (i). Probability distributions obtained from Eq. (5.3.7) with local operation factorize, thus LOSR channels give local probability distributions.

Proof of (ii). Similarly as in the proof of Prop. 9, define

$$\tilde{M}^{a|x} = \text{tr}_{RA_0} \left((\varrho_{RA_0}^x \otimes \text{id}_{E_0}) \cdot (\mathcal{E}_{AE}^\dagger \otimes \text{id}_R)[M_{RA_1}^a \otimes \text{id}_{E_1}] \right), \quad (5.3.8a)$$

$$\tilde{N}^{b|y} = \text{tr}_{SB_0} \left((\sigma_{SB_0}^y \otimes \text{id}_{F_0}) \cdot (\mathcal{E}_{BF}^\dagger \otimes \text{id}_S)[N_{SB_1}^b \otimes \text{id}_{F_1}] \right). \quad (5.3.8b)$$

One verifies that $p(a, b|x, y)$ in Eq. (5.3.7) reads

$$p(a, b|x, y) = \text{tr} \left(\tilde{M}_E^{a|x} \otimes \tilde{N}_F^{b|y} \cdot \tau_{EF} \right). \quad (5.3.9)$$

Proof of (iii). This follows from Proposition 9. □

Two remarks are now in order. On the one hand, note that the converse of Corollary 4 is not true. In particular, channels denoted as *local-limited* in [229] always produce local correlations for any choice of input states and measurements. Examples are LOSE channels where Alice and Bob share an entangled state that cannot produce nonlocal correlations [54]. On the other hand, Eq. (5.3.5) implies that a bipartite LOSE channel produces nonlocal correlations if the shared state τ_{EF} in Eq. (1.2.33) is nonlocal in some Bell protocol. This in turn dictates that if the shared state in one realization of an LOSE channel is Bell nonlocal, the channel cannot be realized by some shared local state.

5.3.3 A Buscemi-like characterization of LOSR operations

The results of Buscemi [56] extend to more general processes [7]. In particular, for a fixed channel $\mathcal{E}$ and a fixed Bell functional γ , we denote as $\gamma_{\max}(\mathcal{E})$ the maximal value of $\gamma(p)$ that can be achieved when the conditional probability distribution p is obtained by a fixed set of input states $\{\varrho^x\}$ and $\{\sigma^y\}$ over all possible measurements $\{M^a\}$ and $\{N^b\}$, i.e.,

$$\gamma_{\max}(\mathcal{E}) := \max_{\{M^a, N^b\}} \sum_{xy, ab} \gamma_{ab, xy} p(a, b|x, y), \quad (5.3.10)$$

where $p(ab|xy)$ is obtained through Eq. (5.3.7).

The setup depicted in Fig. 5.3.1 (b) can be seen as a particular case of so-called *semi-quantum games*, which are known to characterize convertibility of general processes through free operations in a generalized theory of nonlocality [7]. In particular, the setup described in Eq. (5.3.7) and Fig. 5.3.1 (b) provides semi-quantum games characterizing local transformations between LOSR quantum channels. The following follows from [7, Corollary 1]:

Proposition 10. *Given two bipartite quantum operations $\mathcal{E}_{AB}$ and $\mathcal{E}'_{A'B'}$, $\mathcal{E}_{AB}$ can be transformed into $\mathcal{E}'_{A'B'}$ by a convex combination of product superchannels if and only if for all Bell functionals γ and for all sets of input states, $\gamma_{\max}(\mathcal{E}) \geq \gamma_{\max}(\mathcal{E}')$, where $\gamma_{\max}$ is given by Eq. (5.3.10).*

Proof. The proof of the **only if** direction can be easily seen as follows. If $\mathcal{E}$ can be converted into $\mathcal{E}'$ by a product of superchannels (or a convex combination of them), then we can absorb such product of superchannels in the state preparations and measurements of semiquantum games, described in Eq. (5.3.7) and depicted in Fig. 5.3.1 (b). Then it becomes clear that $\gamma_{\max}(\mathcal{E}) \geq \gamma_{\max}(\mathcal{E}')$ for all functionals γ .

Now we will prove the **if** part of Proposition 10. Namely, we will show that if $\mathcal{E}$ is more nonlocal than $\mathcal{E}'$ for all the games specified in Fig. 5.3.1 (b), then it can be transformed to $\mathcal{E}'$ through some superchannels $\Xi = \sum_i \nu_i \Xi_i^A \otimes \Xi_i^B$. We will follow a similar reasoning to the analogous result of Buscemi for bipartite quantum states [56]. To this end, consider the extended Hilbert spaces

$$\mathcal{H}_{E_0} = \mathcal{H}_{A_0} \otimes \mathcal{H}_{R_0} = \mathcal{H}_{A'_0} \otimes \mathcal{H}_{R'_0} \quad \text{and} \quad (5.3.11)$$

$$\mathcal{H}_{F_0} = \mathcal{H}_{B_0} \otimes \mathcal{H}_{S_0} = \mathcal{H}_{B'_0} \otimes \mathcal{H}_{S'_0}. \quad (5.3.12)$$

Note that the input and output dimensions of the channels $\mathcal{E}_{AB}$ and $\mathcal{E}'_{A'B'}$ in consideration can in general be different. Therefore, the subsystems E_0 , R_0 and R'_0 (F_0 , S_0 , and S'_0) shall be chosen accordingly such that the dimensions match. In particular, we set the dimension d_{E_0} (d_{F_0}) to be a common multiple of the input dimensions d_{A_0} and $d_{A'_0}$ (d_{B_0} and $d_{B'_0}$). Namely, there exist natural numbers n , n' , m and m' so that

$$d_{E_0} = nd_{A_0} = n'd_{A'_0} \quad \text{and} \quad (5.3.13)$$

$$d_{F_0} = md_{B_0} = m'd_{B'_0}. \quad (5.3.14)$$

Thus, the input states $\varrho^x \in \mathcal{M}_{d_{E_0}}$ and $\sigma^y \in \mathcal{M}_{d_{F_0}}$ are regarded in different bipartitions when they are used as input states for the different channels $\mathcal{E}$ and $\mathcal{E}'$. For example, suppose we consider a channel $\mathcal{E}_{AB}$ with $d_{A_0} = 4$ and a channel $\mathcal{E}'_{A'B'}$ with $d_{A'_0} = 6$. Then we can consider an input state $\varrho^x \in \mathcal{M}_{12}$, which can be seen either as a bipartite state $\varrho^x_{A_0 R_0}$ with $d_{A_0} = 4$ and $d_{R_0} = 3$, or as a bipartite state $\varrho^x_{A'_0 R'_0}$ with $d_{A'_0} = 6$ and $d_{R'_0} = 2$.

In a similar manner, consider the convex combinations of local measurements in the output systems E_1 and F_1 , $Z_{E_1 F_1}^{a,b} = \sum_i p_i M^{a,i}_{E_1} \otimes N^{b,i}_{F_1}$. Following the same arguments in the proof of Proposition 1 of [56], notice that the probability distributions available for a particular game with different measurements form a convex set. Therefore, $\gamma(\mathcal{E}) \geq \gamma(\mathcal{E}')$ for all γ is equivalent to assuming for any measurement Z of the above form there exists a measurement $\bar{Z}$ in the convex hull of local measurements $\bar{M}_{E_1} \otimes \bar{N}_{F_1}$ such that

$$p(ab|xy) = \text{tr} \left[\bar{Z}_{E_1 F_1}^{a,b} \cdot \text{id}_{RS} \otimes \mathcal{E}_{AB}(\varrho^x_{E_0} \otimes \sigma^y_{F_0}) \right] = \text{tr} \left[Z_{E_1 F_1}^{a,b} \cdot \text{id}_{R'S'} \otimes \mathcal{E}'_{A'B'}(\varrho^x_{E_0} \otimes \sigma^y_{F_0}) \right]. \quad (5.3.15)$$

Define now the input states as

$$\varrho^x_{E_0} := \text{tr}_{E_0^c} [\Theta_{E_0^c}^x \otimes \text{id}_{E_0} \cdot \phi_{E_0^c E_0}^+] \quad \text{and} \quad (5.3.16a)$$

$$\sigma^y_{F_0} := \text{tr}_{F_0^c} [\Upsilon_{F_0^c}^y \otimes \text{id}_{F_0} \cdot \phi_{F_0^c F_0}^+], \quad (5.3.16b)$$

with extended Hilbert spaces $\mathcal{H}_{E_0^c} \mathcal{H}_{E_0}$ and $\mathcal{H}_{F_0^c} \mathcal{H}_{F_0}$, where $\{\Theta^x\}_x$ and $\{\Upsilon^y\}_y$ are informationally complete, positive operator-valued measurements (IC-POVMs) and $\phi_{E_0^c E_0}^+$ is the density matrix of the Bell state $|\phi^+\rangle_{E_0^c E_0} = 1/\sqrt{d_{E_0}} \sum_{i=0}^{d_{E_0}-1} |i\rangle_{E_0^c} |i\rangle_{E_0}$ (and similarly for $\phi_{F_0^c F_0}^+$). For convenience we will denote $\mathcal{E}_{EF} := \text{id}_{RS} \otimes \mathcal{E}_{AB}$ and $\mathcal{E}'_{EF} := \text{id}_{R'S'} \otimes \mathcal{E}'_{A'B'}$. Thus, we can write Eq. (5.3.15) as

$$\begin{aligned} p(ab|xy) &= \text{tr} \left[\Theta_{E_0^c}^x \otimes \bar{Z}_{E_1 F_1} \otimes \Upsilon_{F_0^c}^y \cdot \mathcal{E}_{EF} \otimes \text{id}_{E_0^c F_0^c} (\phi_{E_0^c E_0}^+ \otimes \phi_{F_0^c F_0}^+) \right] \\ &= \text{tr} \left[\Theta_{E_0^c}^x \otimes Z_{E_1 F_1} \otimes \Upsilon_{F_0^c}^y \cdot \mathcal{E}'_{EF} \otimes \text{id}_{E_0^c F_0^c} (\phi_{E_0^c E_0}^+ \otimes \phi_{F_0^c F_0}^+) \right]. \end{aligned} \quad (5.3.17)$$

Now divide the full trace above as $\text{tr}_{E_0^c F_0^c} [\text{tr}_{E_1 F_1} (\cdots)]$. By assumption, $\{\Theta^x\}_x$ and $\{\Upsilon^y\}_y$ are IC-POVMs and therefore form a basis for the linear operators acting on the Hilbert spaces $\mathcal{H}_{E_0^c}$ and $\mathcal{H}_{F_0^c}$. As a result, for the partial trace over E_1 and F_1 , we have the following

$$\begin{aligned} \text{tr}_{E_1 F_1} \left[\text{id}_{E_0^c} \otimes \bar{Z}_{E_1 F_1}^{ab} \otimes \text{id}_{F_0^c} \cdot \mathcal{E}_{EF} \otimes \text{id}_{E_0^c F_0^c} (\phi_{E_0^c E_0}^+ \otimes \phi_{F_0^c F_0}^+) \right] &= \\ \text{tr}_{E_1 F_1} \left[\text{id}_{E_0^c} \otimes Z_{E_1 F_1}^{ab} \otimes \text{id}_{F_0^c} \cdot \mathcal{E}'_{EF} \otimes \text{id}_{E_0^c F_0^c} (\phi_{E_0^c E_0}^+ \otimes \phi_{F_0^c F_0}^+) \right]. \end{aligned} \quad (5.3.18)$$

Let us now recover the notation $E_1 = A_1 R_1 = A'_1 R'_1$ and $F_1 = B_1 S_1 = B'_1 S'_1$ while noticing that $R_1 = R_0$, $R'_1 = R'_0$ and the same for S . Thus, we can write

$$\begin{aligned} \text{tr}_{\substack{A_1 R_0 \\ B_1 S_0}} \left[\text{id}_{A_0^c R_0^c} \otimes \bar{Z}_{\substack{A'_1 R'_0 \\ B'_1 S'_0}}^{ab} \otimes \text{id}_{B_0^c S_0^c} \cdot \mathcal{E}_{AB} \otimes \text{id}_{RS} \otimes \underbrace{\text{id}_{\substack{A_0^c R_0^c \\ B_0^c S_0^c}} \left(\phi_{A_0^c A_0}^+ \otimes \phi_{B_0^c B_0}^+ \otimes \phi_{R_0^c R_0}^+ \otimes \phi_{S_0^c S_0}^+ \right)}_{J_{A_1 B_1 A_0^c B_0^c} \otimes \phi_{R_0^c R_0}^+ \otimes \phi_{S_0^c S_0}^+} \right] &= \\ \text{tr}_{\substack{A'_1 R'_0 \\ B'_1 S'_0}} \left[\text{id}_{A_0^c R_0^c} \otimes Z_{\substack{A'_1 R'_0 \\ B'_1 S'_0}}^{ab} \otimes \text{id}_{B_0^c S_0^c} \cdot \mathcal{E}'_{A'B'} \otimes \text{id}_{R'S'} \otimes \underbrace{\text{id}_{\substack{A_0^c R_0^c \\ B_0^c S_0^c}} \left(\phi_{A_0^c A_0}^+ \otimes \phi_{B_0^c B_0}^+ \otimes \phi_{R_0^c R_0}^+ \otimes \phi_{S_0^c S_0}^+ \right)}_{J'_{A'_1 B'_1 A_0^c B_0^c} \otimes \phi_{R_0^c R_0}^+ \otimes \phi_{S_0^c S_0}^+} \right]. \end{aligned} \quad (5.3.19)$$

On each side of the equality, we identify the Choi states J and J' tensored by two pairs of Bell states ϕ^+ . This allows for a physical interpretation in terms of teleportation similar to [56], with the difference that now two parts out of the four-partite Choi state J' are teleported. That is, here Alice holds the systems A'_1 , B'_1 , R'_0 and S'_0 and Bob holds the systems A_0^c , B_0^c , R_0^c and S_0^c . Using two shared Bell states, $\phi_{R_0^c R_0}^+$ and $\phi_{S_0^c S_0}^+$, Alice teleports her parts A'_1 and B'_1 of the shared Choi state J' to Bob's parts R_0^c and S_0^c .

By assumption, for any Z^{ab} there exists a $\bar{Z}^{ab}$ such that the equality Eq. (5.3.19) holds. In particular, it holds when Alice applies a Bell measurement $\{Z_{A'_1 B'_1 R'_0 S'_0}^{ab} = \{\phi_{A'_1 R'_0}^{ab} \otimes \phi_{B'_1 S'_0}^{ab}\}$ to send her side to Bob, where ϕ^{ab} is the density matrix of each generalized Bell state $|\phi^{ab}\rangle = \sum_{k=0}^{d-1} e^{bk \frac{2\pi i}{d}} |k, k+a\rangle / \sqrt{d}$ [8]. In this case, Bob can apply unitaries $\{U_{R_0^c}^a \otimes V_{S_0^c}^b \otimes \text{id}_{A_0^c B_0^c}\}$ to recover J' in his side. Strictly speaking,

$$\begin{aligned} J'_{A_0^c B_0^c A'_1 B'_1} &= \sum_{a,b} (U_{R_0^c}^a \otimes V_{S_0^c}^b \otimes \text{id}_{A_0^c B_0^c}). \quad (5.3.20) \\ \text{tr}_{\substack{A_1 R_0 \\ B_1 S_0}} \left[\text{id}_{A_0^c R_0^c} \otimes \bar{Z}_{\substack{A'_1 R'_0 \\ B'_1 S'_0}}^{ab} \otimes \text{id}_{B_0^c S_0^c} \cdot J_{A_1 B_1 A_0^c B_0^c} \otimes \phi_{R_0^c R_0}^+ \otimes \phi_{S_0^c S_0}^+ \right] &\cdot (U_{R_0^c}^a \otimes V_{S_0^c}^b \otimes \text{id}_{A_0^c B_0^c})^\dagger, \end{aligned}$$

where we inserted the left-hand side of Eq. (5.3.19) as the (unnormalized) state Bob obtains for each outcome a and b of Alice. Finally, expanding $\bar{Z}^{ab} = \sum_i \nu_i \bar{M}_{A_1 R_0}^{a,i} \otimes \bar{N}_{B_1 S_0}^{b,i}$ with $\nu_i \geq 0$ and

$\sum_i \nu_i = 1$, one gets the following maps

$$\Lambda_i(\tau_{A_1 A_0^c}) := \sum_a (U_{R_0^c}^a \otimes \text{id}_{A_0^c}) \cdot \text{tr}_{A_1 R_0} [\text{id}_{A_0^c R_0^c} \otimes \overline{M}_{A_1 R_0}^{a,i} \cdot \phi_{R_0 R_0^c}^+ \otimes \tau_{A_1 A_0^c}] \cdot (U_{R_0^c}^a \otimes \text{id}_{A_0^c})^\dagger \quad \text{and} \quad (5.3.21)$$

$$\Phi_i(\theta_{B_1 B_0^c}) := \sum_b (V_{S_0^c}^b \otimes \text{id}_{B_0^c}) \cdot \text{tr}_{B_1 S_0} [\text{id}_{B_0^c S_0^c} \otimes \overline{N}_{B_1 S_0}^{b,i} \cdot \phi_{S_0 S_0^c}^+ \otimes \theta_{B_1 B_0^c}] \cdot (V_{S_0^c}^b \otimes \text{id}_{B_0^c})^\dagger. \quad (5.3.22)$$

One can verify that $\text{tr}_{R_0^c}[\Lambda_i(\tau_{A_1 A_0^c})] = \text{id}_{A_0^c}$ if $\text{tr}_{A_1}(\tau_{A_1 A_0^c}) = \text{id}$, thus each Λ_i is a quantum superchannel. Similarly, each Φ_i is a superchannel. It is now clear that the Choi matrix J' can be obtained through local operations assisted by shared randomness from J ,

$$J'_{A_0^c B_0^c A_1 B_1} = \sum_i \nu_i \Lambda_i \otimes \Phi_i(J_{A_1 B_1 A_0^c B_0^c}), \quad (5.3.23)$$

which means that the channel $\mathcal{E}'$ can be obtained from the channel $\mathcal{E}$ by a convex combination of products of superchannels. $\square$

Similarly as in [56], this means that a quantum channel $\mathcal{E}$ is not LOSR if and only if it has a strictly larger value $\gamma_{\max}(\mathcal{E})$ than any LOSR operation for some Bell functional γ and quantum inputs. The following derivations provide explicit necessary and sufficient conditions to certify non-LOSR operations. Consider a set of fixed states $\{\varrho_{RA_0}^x\}$ and $\{\sigma_{SB_0}^y\}$ that are linearly independent and span the whole input spaces of linear operators acting on $\mathcal{H}_{RA_0}$ and $\mathcal{H}_{SB_0}$, and assume $\mathcal{H}_{A_0} \mathcal{H}_R$ and $\mathcal{H}_{B_0} \mathcal{H}_S$. Let $\{\Psi_{RA_1}^a\}$ and $\{\Psi_{SB_1}^b\}$ be projective measurements on a generalized Bell basis and denote

$$p_{\mathcal{E}}(ab|xy) := \text{tr}(\Psi_{RA_1}^a \otimes \Psi_{SB_1}^b \cdot (\mathcal{E}_{AB} \otimes \text{id}_{RS})[\varrho_{RA_0}^x \otimes \sigma_{SB_0}^y]). \quad (5.3.24)$$

This setup allows to formulate the problem of whether or not a channel $\mathcal{E}_{AB}$ is LOSR as follows: A quantum channel $\mathcal{E}_{AB}$ is not an LOSR operation if and only if for any possible LOSR operation $\mathcal{E}'_{A'B'}$ and for any choice of measurements $M_{RA_1}^a$ and $N_{SB_1}^b$, the distribution $p_{\mathcal{E}}(ab|xy)$ of Eq. (5.3.24) cannot be obtained via $\text{tr}(M_{RA_1}^a \otimes N_{SB_1}^b (\mathcal{E}'_{AB} \otimes \text{id}_{RS})[\varrho_{RA_0}^x \otimes \sigma_{SB_0}^y])$. By absorbing the adjoint $\mathcal{E}'_{AB}^\dagger$ in the measurements $M^a \otimes N^b$, namely defining $\tilde{M}^a \otimes \tilde{N}^b := \mathcal{E}'^\dagger \otimes \text{id}(M^a \otimes N^b)$, one arrives at the following criterion.

Corollary 5. *A bipartite quantum channel $\mathcal{E}_{AB}$ is not LOSR if and only for some functional γ*

$$\gamma(p_{\mathcal{E}}) > \max_{\tilde{M}^a, \tilde{N}^b} \sum_{ab, xy} \gamma_{ab, xy} \text{tr}(\tilde{M}^a \varrho^x) \text{tr}(\tilde{N}^b \sigma^y) \quad (5.3.25)$$

where we denote $\gamma(p_{\mathcal{E}}) := \sum_{ab, xy} \gamma_{ab, xy} p_{\mathcal{E}}(ab|xy)$.

This provides a separation between the set of LOSR operations and the rest, which relies only in the input/outcome probabilities and the trusted set of input states ϱ^x and σ^y . Therefore, it constitutes a measurement-device independent characterization of the set of LOSR operations. Note that the left hand side of Eq. (5.3.25) is given by the set of input states and the channel under consideration, and the right hand side can be upper-bounded with semidefinite programming [245].

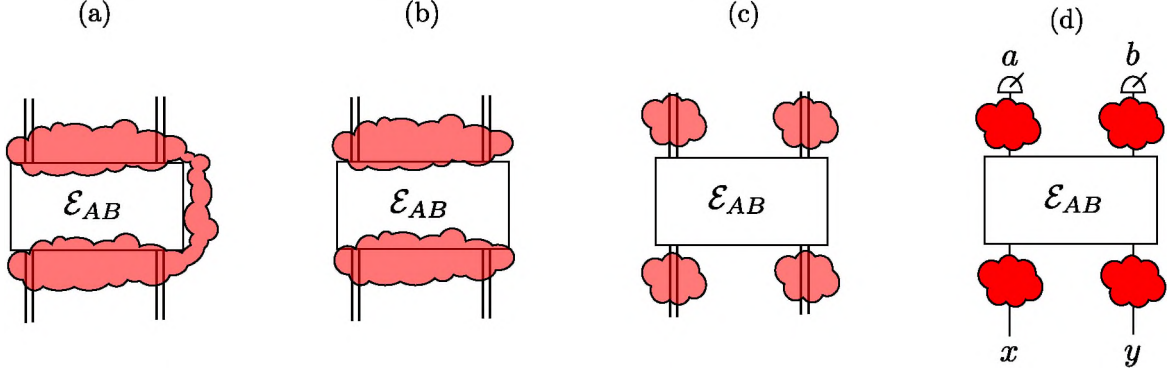


Figure 5.3.3: Models of decoherence occurring by dephasing noise under consideration. (a) Dephasing occurs both in the inputs and the outputs as a superechannel $\Xi_{AB}[\mathcal{E}_{AB}]$. Its action on the Choi matrix is given by a Gram matrix $G_{A_0 A_1 B_0 B_1}$ without any tensor product assumption through a bipartite Eq. (1.1.38) with a bipartite structure. (b) Nonlocal dephasing occurs as independent pre- and post-processing by composition with a dephasing channel, $\mathcal{E}_{AB}^{\mathcal{D}} = \mathcal{D}^G \circ \mathcal{E} \circ \mathcal{D}^{G'}$. Its action on the Choi matrix is given by a product of Gram matrix $G_{A_0 B_0} \otimes G_{A_1 B_1}$ through Eq. (1.1.39) with a bipartite structure. (c) Local dephasing occurs at each party independently, according to Eq. (5.4.1). (d) Local complete dephasing occurs, where the input and output states are diagonal density matrices. The last scheme representing complete decoherence corresponds to Fig 5.3.1(c), since local projective measurements in a fixed basis occur both before and after the channel is applied.

5.4 The role of dephasing noise

5.4.1 Gradual dephasing

Quantum channels, and in particular the protocols described above to certify nonlocality, can suffer from decoherence. Here we ask how dephasing noise affects the nonlocal properties of a channel. Let us first consider local dephasing (Fig. 5.3.3 (c)), which is given by a local four partite Gram matrix of the form

$$G = G_{A_0} \otimes G_{B_0} \otimes G_{A_1} \otimes G_{B_1}. \quad (5.4.1)$$

Observation 7. *If a bipartite channel undergoing local dephasing noise, $\mathcal{D}_{A_1} \otimes \mathcal{D}_{B_1} \circ \mathcal{E}_{AB} \circ \mathcal{D}_{A_0} \otimes \mathcal{D}_{B_0}$, can produce correlations outside of L (Q , NS) from Eq. (5.3.7), then the channel $\mathcal{E}_{AB}$ is not $LOSR$ ($LOSE$, qNS).*

Indeed, note that the Schur product with a Gram matrix preserves both positive semidefiniteness, the trace, and identity resolution of any operators. Therefore, local dephasing is absorbed in the preparation and measurement. This means that if we detect nonlocality on a channel which suffered from local dephasing, then we know that its noiseless version previous to dephasing is also nonlocal.

Let us now consider the case in which the two parties A and B sharing a quantum state ϱ_{AB} or implementing a quantum channel $\mathcal{E}_{AB}$ belong to a joint subsystem. In this case, dephasing can occur as a joint operation (Fig. 5.3.3 (a & b)), and thus one might expect that nonlocal phenomena can occur. This can be summarized as follows.

Proposition 11. *Dephasing noise acting jointly on a bipartite system has the following properties:*

1. *A dephasing channel acting on a separable bipartite state can generate an entangled state.*
2. *A dephasing superchannel acting on an $LOSR$ operation can generate a nonlocal channel.*

At the level of quantum states, an extreme example can be constructed as follows: Consider the states $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Note that the projector $\varrho_{AB} = |++\rangle\langle ++|$ has homogeneous

entries $\varrho_{ij} = 1/4$, thus it evolves under dephasing noise $\mathcal{D}_{AB}^G$ in Eq. (1.1.40) given by some Gram matrix G as

$$\mathcal{D}_{AB}^G(\varrho_{A_0B_0}) = |++\rangle\langle++| \odot G = G/4. \quad (5.4.2)$$

Therefore, it is enough to show that there exists a quantum state that is both maximally entangled and a Gram matrix. Indeed, the Bell state $|\psi\rangle = (|0+\rangle + |1-\rangle)/\sqrt{2}$ has density matrix

$$|\psi\rangle\langle\psi| = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \quad (5.4.3)$$

which defines a Gram matrix $G = 4|\psi\rangle\langle\psi|$ with entries $G_{ij} = \langle v_i | v_j \rangle$ where $|v_1\rangle = |v_2\rangle = |v_3\rangle = -|v_4\rangle = (1, 1, 1, -1)^T/2$. Since $|\psi\rangle$ is maximally entangled, it maximally violates Clauser-Horne-Shimony-Holt (CHSH) inequality,

$$\langle \sigma_Z \otimes (\sigma_+ + \sigma_-) + \sigma_X \otimes (\sigma_+ - \sigma_-) \rangle_\psi = 2\sqrt{2}, \quad (5.4.4)$$

where σ_X and σ_Z are Pauli matrices and $\sigma_\pm = (\sigma_X \pm \sigma_Z)/\sqrt{2}$. From a resource theory perspective, this is possible because the state $|++\rangle := |+\rangle \otimes |+\rangle$ maximizes the resource of coherence (with respect to the computational basis) [57, 246], from which a dephasing channel can generate entanglement.

At the level of quantum channels, this example can be seen as follows: Consider the identity channel id , with the bipartite input state $|++\rangle$. Although the identity channel is an LOSR operation and thus produces probability distributions in local measurement protocols like the one of Eq. (5.3.7) (Fig 5.3.1), dephasing induced by coupling degrees of freedom with the environment can convert the channel id into a channel capable of generating nonlocal quantum probability distributions. This is an example where dephasing occurs as a superchannel with independent pre- and post-processing [62, 64] (Fig. 5.3.3 (b)).

An example where dephasing is induced by a more general superchannel (Fig. 5.3.3 (a)) is as follows: Consider the Choi matrix J^{U_L} of the local unitary $U_L = \text{id} \otimes \sigma_X$. Let the dephasing superchannel be described by a 16×16 Gram matrix $G = G^\dagger \geq 0$ with nonzero coefficients $G_{kk} = 1$, $G_{12,15} = -G_{2,5} = 1$, and $G_{2,12} = G_{5,12} = -G_{2,15} = G_{5,15} = i$ (and their Hermitian adjoints), where i is the imaginary unit. This dephasing leads to the Choi matrix $J^U = J^{U_L} \odot G$ which defines the channel implementing the following unitary,

$$U = \sum_{j=0}^1 |j\rangle\langle j| \otimes \sigma_X(i\sigma_Z)^j. \quad (5.4.5)$$

Observe that this unitary can be converted into the C-NOT gate by local unitaries as follows,

$$J_{NOT} = (\text{id} \otimes H \sigma_X) U (\text{id} \otimes H)^\dagger, \quad (5.4.6)$$

and therefore U can produce extremal signaling probability distributions in the protocol in Fig. 5.3.1 (b) when decoherence depicted in Fig. 5.3.3 (a) occurs.

Let us highlight that the examples above constitute a special case of dephasing channels and superchannels occurring to highly coherent states and channels, whose coherence only suffer from phase factors. Although these examples are consistent with the mathematical model of dephasing operations established in [62], generic dephasing transforms pure states into mixed states, and unitary channels into irreversible channels [21]. Yet, these examples are illustrative for the following comparison: Observation 7 captures the intuitive fact that decoherence might make a nonlocal quantum channel undetectable through measurement protocols, thus giving a *false negative* result to a test of nonlocality. On the contrary, Proposition 11 shows that decoherence can also give

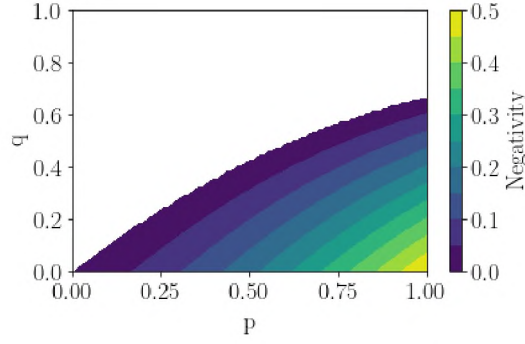


Figure 5.4.1: Entanglement generated by dephasing. The parameters p and q quantify the strength of the dephasing channels $\mathcal{D}_q$ and $\mathcal{D}'_p$ acting on the maximally coherent product state $|++\rangle\langle++|$ [57, 246], which can be mapped to less coherent but entangled states.

rise to *false positive* results if the parties A and B perform a measurement protocol undergoing nonlocal noise.

To better understand the role of gradual decoherence aside from extremal cases, we consider two parametric types of dephasing models. On one hand, $\mathcal{D}_q$ damps all off-diagonal terms equally, $\mathcal{D}_q(\varrho) = (1 - q)\varrho + q \text{diag}(\varrho)$. On the other hand, $\mathcal{D}'_p$ keeps all entries except for the coherences to the fourth population level, namely $\mathcal{D}'_p(\varrho)_{ij} = (1 - 2p)\varrho_{ij}$ if either $i = 4 \neq j$ or $j = 4 \neq i$.

Starting from the maximally coherent separable state $\varrho = |++\rangle\langle++|$, the composition of these two noises reads

$$\mathcal{D}_q \circ \mathcal{D}'_p(\varrho) = \frac{1}{4} \begin{pmatrix} 1 & 1-q & 1-q & x \\ 1-q & 1 & 1-q & x \\ 1-q & 1-q & 1 & x \\ x & x & x & 1 \end{pmatrix} \quad (5.4.7)$$

with $x = (1 - q)(1 - 2p)$. The negativity of this state for the parameters p and q is shown in Fig. 5.4.1.

5.4.2 Complete decoherence

Here we consider the case of complete decoherence depicted in Fig. 1.1.3, where Choi matrices become diagonal. Recall that the set of stochastic matrices and the set of conditional probability distributions are isomorphic. For bipartite channels $\mathcal{E}_{AB}$, the decoherent action is represented by a bipartite stochastic matrix with entries

$$S_{ab,xy}^{\mathcal{E}} = \langle ab | \mathcal{E}(|xy\rangle\langle xy|) | ab \rangle, \quad (5.4.8)$$

where $(a, b, x, y) \in [d_{A_1}] \times [d_{B_1}] \times [d_{A_0}] \times [d_{B_0}]$ define the canonical basis (Fig. 1.1.3). Therefore, the decoherent action (5.4.8) defines probabilities $p(a, b|x, y) = S_{ab,xy}$ of obtaining classical outputs ab conditioned to classical inputs xy , through the scheme of Fig. 5.3.1 (c). In [229, Propositions 3-5] it was shown that the sets of L, Q and NS correlations correspond to the sets of correlations that can be obtained respectively from LOSR, LOSE, and qNS channels as correlations in the inputs and outputs when all measurement bases are allowed. A similar result can be formulated for a fixed basis, namely considering the classical action of Eq. (5.4.8), as follows.

Proposition 12. *The following equivalences hold:*

- (i) *The set of local (L) correlations coincides with the set of decoherent actions of LOSR channels.*
- (ii) *The set of quantum (Q) correlations coincides with the set of decoherent actions of LOSE channels.*

(iii) The set of nonsignaling (NS) correlations coincides with the set of decoherent actions of quantum nonsignaling channels qNS.

Proposition 12 can be obtained from [229, Prop. 3, 4 & 5], by absorbing the basis choice into the existence conditions. An explicit constructive proof is detailed below.

Proof. Let us show first item (ii). Let $p(a, b|x, y) = \text{tr}(\sigma P^{a|x} \otimes Q^{b|y})$ determine a quantum probability distribution, where σ is a quantum state on RS , $\{P^{a|x}\}_a$ projective measurements over R and $\{Q^{b|y}\}_b$ over S . Going to a Stinespring's dilation, there exist unitary maps $U_x : AR \rightarrow AR$ and $V_y : BS \rightarrow BS$ such that

$$P^{a|x} = \langle a| U_x |0\rangle \quad \& \quad Q^{b|y} = \langle b| V_y |0\rangle. \quad (5.4.9)$$

Take $U : CAR \rightarrow CAR$ and $V : DBS \rightarrow DBS$ unitaries on some control systems C and D ,

$$U = \sum_x |x\rangle \langle x| \otimes U_x \quad \& \quad V = \sum_y |y\rangle \langle y| \otimes V_y. \quad (5.4.10)$$

These define a LOSE channel $\mathcal{E} : CD \rightarrow AB$ via

$$\mathcal{E}(\varrho) = \text{tr}_{CDS}[U \otimes V(\varrho \otimes |00\rangle\langle 00| \otimes \sigma)U^\dagger \otimes V^\dagger] \quad (5.4.11)$$

whose decoherent action gives back $p(a, b|x, y)$. Indeed,

$$\langle ab| U \otimes V(|xy\rangle\langle xy| \otimes |00\rangle) = |xy\rangle \otimes P^{a|x} \otimes Q^{b|y}, \quad (5.4.12)$$

thus $\text{tr}[|ab\rangle\langle ab| \mathcal{E}(|xy\rangle\langle xy|)]$ reads

$$\text{tr}[|xy\rangle\langle xy| \otimes P^{a|x} \otimes Q^{b|y} \sigma] = p(a, b|x, y). \quad (5.4.13)$$

Conversely, assume $p(a, b|x, y)$ can be obtained as the decoherent action of a LOSE channel, $\text{tr}[|ab\rangle\langle ab| \otimes \text{id}]U \otimes V(|xy\rangle\langle xy| \otimes \sigma)U^\dagger \otimes V^\dagger]$, where $U : AR \rightarrow AR$ and $V : BS \rightarrow BS$ are unitaries and σ a state on RS . With the isometries

$$U_x = U |x\rangle \quad \& \quad V_y = V |y\rangle \quad (5.4.14)$$

we can define quantum effects

$$E^{a|x} = U_x^\dagger |a\rangle\langle a| U_x \quad \& \quad F^{b|y} = V_y^\dagger |b\rangle\langle b| V_y, \quad (5.4.15)$$

that reproduce the probabilities $p(a, b|x, y)$. Indeed,

$$\text{tr}[U_x^\dagger \otimes V_y^\dagger |xy\rangle\langle xy| U_x \otimes V_y \sigma] = \text{tr}[E^{a|x} \otimes F^{b|y} \sigma]. \quad (5.4.16)$$

Item (i) follows from the above derivations by considering σ to be a separable state. Item (iii) can be seen by realizing that the qNS conditions on a diagonal Choi matrix are equivalent to the NS conditions on the probability distribution given by the decoherent action of the channel. $\square$

This provides a simple criterion to detect certain classes of quantum channels, from local measurements in the computational basis: One certifies that a quantum channel $\mathcal{E}_{AB}$ is outside of the set of LOSR (LOSE, qNS) channels if its decoherent action $S_{AB}^\mathcal{E}$ is outside of the set of L (Q, NS) correlations. This criterion can be implemented in decoherent setups, as it involves only measurements in the computational basis.

Note that the converse is not true: there exist non-LOSR quantum channels with local decoherent action. As an example, consider the 2-qubit channel with unitary (single) Kraus op-

erator $U = \sum_{i,j=0}^1 |\phi_{ij}\rangle \langle i+j, j|$, which maps the computational basis to the Bell states $|\phi_{ij}\rangle = \sigma_X^i \sigma_Z^j \otimes \mathbb{1}_2 (|00\rangle + |11\rangle) / \sqrt{2}$ where σ_X and σ_Z are Pauli matrices. Even though U can generate entanglement, its decoherent action $S^U = U \odot U^* = \frac{1}{2} (\mathbb{1}_2^{\otimes 2} + \sigma_X^{\otimes 2})$ is local. However, note that this criterion can be enhanced with any operations that preserve the set of LOSR. In the example above, the freedom of local measurement basis considered in [229] is enough: the bipartite unitary U can be transformed by local unitaries to $\tilde{U} = (H \otimes H)U(\text{id} \otimes H)$, where H is the Hadamard gate with $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$. We have $S^{\tilde{U}} = \tilde{U} \odot \tilde{U}^* = |0\rangle\langle 0| + |1\rangle\langle 3| + |2\rangle\langle 2| + |3\rangle\langle 1|$ which defines extremal nonsignaling correlations.

Note that the nonsignaling conditions establish equality constraints, and thus define a measure zero subset of channels. Moreover, the sets of local and nonsignaling correlations form a polytope. Therefore, detecting channels with nonlocal decoherent action can be done with a linear program, asking whether there exists a convex combination of extremal local stochastic matrices decomposing the decoherent action. For local correlations, finitely-many Bell games [52, 247, 248] equivalently define the faces of their polytope. Bounding the set of quantum correlations is a more involved problem and relies on a convergent hierarchy [249]. This problem has been tackled through analytical solutions for the smallest setting [250–252], conjectured generalizations [68, 253] and additional results using semidefinite programming techniques [254, 255]. Outer approximations are given by the maximum quantum value of Bell games [256].

By Proposition 12, Bell functionals on the decoherent action of a quantum channel can be used to construct linear nonlocality witnesses, similar to entanglement witnesses [4]. This mechanism is illustrated in Fig. 5.3.2. Let $\mathcal{L}$ be one of the three subsets of bipartite quantum channels: LOSR, LOSE, and qNS. Let S be the corresponding set of bipartite probability distributions under the correspondence in Prop. 12, namely, L, Q, and NS. Consider

$$W_{\mathcal{L}} = \gamma_S / (d_{A_0} d_{B_0}) \text{id} - \Omega, \quad (5.4.17)$$

where Ω is defined from the Bell functional through

$$\Omega = \sum_{x,y,a,b=0}^1 \gamma_{ab,xy} |xayb\rangle \langle xayb|_{A_0 A_1 B_0 B_1} \quad (5.4.18)$$

and γ_S denotes its maximum value on the subset S . Thus, a channel $\mathcal{E}_{AB}$ with Choi matrix $J^{\mathcal{E}}$ and decoherent action $S^{\mathcal{E}}$ is outside $\mathcal{L}$ if

$$\text{tr}(J^{\mathcal{E}} W_{\mathcal{L}}) = \gamma_S - \sum \gamma_{ab,xy} S_{ab,xy}^{\mathcal{E}} < 0. \quad (5.4.19)$$

For the CHSH functional, $\gamma_{ab,xy} = (-1)^{a+b}(-1)^{xy}$, it holds that $\gamma_L = 2$, $\gamma_Q = 2\sqrt{2}$ and $\gamma_{NS} = 4$.

On the other hand, by the isomorphism of Proposition 12, any bipartite stochastic matrix in a distinguished subset of correlations can be obtained from decoherence of a quantum channel in a corresponding subset. An explicit example is the convex hull of bipartite stochastic matrices (namely, bipartite probability distributions) $pR + (1-p)S$ with

$$R = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad S = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.4.20)$$

which attains the maximal quantum value $\gamma_Q = 2\sqrt{2}$ of CHSH for $p = (1 + 1/\sqrt{2})/2$. In Fig. 5.4.2 we show a two dimensional slice of the polytope of distributions consisting of the convex hull of R , S and id containing different subsets. The sets of local (L) and nonsignaling (NS) distributions are bounded by linear programming, while the set of quantum (Q) distributions is approximated by the first two levels of the Navascués-Pironio-Acín (NPA) hierarchy [249]. The subsets L, Q and

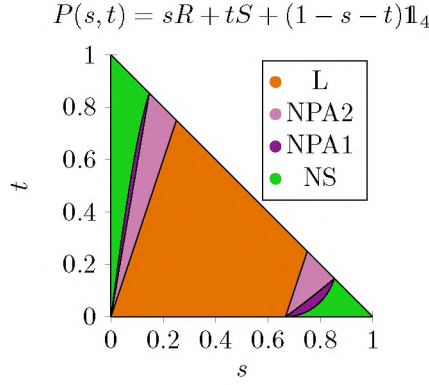


Figure 5.4.2: Partial cross section of the space of bipartite distributions spanned by id_4 , R and S in Eq. (5.4.20), $P(s, t) = sR + tS + (1 - s - t)\text{id}_4$ where $s + t \leq 1$. The distribution id_4 is local, while R and S are extremal non-signaling (PR-boxes). The orange region corresponds with distributions that are local, while dark and light purple regions denote first (NPA1) and second (NPA2) level of Navascués-Pironio-Acín (NPA) approximation [249] to the set Q of quantum distributions (not plotted). Green region depicts the remaining non-signaling distributions. Each of these subsets of distributions can be obtained by LOSR, LOSE and qNS channels under complete decoherence.

NS of distributions are obtained by decoherence of the subsets LOSR, LOSE and qNS of channels.

5.5 Conclusions and open questions

We infer nonlocal properties of bipartite quantum channels based on the bipartite stochastic matrices produced by them under measurement protocols. In particular, our approach allows us to bound the sets of local operations assisted by shared randomness (LOSR), local operations assisted by shared entanglement (LOSE) and quantum non-signaling channels (qNS), from the conditional probabilities of local measurement outcomes. This framework is suitable to experimental verification of nonlocality of bipartite channels from local measurements, even in the presence of complete decoherence. In contrast, we show that both entanglement and nonlocality of bipartite quantum channels can be enhanced under nonlocal gradual decoherence.

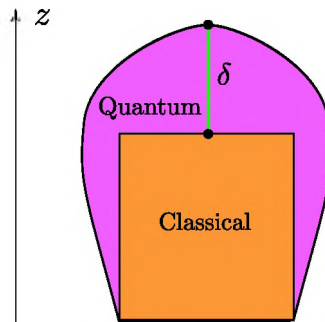


Figure 5.5.1: First glimpse at the structure of multidimensional convex sets of *quantum* and *classical* states, operations and correlations. The transition from quantum to classical is induced by decoherence. The same figure can also represent sets of quantum algorithms or quantum technologies. The distance δ between the extremal points of both sets along the vertical axis z , denoting a selected utility function, represents the *quantum advantage*.

This work relates the geometry of the set of bipartite quantum states, channels, superchannels, and their classical counterparts: probability vectors, stochastic matrices and classical superchannels. In particular, several measures for nonlocality in the space of correlations have been proposed

in the literature [74, 257–259]. Through our framework, these naturally give insight into the geometry of quantum operations [77]. At the same time, our work leaves several research lines open:

1. It is unknown whether local-limited and quantum-limited channels as defined in [229] can be detected by making use of a shared entangled state through Proposition 9.
2. Corollary 5 provides a first step to understanding the set of functionals witnessing non-LOSR operations, analogous to entanglement witnesses for nonseparable states. However, further research is needed to better understand the performance of such witnesses in practice.

The analysis presented for the case of complete decoherence highlights the role of Choi–Jamiołkowski isomorphism to analyze geometric features of quantum correlations and channel nonlocality. The standard form of the isomorphism relates bipartite states with quantum operations, while its generalized version links four-partite quantum states with bipartite quantum channels. Due to decoherence the latter suffer a transition to bipartite stochastic matrices $\mathcal{S}_{d^2}$, which describe correlations between both systems. Therefore, the overall geometric structure of the sets of classical and quantum states, channels and correlations is rooted in the geometric structure of the set $\mathcal{M}_{d^4}$ four-partite states (Fig. 1.1.3).

Any generalized quantum theory can be recovered from basic principles that force the corresponding set of states to be convex [3]. A closer look reveals its convex subset of separable states, defined as classical mixtures of product states – see Fig. 5.5.1. A related structure is characteristic also to the space of quantum maps and quantum correlations. A similar picture can also be used to describe the sets of quantum algorithms and quantum technologies and to visualize the notion of a *quantum advantage*: The difference between the utility function describing, for instance, efficiency of a given protocol, optimized over all quantum and classical scenarios.

6

Nonlocality of decoherent quantum channels

The implementation of realistic quantum devices requires a solid understanding of the nonlocal resources present in quantum channels, and the effects of decoherence on them. In this chapter we quantify nonlocality of bipartite quantum channels and identify its component resisting the effects of dephasing noise. Despite its classical nature, we demonstrate that the latter plays a relevant role in performing quantum protocols, such as state transformations and quantum coding for noisy communication. In the converse direction, we show that simulating certain stochastic processes with quantum channels undergoing decoherence has a communication advantage with respect to their classical simulation.

6.1 Introduction

The phenomenon of Bell nonlocality lies at the core of quantum theory, as quantum distributions cannot always be explained by hidden variables assuming local realism [51, 53]. Nonlocality constitutes a resource for device-independent protocols in cryptography and information theory, such as public key distribution schemes, conjugate coding, and randomness verification [225–227]. Nonlocal resources can be present both in the quantum state describing a system [52], and in a quantum operation manipulating such state [7, 111, 229, 231, 235].

In practice, both the storage of a quantum state and its manipulation through a quantum channel suffer from the phenomenon of *decoherence* [21, 237]: the interaction with the environment can fully or partially spoil quantum properties [58, 238, 239]. Therefore, successful realization of noisy quantum devices demands a deep understanding of quantum properties under decoherence [240, 260–262]. For this purpose, a well-established model of decoherence is the dephasing noise $\mathcal{D}$, which damps the off-diagonal terms in a density matrix while keeping intact the terms diagonal in a preferred basis [57]. In the worst-case scenario, complete decoherence leads to vanishing of the off-diagonal terms. This reduces a quantum state to a probability vector [57, 240–242] and a quantum channel to a stochastic map described by a stochastic matrix [62].

In recent years, increasing attention has been paid to understanding the nonlocal properties of quantum operations. When no shared resources are allowed, the operator Schmidt decomposition of a bipartite unitary gate [263, 264] or a quantum channel [265] determines whether it can be performed locally. On the other hand, linear tests for nonlocality of states [52, 56] can be lifted to the level of channels [7]. There, the probability distributions obtained by local quantum measurements [7, 15, 229, 231] reveal whether a bipartite channel can be implemented locally with shared randomness. While existing methods provide a qualitative characterization, a quantitative analysis contemplating noisy effects is desired for a further step in understanding both the nonlocal resources of channels and the effects of decoherence on them.

In this work we quantify the nonlocal resources of bipartite channels with several measures,

and their corresponding components remaining when complete decoherence occurs. The latter determine the nonlocal transformations that a channel can induce on the computational basis, e.g. bit pairs. Using a semidefinite programming hierarchy for bilinear optimization [245], we evaluate the role of nonlocal resources in performing quantum protocols. Despite accounting for a classical resource, the nonlocality resisting decoherence is found to be required for optimal performance. In a similar direction, we compare two different protocols to simulate discrete Markov chains [58], namely using standard classical methods with respect to using a quantum channel that undergoes complete decoherence. We show that the latter gives advantage in the communication needed.

Some preliminaries will be reminded to the reader for completeness of the chapter. This work relies on the bridge between two seemingly different structures [15, 229, 231, 235, 244, 266]: the polytope of bipartite distributions [52], and the set of bipartite quantum channels [109, 112, 113, 267]. A bipartite distribution $p(ab|xy)$ is *local* (L) if it decomposes as $p(a, b|x, y) = \sum_{\lambda} p_{\lambda} p(a|x, \lambda) p(b|y, \lambda)$ with $p_{\lambda} \geq 0$ and $\sum_{\lambda} p_{\lambda} = 1$, and nonlocal otherwise [52]. Distributions obtained from bipartite quantum states ϱ_{AB} via Born's rule, $p(a, b|x, y) = \text{tr}(P^{a|x} \otimes Q^{b|y} \varrho_{AB})$ for local measurement effects $P^{a|x}$ and $Q^{b|y}$, are called *quantum* (Q). Nonlocal resources present in quantum states [51, 54] cannot increase under local operations with shared randomness (LOSR) [7, 56, 111, 236], $\Phi_{AB} = \sum_{\lambda} p_{\lambda} \Phi_A^{\lambda} \otimes \Phi_B^{\lambda}$. Bipartite channels outside the set of LOSR are called *nonlocal* [15]. The set of LOSR operations is preserved under local superchannels with shared randomness, $\sum_{\lambda} p_{\lambda} \Xi_A^{\lambda} \otimes \Xi_B^{\lambda}$, which are the free operations in the resource theory of nonlocality of channels [7].

6.2 Quantum and classical nonlocality

6.2.1 Quantifying nonlocality of quantum channels

The nonlocal resources available in a quantum channel $\mathcal{E} : A_0 B_0 \rightarrow A_1 B_1$ will be treated as a convex resource theory [268], where LOSR operations are the free elements. Therefore, we quantify the nonlocal resources of a channel as

$$\mathcal{N}_{\Delta}(\mathcal{E}) = \min\{\Delta(\mathcal{E}, \Phi) : \Phi \in \text{LOSR}\} \quad (6.2.1)$$

for a suitable non-negative functional Δ .

Proposition 13. *Desired properties of the each measure $\mathcal{N}_{\Delta}$ are determined by those of the corresponding function Δ , as follows.*

- (i) *If Δ is continuous, then $\mathcal{N}_{\Delta}$ attains the minimum.*
- (ii) *If Δ is convex, then $\mathcal{N}_{\Delta}$ is convex.*
- (iii) *If Δ is non-degenerate, then $\mathcal{N}_{\Delta}$ is faithful.*
- (iv) *If Δ is contractive under superchannels, then $\mathcal{N}_{\Delta}$ is monotonic under local superchannels.*

Proof. Each item will be shown separately.

- (i) Ref. [15, Corollary 4] readily implies LOSR is a closed bounded subset of the euclidean space, and thus compact. Therefore the minimum in Eq. (6.2.1) is attained.
- (ii) First assume that Δ is convex. This means, $\Delta(\sum_i \lambda_i \mathcal{E}_i, \sum_i \lambda_i \Phi_i) \leq \sum_i \lambda_i \Delta(\mathcal{E}_i, \Phi_i)$ for non-negative λ_i adding up to one. In this case, we check that $\mathcal{N}_{\Delta}(\sum_i \lambda_i \mathcal{E}_i) \leq \sum_i \lambda_i \mathcal{N}_{\Delta}(\mathcal{E}_i)$ thus $\mathcal{N}_{\Delta}$ is also convex. Indeed, for any collection $\Phi_i \in \text{LOSR}$, also $\sum_i \lambda_i \Phi_i \in \text{LOSR}$ by convexity of LOSR.
- (iii) Next, assume Δ is non-degenerate, namely $\Delta(\mathcal{E}, \Phi) = 0$ if and only if $\mathcal{E} = \Phi$. Then, $\mathcal{N}_{\Delta}(\mathcal{E}) = 0$ if and only if $\mathcal{E} \in \text{LOSR}$ is faithful with respect to LOSR.

- (iv) Last, assume Δ is weakly contractive under superchannels, $\Delta(\Xi(\mathcal{E}), \Xi(\Phi)) \leq \Delta(\mathcal{E}, \Phi)$ for any superchannel Ξ . Assume $\Xi = \sum_{\lambda} p_{\lambda} \Xi_A^{\lambda} \otimes \Xi_B^{\lambda}$ is local. In this case, $\Xi(\Phi) \in \text{LOSR}$ for each $\Phi \in \text{LOSR}$. In particular, take $\Phi \in \text{LOSR}$ such that $\Delta(\mathcal{E}, \Phi) = \mathcal{N}_{\Delta}(\mathcal{E})$. Then,

$$\mathcal{N}_{\Delta}(\Xi(\mathcal{E})) \leq \Delta(\Xi(\mathcal{E}), \Xi(\Phi)) \leq \Delta(\mathcal{E}, \Phi) = \mathcal{N}_{\Delta}(\mathcal{E}). \quad (6.2.2)$$

Thus, $\mathcal{N}_{\Delta}$ is monotone under local superchannels.

□

It is well known that the relative entropy [269] satisfies all these properties. Moreover, any distance defined from a norm $\Delta(\mathcal{E}, \Phi) = \|\mathcal{E} - \Phi\|$ satisfies the first three properties, also under the Choi-Jamiołkowski isomorphism. Additionally, the fourth property is also satisfied when the norm is contractive under linear operations. This is the case for the trace norm and the diamond norm [28].

Examples are $\mathcal{N}_1$ defined from the trace distance $\Delta(\mathcal{E}, \Phi) = \|J^{\mathcal{E}} - J^{\Phi}\|_1$ with $\|X\|_1 = \text{tr} \sqrt{XX^{\dagger}}$, which admits semidefinite programming bounds for unitary channels as shown in Observation 8; $\mathcal{N}_{\diamond}$ from the diamond distance $\Delta(\mathcal{E}, \Phi) = \|\mathcal{E} - \Phi\|_{\diamond}$ with $\|\mathcal{E}\|_{\diamond} = \sup\{\|\mathcal{E} \otimes \text{id}(X)\|_1 : \|X\|_1 \leq 1\}$; and $\mathcal{N}_H$ from the relative entropy $\Delta(\mathcal{E}, \Phi) = H(J^{\mathcal{E}}|J^{\Phi}) = \text{tr} J^{\mathcal{E}}(\log J^{\mathcal{E}} - \log J^{\Phi})$, where $J^{\mathcal{E}}$ is the unnormalized four-partite Choi state [26, 27]. Faithful monotones can similarly be obtained from Eq. (6.2.1) using quantum Rényi divergences [258, 270, 271]. Physically, $\mathcal{N}_{\diamond}(\mathcal{E})$ corresponds to the maximal success probability of discriminating $\mathcal{E}$ with respect to the least distinguishable LOSR operation Φ [28, 272]. This can be bounded with $\mathcal{N}_1(\mathcal{E})$ through $\mathcal{N}_1(\mathcal{E}) \leq \mathcal{N}_{\diamond}(\mathcal{E}) \leq \dim A_0 B_0 \mathcal{N}_1(\mathcal{E})$. Similarly, $\mathcal{N}_H(\mathcal{E})$ measures the smallest relative entropy between the Choi matrices of $\mathcal{E}$ and some LOSR channel Φ , which is a frequent measure for their distinguishability [28].

Observation 8. *For unitary channels, bounds for the quantity $\mathcal{N}_1$ can be obtained with semidefinite programming for unitary channels.*

Proof. Recall that the Choi matrix J^U of a unitary channel has rank 1. Therefore, the Fidelity with respect to another Choi state J^{Φ} is a linear function, $F = \text{tr}(J^U J^{\Phi})$, which in particular can be optimized over $\Phi \in \text{LOSR}$ with the (strengthened) hierarchy of Appendix 6.3.2. Being $d_{A_0 B_0}$ the dimension of the input system, the Fuchs–van de Graaf inequality [273] applied to Choi matrices establishes that $\|J^U - J^{\Phi}\|_1$ is lower bounded by $d_{A_0 B_0} - \sqrt{\text{tr}(J^U J^{\Phi})}$ and upper bounded by $\sqrt{d_{A_0 B_0} - \text{tr}(J^U J^{\Phi})}$, whose optimal values can be approximated with semidefinite programming. This in turn allows to lower and upper bound $\mathcal{N}_1$, since the square root function is monotonic over its domain. □

6.2.2 Quantifying nonlocal resources in the presence of decoherence

Consider a bipartite channel $\mathcal{E}_{AB}$ which undergoes dephasing noise $\mathcal{D}(|i\rangle\langle j|) = |i\rangle\langle j|\delta_{ij}$ both in its input and output states, $\mathcal{E}^{\mathcal{D}} := \mathcal{D} \circ \mathcal{E} \circ \mathcal{D}$, where δ_{ij} is the Kronecker delta and $\circ$ denotes channel concatenation. Then, one can describe the effective action of $\mathcal{E}^{\mathcal{D}}$ by the *decoherent action* (or *classical action*) of $\mathcal{E}$, which is a classical channel acting on digits [110, 274–276] given by a bipartite stochastic matrix $S^{\mathcal{E}}$ with entries [15, 62, 63, 65–67]

$$S_{ab,xy}^{\mathcal{E}} = \langle ab | \mathcal{E}(|xy\rangle\langle xy|) | ab \rangle, \quad (6.2.3)$$

where indices (a, b, x, y) label the canonical basis of $A_1 B_1 A_0 B_0$. The concatenation of the columns of $S^{\mathcal{E}}$ coincides with the diagonal of the Choi matrix of the channel $\mathcal{E}$, which is preserved under complete decoherence. As bipartite distributions correspond with bipartite stochastic matrices, the decoherent action (5.4.8) defines probabilities $p(a, b|x, y) = S_{ab,xy}^{\mathcal{E}}$ of obtaining outputs a, b conditioned to inputs x, y through a measurement protocol performed in the computational basis [7,

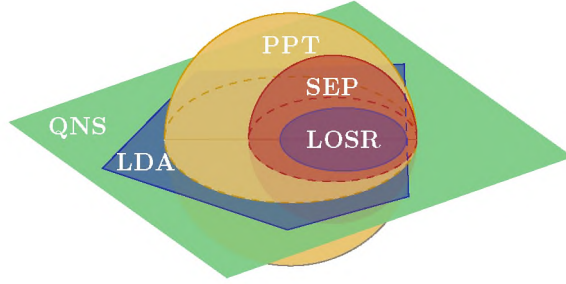


Figure 6.2.1: Subsets of bipartite channels and their cross sections. The orange convex set (PPT) contains PPT-preserving channels [114]. The green hyperplane represents the quantum nonsignaling constraints (QNS) in Eqs. (1.2.34a) and (1.2.34b), and contains the purple set of local operations with shared randomness (LOSR). The intersection between QNS and separable channels [31] (SEP) depicted in red outer-approximates LOSR through the first level of the hierarchy [245] given in Eqs. (6.3.2)-(6.3.7). This approximation can be strengthened with linear constraints imposing a local decoherent action (LDA), which give the polytope depicted in blue.

15,229]. Correspondingly, we denote as *decoherent nonlocality* of a channel $\mathcal{E}$ the nonlocal resources present in the distribution defining its decoherent action $S^\mathcal{E}$.

Nonlocality of bipartite conditional distributions (or classical channels) can be quantified in several ways [110,232,233,277,278]. For any continuous, convex, non-degenerate and contractive non-negative functional δ we introduce the measure

$$\nu_\delta(S) = \min\{\delta(S, T) : T \in \mathcal{L}\}. \quad (6.2.4)$$

Note that $\nu_\delta(S^\mathcal{E})$ quantifies the decoherent nonlocality of a bipartite channel $\mathcal{E}$. Indeed, each measure ν_δ arises from a measure $\mathcal{N}_\Delta$ by taking $\delta(S, T) = \Delta(\mathcal{E}, \Phi)$ for channels $\mathcal{E} = \mathcal{D} \circ \mathcal{E} \circ \mathcal{D}$ and $\Phi = \mathcal{D} \circ \Phi \circ \mathcal{D}$ with decoherent actions S and T . A stronger correspondence between Eqs. (6.2.1) and (6.2.4) holds:

Proposition 14. *The following relations hold between each measure $\mathcal{N}_\Delta$ of channel nonlocality and its corresponding measure ν_δ of decoherent nonlocality.*

$$(i) \quad \mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) = \nu_\delta(S^\mathcal{E})$$

$$(ii) \quad \mathcal{N}_\Delta(\mathcal{E}) \geq \nu_\delta(S^\mathcal{E}).$$

Proposition 14 allows for bounding the quantum nonlocal resources of a channel from its nonlocality resisting decoherence, which is arguably much simpler to estimate both in theory and experiments. Geometrically, item (i) holds because the dephasing channel $\mathcal{D}$ induces a cross section in the space of channels which is also a projection, and item (ii) follows from (i). A detailed proof is given below.

Proof. Consider a bipartite channel $\mathcal{E}$ with decoherent action $S^\mathcal{E}$ and a function Δ with corresponding decoherent measure δ . Let $\tilde{\Phi} \in \text{LOSR}$ minimize $\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \Phi)$ over $\Phi \in \text{LOSR}$. On the one hand, we have

$$\begin{aligned} \mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) &= \Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \tilde{\Phi}) \\ &\leq \Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \mathcal{D} \circ \tilde{\Phi} \circ \mathcal{D}), \end{aligned} \quad (6.2.5)$$

since $\mathcal{D} \circ \tilde{\Phi} \circ \mathcal{D}$ is an LOSR operation which is in general different from the minimizing LOSR operation $\tilde{\Phi}$. On the other hand, recall that by assumption Δ is contractive under superchannels Ξ , and in particular under concatenation with the dephasing $\mathcal{D}$. Therefore, we have the following

inequality,

$$\begin{aligned}\mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) &\geq \Delta(\mathcal{D}^2 \circ \mathcal{E} \circ \mathcal{D}^2, \mathcal{D} \circ \tilde{\Phi} \circ \mathcal{D}) \\ &= \Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \mathcal{D} \circ \tilde{\Phi} \circ \mathcal{D}),\end{aligned}\quad (6.2.6)$$

where the equality holds because $\mathcal{D}$ is a projection and therefore $\mathcal{D}^2 := \mathcal{D} \circ \mathcal{D} = \mathcal{D}$. Eqs. (6.2.5) and (6.2.6) together imply $\mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) = \Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \mathcal{D} \circ \tilde{\Phi} \circ \mathcal{D})$. This means that for the channel $\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}$, there exists a decohered LOSR channel attaining the minimum in Eq. (6.2.1).

Now, by definition of each δ from decoherence of Δ , we can write

$$\nu_\delta(S^\mathcal{E}) = \min_{\Phi \in \text{LOSR}} \Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}, \mathcal{D} \circ \Phi \circ \mathcal{D}), \quad (6.2.7)$$

since the whole local polytope can be obtained by decoherence of the set of LOSR operations [15]. Thus we arrive at $\mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) = \nu_\delta(S^\mathcal{E})$, which shows item (i). To see item (ii) we note that

$$\nu_\delta(S^\mathcal{E}) = \mathcal{N}_\Delta(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) \leq \mathcal{N}_\Delta(\mathcal{E}). \quad (6.2.8)$$

The equality holds due to item (i); the inequality holds because Δ is contractive under superchannels and therefore $\mathcal{N}_\Delta$ is monotonic under local superchannels. In particular, $\mathcal{D}$ is a local superchannel. $\square$

6.2.3 Quantum to classical measures: examples

Claim 1. Consider $\mathcal{E} : \mathcal{B}(A_0 \otimes B_0) \rightarrow \mathcal{B}(A_1 \otimes B_1)$ a channel and $S : \ell_1(A_0 B_0) \rightarrow \ell_1(A_1 B_1)$ its decoherent action. Then,

$$\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}\|_\diamond = \|S\|_{11}, \quad (6.2.9)$$

where $\|\mathcal{E}\|_\diamond = \sup\{\|\mathcal{E} \otimes \text{id}_{A_0 B_0}(x)\|_1 : \|x\|_1 \leq 1\}$ and $\|S\|_{11} = \sup\{\|S(a)\|_{\ell_1} : \|a\|_{\ell_1} \leq 1\}$ is the operator norm between sequence spaces ℓ_1 .

Proof. First, given $a = \sum_i a_i |i\rangle \in \ell_1(A_0 \times B_0)$ with $\|a\|_{\ell_1} \leq 1$ then $x = \sum_i a_i |i\rangle\langle i| \otimes |0\rangle\langle 0|$ has $\|x\|_1 \leq 1$. It is immediate to check that

$$\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D} \otimes \text{id}_{XY}(x)\|_1 = \|S(a)\|_{\ell_1}, \quad (6.2.10)$$

thus $\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}\|_\diamond \geq \|S\|_{11}$.

Conversely, given $x = \sum_{ij} x_{ij} |i\rangle\langle j| \otimes A_{ij}$ with $\|x\|_1 \leq 1$ we have that $\mathcal{D}(x) = \sum_i x_{ii} |i\rangle\langle i| \otimes A_{ii}$ satisfies $\|\mathcal{D}(x)\|_1 = \sum_i |x_{ii}| \|A_{ii}\|_1 \leq 1$. This is a consequence of Gelfand-Naimark theorem [183, Theorem III.4.5], since $\mathcal{D}$ is a projection. Then $a = \sum_i x_{ii} \|A_{ii}\|_1 |i\rangle$ satisfies $\|a\|_{\ell_1} \leq 1$. Furthermore, we can check

$$\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D} \otimes \text{id}_{XY}(x)\|_1 \leq \|S(a)\|_{\ell_1}, \quad (6.2.11)$$

thus, $\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}\|_\diamond \leq \|S\|_{11}$. Indeed,

$$\sum_j \left\| \sum_i x_{ii} S_{ji} A_{ii} \right\|_1 \leq \sum_j \left| \sum_i x_{ii} S_{ji} \|A_{ii}\|_1 \right|. \quad (6.2.12)$$

$\square$

Notice that from Eq. (6.2.9) and the fact that LOSR is mapped to L under decoherence it follows that $\mathcal{N}_\diamond(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}) = \nu_\diamond(S^\mathcal{E})$ as in Proposition 14. The same holds for the operator norm between spaces with the trace norm, $\|\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}\|_{11} = \|S\|_{11}$. That is, the dilation in the diamond norm makes no difference under decoherence.

Claim 2. *The trace norm of the Choi state reduces under decoherence to the element-wise one norm of the decoherent action,*

$$\|J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D})\|_1 = \|S^\mathcal{E}\|_{\ell_1}, \quad (6.2.13)$$

where $\|S\|_{\ell_1} = \sum_{ij} |S_{ij}|$.

Proof. This follows from the fact that the decoherent action $S^\mathcal{E}$ coincides with the diagonal of $J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D})$ and the off-diagonal terms are zero. $\square$

Again, from Eq. (6.2.13) and the fact that LOSR is mapped to L under decoherence it follows that $\mathcal{N}_1(J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D})) = \nu_1(S^\mathcal{E})$ as in Proposition 14.

Claim 3. *The relative entropy of the Choi state reduces under decoherence to the classical relative entropy of the decoherent action,*

$$H(J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D}), J(\mathcal{D} \circ \Phi \circ \mathcal{D})) = h(S^\mathcal{E} | S^\Phi), \quad (6.2.14)$$

where $h(S|T) = \sum_{ij} S_{ij} (\log S_{ij} - \log T_{ij})$.

Proof. This follows from the fact that the decoherent action $S^\mathcal{E}$ coincides with the diagonal of $J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D})$ and the off-diagonal terms are zero. $\square$

Again, from Eq. (6.2.14) and the fact that LOSR is mapped to L under decoherence it follows that $\mathcal{N}_H(J(\mathcal{D} \circ \mathcal{E} \circ \mathcal{D})) = \nu_H(S^\mathcal{E})$ as in Proposition 14.

6.3 Nonlocality resisting decoherence

6.3.1 A novel interpretation

The measures ν_δ of decoherent nonlocality can be seen as the amount of nonlocal resources that remain after complete decoherence occurs. In particular, these serve as a witness for nonlocality of the channel $\mathcal{E}$ in hand: if $\nu_\delta(S^\mathcal{E}) > 0$ ($S^\mathcal{E} \notin \text{L}$), then $\mathcal{N}_\Delta(\mathcal{E}) > 0$ ($\mathcal{E} \notin \text{LOSR}$) [15, 229]. Consider for example the quantum channel $\mathcal{E}_{CN}^U(\varrho) = U_{CN} \varrho U_{CN}^\dagger$ defined by the controlled-not gate, $U_{CN} = \sum_{i=0}^1 |i\rangle\langle i| \otimes \sigma_X^i$. One verifies that its decoherent action satisfies $S^{U_{CN}} = U_{CN} \odot U_{CN} = U_{CN}$, where $\odot$ is the Schur (entry-wise) product [15]. This means that after dephasing occurs, the action of the controlled-not gate in two classical bits remains untouched. In fact, $S^{U_{CN}}$ defines an extremal one-way signaling bipartite distribution and maximizes ν_1 .

This raises the question of whether the nonlocal resources resisting decoherence, which constitute the decoherent nonlocality, are related with the performance of a channel in quantum protocols which do not suffer from decoherence. The convergent hierarchy detailed below will be used to tackle this question numerically with semidefinite programming.

6.3.2 Bounding the set of LOSR operations

With the Choi-Jamiołkowski isomorphism [26, 27], deciding whether or not a given bipartite channel $\mathcal{E}_{AB}$ is LOSR is a constrained version of the separability problem: one asks that the Choi matrix defines a *separable operation* (SEP),

$$J_{A_0 B_0 A_1 B_1} = \sum_{\lambda} p_{\lambda} J_{A_0 A_1}^{\lambda} \otimes J_{B_0 B_1}^{\lambda}, \quad (6.3.1)$$

where $p_{\lambda} \geq 0$ and $J_{A_0 A_1}^{\lambda}, J_{B_0 B_1}^{\lambda} \geq 0$ [83, 84], and the additional linear constraints $\text{tr}_{A_1}(J_{A_0 A_1}^{\lambda}) = \text{id}_{A_0}$ and $\text{tr}_{B_1}(J_{B_0 B_1}^{\lambda}) = \text{id}_{B_0}$ to ensure that each factor defines a completely positive and trace preserving (CPTP) map. A solution [279] can be found by applying the hierarchy of [245], which consists of

strengthening the standard Doherty-Parrilo-Spedalieri (DPS) hierarchy [83, 84] with further linear constraints at each level n :

$$\text{find } X \geq 0 \quad (6.3.2)$$

$$\text{tr}_{A_1 B_1^0 B^1 \dots B^n}(X) = \text{id}_{A_0 B_0} \quad (6.3.3)$$

$$V_{B^i, B^n} X V_{B^i, B^n}^\dagger = X \quad \forall i \quad (6.3.4)$$

$$X^{T_{B^n}} \geq 0 \quad (6.3.5)$$

$$\text{tr}_{B_1^n}(X) = \text{tr}_{B_0^n B_1^n} X \otimes \text{id}_{B_0^n} \quad (6.3.6)$$

$$\text{tr}_{A_0}(X) = \text{id}_A \otimes \text{tr}_{A_0 A_1}(X) \quad (6.3.7)$$

where X acts on an extended Hilbert space $\mathcal{H}_{A_0 B_0 A_1 B_1} \otimes (\mathcal{H}_{B_0 B_1})^{\otimes n-1}$, we denote $A = A_0 A_1$ and $B^{(0)} = B_0 B_1$, $X_{A_0 B_0 A_1 B_1} = \text{tr}_{B^1 \dots B^n}(X)$, $V_{B^{(i)}, B^{(n)}}$ is the SWAP operator between the i 'th and n 'th party and $T_{B^{(n)}}$ is the partial transpose on the party $B^{(n)}$. Here Eqs. (6.3.2) and (6.3.3) impose that the marginal $X_{A_0 B_0 A_1 B_1}$ is the Choi matrix of a CPTP map; Eqs. (6.3.4) and (6.3.5) converge to the condition that $X_{A_0 B_0 A_1 B_1}$ decomposes as in Eq. (6.3.1) [83, 84], which means that $X_{A_0 B_0 A_1 B_1}$ defines a separable operation, $X_{A_0 B_0 A_1 B_1} = \sum_\lambda X_{A_0 A_1}^\lambda \otimes X_{B_0 B_1}^\lambda$; Eqs. (6.3.6) and (6.3.7) impose that $X_{A_0 B_0 A_1 B_1}$ is the Choi matrix of a qNS channel, and converge at some n to the condition that the marginals of each term satisfy $X_{A_0}^\lambda = \text{id}$ and $X_{B_0}^\lambda = \text{id}$ [245], which means that they define trace preserving channels and therefore $X_{A_0 B_0 A_1 B_1}$ defines an LOSR channel.

Although the hierarchy presented above is complete, only the first levels are available in practice and further linear constraints are desired to solve realistic problems. Here we will simply add the following linear constraint at each level,

$$\text{diag}(J_{A_0 B_0 A_1 B_1}) \in \text{L}, \quad (6.3.8)$$

which requires that the probability distribution given by the decoherent action lies in the local polytope (Fig. 6.2.1) and thus enforces $\nu_\delta = 0$. Straightforward examples of the strengthening of Eqs. (6.3.2)-(6.3.7) through Eq. (6.3.8) are diagonal Choi matrices given by $J^\mathcal{E} = \sum_{xyab} p(ab|xy) |xyab\rangle \langle xyab|$ where $p = S^\mathcal{E}$ is nonsignaling, as explained below.

Consider a diagonal Choi matrix given by $J^\mathcal{E} = \sum_{xyab} p(ab|xy) |xyab\rangle \langle xyab|$ with $p = S^\mathcal{E} \in \text{NS}$, e.g. given by the Popescu-Rohrlich (PR-) box

$$S^\mathcal{E} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad (6.3.9)$$

Such Choi matrix is in the set SEP because it is diagonal in product basis, and thus is not detected by the Doherty-Parrilo-Spedalieri (DPS) hierarchy (Eqs. (6.3.2)-(6.3.5)). At the first level, equations (6.3.6) and (6.3.7) are also satisfied because a classical map given by a stochastic matrix in the nonsignaling polytope defines a quantum nonsignaling channel. However, such an example does not satisfy Eq. (6.3.8).

6.4 The role of decoherent nonlocality in performing quantum protocols

6.4.1 Bell functionals on quantum channels

We will introduce functionals on Choi matrices that reach different maximum values for different subsets of bipartite channels. This generalizes through Proposition 14 the analogous property of ν_δ for bipartite distributions. Given a linear functional Ω on the space of Choi matrices for

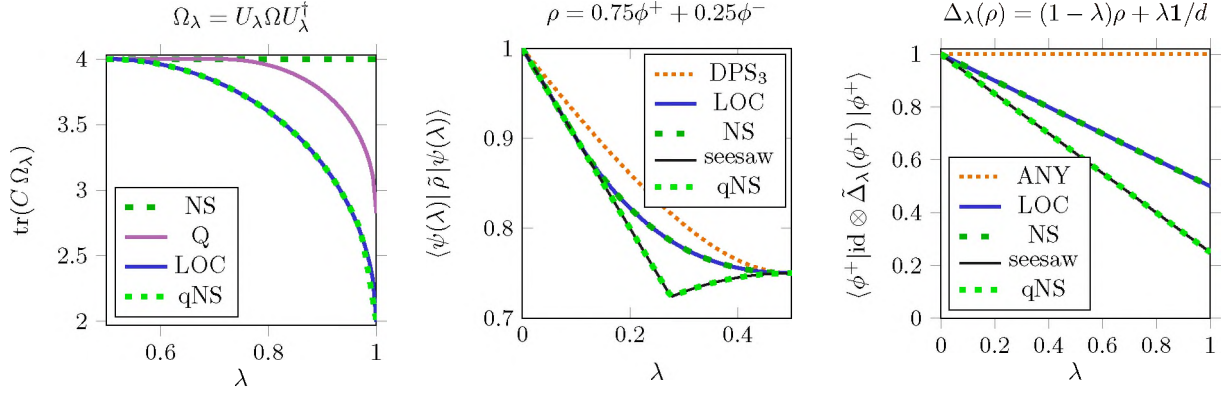


Figure 6.3.1: Optimal performance of bipartite channels $\mathcal{E}$ in restricted subsets at quantum protocols with parameter μ . Different bounds are obtained when $S^\mathcal{E}$ defines local (LDA), quantum (NPA₃), nonsignaling (NSDA), and general distributions, where NPA₃ denotes third-level approximation with [249]. Bounds for LOSR are obtained when those obtained with quantum-nonsignaling channels (QNS) and with a seesaw algorithm coincide. We depict the maximum value of: (a) The Bell functional Ω_μ in Eq. (6.4.1). (b) The fidelity of $\tilde{\rho} = \mathcal{E}(\rho)$ to a target pure state $|\psi_\mu\rangle$, where DPS₃ approximates SEP at the third level of [83, 84]. (c) The fidelity of the depolarizing channel Λ_μ assisted according to Eq. (6.4.3) and Fig. 6.4.1. General superchannels (CPTP) [28, 64] allow for unit fidelity through $\Xi[\Lambda_\mu] = \text{id}$.

channels $\mathcal{E}_{AB}$ [15], consider the maximum value that can be obtained with quantum channels whose decoherent action belongs to a preferred subset $\mathcal{L}$,

$$\gamma_{\mathcal{L}}(\Omega) := \max\{\text{tr}(J^\mathcal{E}\Omega) : S^\mathcal{E} \in \mathcal{L}\}. \quad (6.4.1)$$

Let us fix $\Omega_\mu = U_\mu \Omega U_\mu^\dagger$, which is a coherified version of the diagonal functional $\Omega = \sum_{x,y,a,b=0}^1 (-1)^{a+b} (-1)^{xy} |xy\rangle\langle xy|$ used in [15] to evaluate the Clauser-Horne-Shimony-Holt inequality [280] on the decoherent action. Nonzero off-diagonal terms are given by the unitary $U_\mu = (|\phi_\mu^+\rangle\langle 00| + |\psi_\mu^+\rangle\langle 01| + |\psi_\mu^-\rangle\langle 10| + |\phi_\mu^-\rangle\langle 11|)_{A_0 B_0} \otimes \mathbb{1}_{A_1 B_1}$ where $|\psi_\mu^\pm\rangle = X^\pm(\sqrt{\mu}|01\rangle \pm \sqrt{1-\mu}|10\rangle)\sqrt{2}$ with $X^\pm = (\sigma_X \otimes \sigma_X)^{(\pm 1 - 1)/2}$. The states $|\psi_\mu^\pm\rangle$ and $|\phi_\mu^\pm\rangle$ are entangled for $\mu \neq 0, 1$ and form an orthonormal basis for all μ , e.g. the computational basis for $\mu = 1$ and the Bell basis for $\mu = 1/2$. Fig. 6.3.1 (a) shows that the optimal value of Eq. (6.4.1) depends on the nonlocality of the decoherent action throughout the whole range $\mu \in (1/2, 1]$: a continuous gap exists between the sets of channels whose decoherent action is in the set of local, quantum (an outer approximation [249]), and nonsignaling distributions.

6.4.2 State interconversion

Consider the fidelity of converting a Bell state which suffered from the unitary noise $\text{id} \otimes \sigma_Z$ with probability $1-p$, $\rho = p\phi^+ + (1-p)\phi^-$, into a pure state with entanglement entropy $E = \mu \log_2 \mu^{-1} + (1-\mu) \log_2 (1-\mu)^{-1}$, namely $|\psi_\mu\rangle = \sqrt{\mu}|00\rangle + \sqrt{1-\mu}|11\rangle$ up to local unitaries. The figure of merit is the fidelity between the transformed state by a channel $\mathcal{E}$ and the target state $|\psi_\mu\rangle$,

$$\mathcal{F} = \langle \psi_\mu | \mathcal{E}(\rho) | \psi_\mu \rangle. \quad (6.4.2)$$

Standard physical scenarios require that transformations can be done through quantum operations with a limited amount of resources the nonlocality present in the decoherent action of the channel $\mathcal{E}$ (Eq. (5.4.8)) can increase the fidelity of the transformation in Eq. (6.4.2).

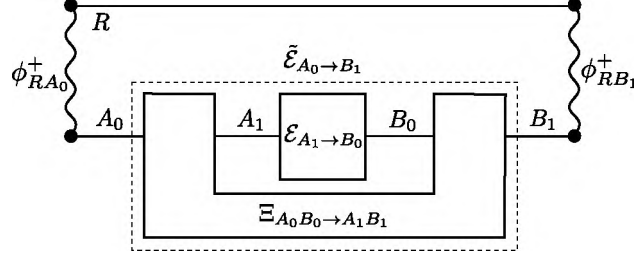


Figure 6.4.1: Assisted communication. In a standard communication protocol, one part A_1 of a Bell state ϕ^+ is sent to another part B_0 through a noisy channel $\mathcal{E}$ with input A_1 and output B_0 . In an assisted communication protocol, the noisy channel is corrected to $\tilde{\mathcal{E}}_{A_0 \rightarrow B_1} = \Xi_{A_0 B_0 \rightarrow A_1 B_1}(\mathcal{E}_{A_1 \rightarrow B_0})$ (dashed outline) by a superchannel Ξ . The fidelity of the assisted channel is then the fidelity between the input Bell state and its output image, $\mathcal{F} = \langle \phi^+ | \tilde{\mathcal{E}}(\phi^+) | \phi^+ \rangle$.

6.4.3 Quantum codes

Consider the task of sending information from Alice to Bob through a noisy channel $\mathcal{E}$. The assisted fidelity of sending one part of a maximally entangled state,

$$\mathcal{F} = \text{tr} \left(\phi^+_{RB_1} \cdot (\text{id}_R \otimes \tilde{\mathcal{E}}_{A_0 \rightarrow B_1})(\phi^+_{RA_0}) \right), \quad (6.4.3)$$

quantifies how well information is transmitted through $\tilde{\mathcal{E}}_{A_0 \rightarrow B_1} := \Xi(\mathcal{E}_{A_1 \rightarrow B_0})$, which is the transformed version of $\mathcal{E}_{A_1 \rightarrow B_0}$ by the superchannel Ξ (Fig. 6.4.1) [114, 163, 281–284]. The superchannel Ξ is equivalent to a bipartite channel $\Xi_{A_0 B_0 \rightarrow A_1 B_1}$ whose Choi matrix satisfies $J \geq 0$ with the marginals $J_{A_0 B_0} = \text{id}_{A_0 B_0}$ and the one-way nonsignaling constraint $J_{A_0 A_1 B_0} = J_{A_0 A_1} \otimes \text{id}_{B_0} / d_{B_0}$ [28, 112]. These constraints alone are compatible with $\mathcal{F} = 1$ since they admit a contractive superchannel Ξ such that $\tilde{\mathcal{E}} = \text{id}$ is the noiseless identity channel. However, physically realistic scenarios can be modelled by imposing extra constraints to Ξ , leading to nontrivial bounds [114, 163, 279, 281, 285]. Bounds for the depolarizing channel Λ_μ with mixing parameter μ obtained through the strengthened hierarchy of Section 6.3.2 are depicted in Fig 6.3.1 (c). The results of the three protocols depicted in Fig 6.3.1 imply the following.

Proposition 15. *The nonlocal resources of a quantum channel that remain available when coherences are lost (quantified by ν_δ) are required for optimal performance of coherent quantum protocols.*

6.5 Quantum simulation of stochastic matrices

Stochastic matrices, also called classical channels [110, 184, 274–276], model the evolution of discrete probabilistic systems and are the building blocks of certain stochastic processes like discrete Markov chains [58]. Although these are classical, it has been shown that the quantum simulation of stochastic processes provides advantage in certain resources such as memory and space-time cost with respect to their classical simulation [66, 67]. A common tool for this purpose is the decoherent action, using the fact that complete decoherence preserves the action of a channel on classical systems.

Here we will follow a similar approach: one verifies that two parties that share a common classical source of randomness but cannot communicate, can only simulate bipartite stochastic matrices of the form

$$S_{AB} = \sum_{\lambda} p_{\lambda} T_A^{\lambda} \otimes R_B^{\lambda} \quad (6.5.1)$$

where T_A^{λ} and R_B^{λ} are stochastic matrices. This is because only local bipartite distributions can be obtained in this case [286]. In contrast, larger sets of bipartite stochastic matrices can be simulated using shared quantum resources undergoing noise:

Proposition 16. *There are stochastic processes that require communication to be simulated classically, but can be simulated quantumly without communication through a channel undergoing decoherence.*

For the purpose of quantum advantage, one can use quantum channels which are implemented by using a shared entangled state and applying local operations. The following shows how stochastic matrices corresponding to nonlocal distributions can be obtained when the channel in hand undergoes dephasing noise, in which case measurements can only be performed in the computational basis.

Proof. Consider an LOSE operation as in Eq. (1.2.33). Let the Kraus representation of the channels $\mathcal{E}_{AR}$ and $\mathcal{E}_{BS}$, which act through $\mathcal{E}_{AR}(\cdot) = \sum_{i,j} K_{i,j}^{\mathcal{E}_{AR}} \cdot K_{i,j}^{\mathcal{E}_{AR}\dagger}$ and $\mathcal{E}_{BS}(\cdot) = \sum_{k,l} Q_{k,l}^{\mathcal{E}_{BS}} \cdot Q_{k,l}^{\mathcal{E}_{BS}\dagger}$, be $K_{i,j}^{\mathcal{E}_{AR}} = |i\rangle\langle j|_A \otimes L^{i|j}_R$ and $Q_{k,l}^{\mathcal{E}_{BS}} = |k\rangle\langle l|_B \otimes T^{k|l}_S$, where $M^{i|j} := L^{i|j\dagger} L^{i|j}$ and $N^{k|l} := T^{k|l\dagger} T^{k|l}$ are conditional measurement operators. One verifies that the decoherent action $S^\mathcal{E}$ of $\mathcal{E}_{AB}$ reads entrywise $S_{ab,xy}^\mathcal{E} = \text{tr}(M^{a|x} \otimes N^{b|y} \tau_{RS})$, obtaining a similar result as if Alice and Bob were able to measure in coherent bases to implement a standard Bell game [52]. Therefore, Alice and Bob can simulate a classical channel beyond the set of matrices given by Eq. (6.5.1) without communication if they have access to quantum resources. $\square$

Proposition 16 implies that quantum advantage in communication remains accessible when coherences are lost in the inputs and outputs of the channel in use. This allows us to use the techniques of [66], where quantum advantage was shown in noisy channels, in the spirit of [287], where entanglement reduces communication needed to implement a classical channel.

6.6 Conclusions

We introduce several measures of nonlocality for bipartite channels with desirable properties, and identify their component remaining after decoherence. Despite accounting for classical resources at the level of channels, we show that nonlocality resisting decoherence is required for certain quantum protocols: maximizing Bell functionals for channels, state interconversion fidelity, and assisted codes fidelity for noisy channel communication. This analysis also demonstrates a gap between different subsets of bipartite quantum channels. These results provide significant insight into the nonlocal properties resisting dephasing noise and their role in quantum implementations: Bell nonlocality of correlations is understood as the classical component of channel nonlocality, that can be implemented with fully classical devices.

On a more conceptual level, nonlocality of channels (quantum or classical) accounts for the resources needed for their implementation in distant labs. These resources include classical communication or shared entanglement. Combining these two, any quantum operation can be locally implemented, while classical communication is enough to locally realize any classical operation. Our analysis suggests a natural way to simulate bipartite stochastic matrices, from bipartite quantum channels which undergo complete decoherence. We show that this quantum simulation method has a communication advantage with respect to classical simulation. Our results open an avenue to understanding the tradeoff between entanglement and communication costs of implementing a bipartite quantum operation.

Outlook and open problems

Quantum information has been developed as a major mathematical ground to understand quantum physics occurring at small scales, in the non-relativistic limit. Currently, it constitutes the main tool employed to establish a solid basis for the development of quantum technologies: it allows to design of quantum algorithms overperforming their classical counterparts, develop methods to certify the presence of intrinsically quantum properties, and understand the potential and limitations of quantum devices, among other applications. Therefore, quantum information is currently regarded as a promising starting point of a potentially upcoming technological revolution, in which current devices might evolve to improved versions obeying the laws of quantum physics in their functioning. This is expected to allow to find the solution to currently challenging problems, which are inaccessible with standard classical computing tools.

In turn, quantum information provides with a powerful framework to both understand the underlying principles of physics, and enable a technological step growth towards a new era. However, its powerful nature is inherent to a complex structure that challenges its implementation. The vector space describing a state space of physical systems grows exponentially with the system size, and precise strategies are required for a proper use of quantum theory. In particular, identifying symmetries in a problem allows for a simpler formulation; and a geometric description allows to identify the relevant structures in specific problems, such as the distinction between classical and quantum structures. This thesis is focused on symmetric and geometric tools that allow for a better understanding of selected questions concerning the inherently convex structure of quantum theory. These tackle natural questions directly built upon the basic Postulates of quantum theory, presented at the beginning of Chapter 1 and further developed throughout the chapter.

Chapter 2 is focused of the problem of *entanglement detection*: Given a quantum state describing a physical system, one asks whether it is separable or entangled. Answering this question is both challenging and essential for a proper implementation of quantum devices, and thus constitutes an active topic of research. A standard tool to detect entanglement are entanglement witnesses, which are observables whose expectation value can separate a state from the convex hull of product states. In Chapter 2, standard witnesses are used in a novel manner. The method is based on identical copies of the (known) state in hand, which results in a nonlinear protocols that make use of the linear toolkit of witnesses. These are tested in toy-model multipartite states and symmetric states, whose evaluation becomes accessible for nontrivial system sizes. The use of symmetric witnesses allows for entanglement tests invariant under local unitaries, which can moreover be implemented in the laboratory through randomized measurements. This chapter provides thus a novel method to detect entanglement. Yet, its power is at the moment not completely explored. Further research can be done in understanding the set of states that can be detected for a fixed number of copies k using witnesses arising from a complete set of immanant inequalities.

Chapter 3 provides a detailed analysis of a practical method to solve convex optimization problems in the presence of symmetries of the form $U^{\otimes n}$. First, a general semidefinite program formulation for such problems is discussed, starting from the Hilbert space decomposition through

Schyr-Weyl duality. Then we analyze in depth the distinct methods of exploiting the symmetries displayed by the physical system under consideration. A detailed discussion on the optimal methods is given, based on a numerical implementation through symbolic algebra and a numerical representation theory toolkit. A comparison between the size of the positivity constraint, the number of variables adapted to a specific basis spanning the reduced space, and a specific formulation exploiting the $U^{\otimes n}$ -type symmetry are given. The application of these tools is discussed in particular to solve the quantum Max-Cut problem for an arbitrary Hamiltonian graph, the marginal problem for Werner states, and to find the ground state of an infinite spin lattice.

Chapter 4 is concerned with the geometrical structure of positive-operator-valued measurements (POVM), also known as generalized quantum measurements. While numerous studies have been done to understand their static structure, our work opens a path to understand their dynamical properties. Mathematically, the starting point is the standard evolution of their simple classical analogues: discrete probability distributions evolving by stochastic matrix multiplication, which generalizes to the quantum theory as the endomorphism in the space of quantum measurements generated by a Cartesian product of elements within. Physically, this corresponds to understanding the set of measurements that can be obtained by a sequence of previous measurements. Naturally, the answer relies on joint measurability, and standard rules of probability transformation apply to the quantum setup: we prove a quantum analog of the Ostrowski characterization of the Birkhoff polytope, by introducing a novel notion of majorization between quantum measurements through their effects. While developing this project, further questions arose concerning the set of blockwise bistochastic matrices, and the answers we found lead to a follow-up project. In particular, the geometry of such set for small system sizes is currently under investigation together with my collaborators.

Chapter 5 tackles the following problem: Given a bipartite quantum operation, can we certify nonlocal properties with accessible measurement protocols? While this question has started to be a current topic of research recently, a realistic design of quantum devices requires considering the case of noise affecting quantum operations. We provide novel criteria for the certification of nonlocal resources, and focus on the effect of noise produced by decoherence. First we discuss extensions of existing certification methods. Then we analyze different models of decoherence and study their effects in terms of gate nonlocality and entanglement generation. We finally analyze the case of complete decoherence, in which the complex structure of the quantum state space and its convex hull is reduced to a probability simplex. The analysis provided sheds light into the quantum-to-classical transition induced by decoherence, and into the certification methods that can be implemented even if all coherences are lost.

Chapter 6 follows up and expands on the last discussion of Chapter 5 concerning complete decoherence occurring to a quantum operation. The transition from quantum operations in a vector state space to classical stochastic operations in a probability simplex leads to a distinction between two types of nonlocal resources of quantum operations: quantum and classical. Both the overall nonlocal resources and their fraction withstanding decoherence are quantified. Despite their classical nature, the latter are shown to play a crucial role in optimal performance of certain standard quantum protocols, even unaffected by decoherence. The resource distinction presented in this chapter provides a starting point to a new research avenue towards understanding the distinction and the tradeoff between the two resources required to implement any nonlocal quantum operation: shared entanglement, and classical communication.

Open problems

The research work gathered in this thesis is focused on providing a new perspective and new connections in quantum information theory. This new angle allows to solve some existing problems, like the generalized Ostrowski's characterization of the Birkhoff polytope applied to quantum measurements [186]. But it also introduces new techniques, like the use of immanent inequalities to detect entanglement, leading to novel research avenues. Naturally, this new insight leaves central

questions in the fundamental and applied pillars of quantum information theory:

1. Entanglement is a physical resource stored in composite quantum systems. By definition, it is invariant under local unitary operations. However, linear methods for entanglement detection of the form $\text{tr}(W\rho)$ are in general sensitive to rotations of the local measurement apparatuses, and thus cannot constitute a systematic general method. Following Chapter 2 we may consider expressions of the form $\theta_{\mathcal{W}}(\rho) := \text{tr}(\mathcal{W}\rho^{\otimes k})$, which display both the desired local unitary invariance and translation invariance. Thus one can ansatz $\mathcal{W} = \sum_{\lambda \vdash k} \omega_{\lambda} P_{\lambda}$ with $\omega_{\lambda} \in \mathbb{C}$, where P_{λ} is a Young projector. While this case contains the approach of Chapter 2 through Immanant inequalities, a more general question remains open: *Can one find a complete set of local unitary invariant expressions $\theta_{\mathcal{W}}(\rho)$ detecting all classes of entanglement?*
2. The implementation of a shared quantum operation requires shared resources. In particular, two types of nonlocal resources suffice: shared entanglement, and classical communication. Together, these allow for example to perform gate teleportation. However, through the canonical descriptions of quantum channels used at the moment, both the separation between these resources and their exact use in a channel implementation are unclear. In chapter 6 we distinguished them by introducing a measure of total nonlocality and a measure of decoherent nonlocality, whose difference quantifies the strictly quantum shared resources. While this analysis constitutes a promising starting point, the following problem is left open: *How can one systematically describe the implementation of arbitrary bipartite channels, using the distinction between quantum and classical shared resources?*

Bibliography

- [1] P. W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM Journal on Computing*, vol. 26, no. 5, 1997. [Online]. Available: <http://dx.doi.org/10.1137/S0097539795293172>
- [2] C. H. Bennett, G. Brassard, and J.-M. Robert, “Privacy amplification by public discussion,” *SIAM J. Comput.*, vol. 17, no. 2, 1988. [Online]. Available: <https://doi.org/10.1137/0217014>
- [3] B. Mielnik, “Generalized quantum mechanics,” *Commun. Math. Phys.*, vol. 37, 1974. [Online]. Available: <https://link.springer.com/article/10.1007/BF01646346>
- [4] O. Gühne and G. Tóth, “Entanglement detection,” *Physics Reports*, vol. 474, no. 1–6, 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.physrep.2009.02.004>
- [5] F. Huber and N. Wyderka, “Refuting spectral compatibility of quantum marginals,” *arXiv: 2211.06349*, 2023. [Online]. Available: <https://arxiv.org/abs/2211.06349>
- [6] T. Heinosaari and M. M. Wolf, “Nondisturbing quantum measurements,” *Journal of mathematical physics*, vol. 51, no. 9, 2010. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/51/9/092201/896891>
- [7] D. Schmid, D. Rosset, and F. Buscemi, “The type-independent resource theory of local operations and shared randomness,” *Quantum*, vol. 4, 2020. [Online]. Available: <https://doi.org/10.22331/q-2020-04-30-262>
- [8] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels,” *Phys. Rev. Lett.*, vol. 70, 1993. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.70.1895>
- [9] S. Lloyd, “Capacity of the noisy quantum channel,” *Phys. Rev. A*, vol. 55, 1997. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.55.1613>
- [10] P. W. Anderson, “Limits on the energy of the antiferromagnetic ground state,” *Phys. Rev.*, vol. 83, 1951. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.83.1260>
- [11] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin, “Invariant semidefinite programs,” *Handbook on semidefinite, conic and polynomial optimization*, 2012. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4614-0769-0_9
- [12] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: an Introduction to Quantum Entanglement*. Cambridge University Press, 2006.
- [13] A. Rico and F. Huber, “Entanglement detection with trace polynomials,” *Phys. Rev. Lett.*, vol. 132, 2024. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.132.070202>
- [14] A. Rico and K. Życzkowski, “Discrete dynamics in the set of quantum measurements,” *arXiv: 2308.05835*, 2023. [Online]. Available: <https://arxiv.org/abs/2308.05835>

- [15] A. Rico, M. B. Morán, F. Shahbeigi, and K. Życzkowski, “Certifying nonlocal properties of noisy quantum operations,” *arXiv: 2407.02570*, 2024. [Online]. Available: <https://arxiv.org/abs/2407.02570>
- [16] A. Rico, M. B. Morán, F. Shahbeigi, and K. Życzkowski, “Channel nonlocality under decoherence,” *arXiv: 2408.10317*, 2024. [Online]. Available: <https://arxiv.org/abs/2408.10317>
- [17] T. Young, “I. The Bakerian Lecture. Experiments and calculations relative to physical optics,” *Philosophical transactions of the Royal Society of London*, no. 94, 1804. [Online]. Available: <https://royalsocietypublishing.org/doi/10.1098/rstl.1804.0001>
- [18] W. Heisenberg, “Über den Anschaulichen Inhalt der Quantentheoretischen Kinematik und Mechanik,” *Zeitschrift für Physik*, vol. 43, no. 3, 1927. [Online]. Available: <https://link.springer.com/article/10.1007/BF01397280>
- [19] B. S. Dewitt and N. Graham, *The many-worlds interpretation of quantum mechanics*. Princeton University Press, 2015, vol. 61.
- [20] D. Bohm, “A suggested interpretation of the quantum theory in terms of “hidden” variables. i,” *Phys. Rev.*, vol. 85, 1952. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.85.166>
- [21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. [Online]. Available: <https://doi.org/10.1017/CBO9780511976667>
- [22] C. E. Shannon, “A mathematical theory of communication,” *The Bell system technical journal*, vol. 27, no. 3, 1948. [Online]. Available: <https://ieeexplore.ieee.org/document/6773024>
- [23] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The quantum theory of measurement*. Springer, 1996.
- [24] J. A. Wheeler and W. H. Zurek, *Quantum theory and measurement*. Princeton University Press, 2014, vol. 81.
- [25] W. F. Stinespring, “Positive functions on C*-algebras,” *Proceedings of the American Mathematical Society*, vol. 6, no. 2, 1955. [Online]. Available: <http://www.jstor.org/stable/2032342>
- [26] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.*, vol. 10, no. 3, 1975. [Online]. Available: [https://doi.org/10.1016/0024-3795\(75\)90075-0](https://doi.org/10.1016/0024-3795(75)90075-0)
- [27] A. Jamiolkowski, “Linear transformations which preserve trace and positive semidefiniteness of operators,” *Rep. Math. Phys.*, vol. 3, no. 4, 1972. [Online]. Available: [https://doi.org/10.1016/0034-4877\(72\)90011-0](https://doi.org/10.1016/0034-4877(72)90011-0)
- [28] G. Gour, “Comparison of quantum channels by superchannels,” *IEEE Trans. Inf. Theory*, vol. 65, no. 9, 2019. [Online]. Available: <https://doi.org/10.1109/TIT.2019.2907989>
- [29] J. Von Neumann, *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, 1932, vol. 38.
- [30] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?” *Phys. Rev.*, vol. 47, 1935. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.47.777>

- [31] 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*, vol. 328, 2014. [Online]. Available: <https://link.springer.com/article/10.1007/s00220-014-1953-9>
- [32] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.*, vol. 81, 2009. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.81.865>
- [33] W. Dür, G. Vidal, and J. I. Cirac, “Three qubits can be entangled in two inequivalent ways,” *Phys. Rev. A*, vol. 62, 2000. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.62.062314>
- [34] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, “Experimental quantum teleportation,” *Nature*, vol. 390, no. 6660, 1997. [Online]. Available: <https://www.nature.com/articles/37539>
- [35] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, ““event-ready-detectors” bell experiment via entanglement swapping,” *Phys. Rev. Lett.*, vol. 71, 1993. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.71.4287>
- [36] J.-W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, “Experimental entanglement swapping: Entangling photons that never interacted,” *Phys. Rev. Lett.*, vol. 80, 1998. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.80.3891>
- [37] R. Raussendorf and H. J. Briegel, “A one-way quantum computer,” *Phys. Rev. Lett.*, vol. 86, 2001. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.86.5188>
- [38] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger, “Experimental one-way quantum computing,” *Nature*, vol. 434, no. 7030, 2005. [Online]. Available: <https://www.nature.com/articles/nature03347>
- [39] L. Gurvits, “Classical complexity and quantum entanglement,” *Journal of Computer and System Sciences*, vol. 69, no. 3, 2004, special Issue on STOC 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022000004000893>
- [40] S. Gharibian, “Strong np-hardness of the quantum separability problem,” *Quantum Info. Comput.*, vol. 10, no. 3, 2010. [Online]. Available: <https://dl.acm.org/doi/10.5555/2011350.2011361>
- [41] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps,” *Physics Letters A*, vol. 283, no. 1-2, 2001. [Online]. Available: [http://dx.doi.org/10.1016/S0375-9601\(01\)00142-6](http://dx.doi.org/10.1016/S0375-9601(01)00142-6)
- [42] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, “Mixed-state entanglement from local randomized measurements,” *Physical Review Letters*, vol. 125, 2020. [Online]. Available: <https://link.aps.org/%doi/10.1103/PhysRevLett.125.200501>
- [43] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, and B. Kraus, “Symmetry-resolved entanglement detection using partial transpose moments,” *npj Quantum Information*, vol. 7, no. 1, 2021. [Online]. Available: <https://doi.org/10.1038%2Fs41534-021-00487-y>
- [44] M. Horodecki, P. Horodecki, and R. Horodecki, “Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature?” *Physical Review Letters*, vol. 80, no. 24, 1998. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.80.5239>

- [45] M. Horodecki and P. Horodecki, “Reduction criterion of separability and limits for a class of distillation protocols,” *Phys. Rev. A*, vol. 59, 1999. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.59.4206>
- [46] H.-P. Breuer, “Optimal entanglement criterion for mixed quantum states,” *Phys. Rev. Lett.*, vol. 97, 2006. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.97.080501>
- [47] W. Hall, “A new criterion for indecomposability of positive maps,” *Journal of Physics A: Mathematical and General*, vol. 39, no. 45, 2006. [Online]. Available: <https://iopscience.iop.org/article/10.1088/0305-4470/39/45/020>
- [48] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions,” *Phys. Lett. A*, vol. 223, no. 1, 1996. [Online]. Available: [https://doi.org/10.1016/S0375-9601\(96\)00706-2](https://doi.org/10.1016/S0375-9601(96)00706-2)
- [49] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, “Optimization of entanglement witnesses,” *Phys. Rev. A*, vol. 62, 2000. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.62.052310>
- [50] P. Hyllus, O. Gühne, D. Bruß, and M. Lewenstein, “Relations between entanglement witnesses and Bell inequalities,” *Phys. Rev. A*, vol. 72, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.72.012321>
- [51] J. S. Bell, “On the Einstein-Podolsky-Rosen paradox,” *Physics Physique Fizika*, vol. 1, no. 3, 1964. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195>
- [52] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, “Bell nonlocality,” *Rev. Mod. Phys.*, vol. 86, 2014. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.86.419>
- [53] J. S. Bell, “On the problem of hidden variables in quantum mechanics,” *Rev. Modern Phys.*, vol. 38, no. 3, 1966. [Online]. Available: <https://doi.org/10.1103/RevModPhys.38.447>
- [54] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” *Phys. Rev. A*, vol. 40, no. 8, 1989. [Online]. Available: <https://doi.org/10.1103/PhysRevA.40.4277>
- [55] T. Vértesi and N. Brunner, “Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement,” *Nat. Commun.*, vol. 5, no. 1, 2014. [Online]. Available: <http://dx.doi.org/10.1038/ncomms6297>
- [56] F. Buscemi, “All entangled quantum states are nonlocal,” *Phys. Rev. Lett.*, vol. 108, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.108.200401>
- [57] A. Winter and D. Yang, “Operational resource theory of coherence,” *Phys. Rev. Lett.*, vol. 116, 2016. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.116.120404>
- [58] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, 2007. [Online]. Available: <https://doi.org/10.1093/acprof:oso/9780199213900.001.0001>
- [59] S.-H. Kye, “Positive linear maps between matrix algebras which fix diagonals,” *Linear Algebra Appl.*, vol. 216, 1995. [Online]. Available: [https://doi.org/10.1016/0024-3795\(93\)00140-U](https://doi.org/10.1016/0024-3795(93)00140-U)
- [60] C.-K. Li and H. J. Woerdeman, “Special classes of positive and completely positive maps,” *Linear Algebra Appl.*, vol. 255, no. 1, 1997. [Online]. Available: [https://doi.org/10.1016/S0024-3795\(96\)00776-8](https://doi.org/10.1016/S0024-3795(96)00776-8)

- [61] J. Levick, D. W. Kribs, and R. Pereira, “Quantum privacy and schur product channels,” *Rep. Math. Phys.*, vol. 80, no. 3, 2017. [Online]. Available: [https://doi.org/10.1016/S0034-4877\(18\)30005-3](https://doi.org/10.1016/S0034-4877(18)30005-3)
- [62] Z. Puchala, K. Korzekwa, R. Salazar, P. Horodecki, and K. Życzkowski, “Dephasing superchannels,” *Phys. Rev. A*, vol. 104, 2021. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.104.052611>
- [63] K. Korzekwa, S. Czachórski, Z. Puchala, and K. Życzkowski, “Coherifying quantum channels,” *New J. Phys.*, vol. 20, no. 4, 2018. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1367-2630/aaaff3>
- [64] G. Chiribella, G. M. D'Ariano, and P. Perinotti, “Transforming quantum operations: Quantum supermaps,” *EPL*, vol. 83, no. 3, 2008. [Online]. Available: <https://doi.org/10.1209%2F0295-5075%2F83%2F30004>
- [65] F. Shahbeigi, D. Amaro-Alcalá, Z. Puchala, and K. Życzkowski, “Log-convex set of Lindblad semigroups acting on N-level system,” *J. Math. Phys.*, vol. 62, no. 7, 2021. [Online]. Available: <http://dx.doi.org/10.1063/5.0009745>
- [66] K. Korzekwa and M. Lostaglio, “Quantum advantage in simulating stochastic processes,” *Phys. Rev. X*, vol. 11, 2021. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevX.11.021019>
- [67] F. Shahbeigi, C. T. Chubb, R. Kukulski, Ł. Paweła, and K. Korzekwa, “Quantum-embeddable stochastic matrices,” *Quantum*, vol. 8, 2024. [Online]. Available: <https://doi.org/10.22331/q-2024-07-10-1404>
- [68] K. T. Goh, J. Kaniewski, E. Wolfe, T. Vértesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, “Geometry of the set of quantum correlations,” *Phys. Rev. A*, vol. 97, 2018. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.97.022104>
- [69] A. U. Rahman, Y. Khedif, M. Javed, H. Ali, and M. Daoud, “Characterizing two-qubit non-classical correlations and non-locality in mixed local dephasing noisy channels,” *Ann. Phys.*, vol. 534, no. 10, 2022. [Online]. Available: <https://doi.org/10.1002/andp.202200197>
- [70] I. Bengtsson, S. Weis, and K. Życzkowski, *Geometry of the Set of Mixed Quantum States: An Apophatic Approach*. Springer Basel, 2012. [Online]. Available: http://dx.doi.org/10.1007/978-3-0348-0448-6_15
- [71] C. Eltschka, M. Huber, S. Morelli, and J. Siewert, “The shape of higher-dimensional state space: Bloch-ball analog for a qutrit,” *Quantum*, vol. 5, 2021. [Online]. Available: <https://doi.org/10.22331/q-2021-06-29-485>
- [72] M. Kuś and K. Życzkowski, “Geometry of entangled states,” *Phys. Rev. A*, vol. 63, 2001. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.63.032307>
- [73] Z. Puchala, J. A. Miszczak, P. Gawron, C. F. Dunkl, J. A. Holbrook, and K. Życzkowski, “Restricted numerical shadow and the geometry of quantum entanglement,” *J. Phys. A*, vol. 45, no. 41, 2012. [Online]. Available: <http://dx.doi.org/10.1088/1751-8113/45/41/415309>
- [74] A. Cabello, “How much larger quantum correlations are than classical ones,” *Phys. Rev. A*, vol. 72, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.72.012113>
- [75] E. Wolfe and S. F. Yelin, “Quantum bounds for inequalities involving marginal expectation values,” *Phys. Rev. A*, vol. 86, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.86.012123>

- [76] C. Duarte, S. Brito, B. Amaral, and R. Chaves, “Concentration phenomena in the geometry of Bell correlations,” *Phys. Rev. A*, vol. 98, 2018. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.98.062114>
- [77] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, “Volume of the set of separable states,” *Phys. Rev. A*, vol. 58, 1998. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.58.883>
- [78] P. B. Slater, “A priori probabilities of separable quantum states,” *J. Phys. A*, vol. 32, no. 28, 1999. [Online]. Available: <http://dx.doi.org/10.1088/0305-4470/32/28/306>
- [79] S. Milz and W. T. Strunz, “Volumes of conditioned bipartite state spaces,” *J. Phys. A*, vol. 48, no. 3, 2014. [Online]. Available: <http://dx.doi.org/10.1088/1751-8113/48/3/035306>
- [80] A. Lovas and A. Andai, “Invariance of separability probability over reduced states in 4×4 bipartite systems,” *J. Phys. A*, vol. 50, no. 29, 2017. [Online]. Available: <http://dx.doi.org/10.1088/1751-8121/aa7176>
- [81] A. Sauer, J. Z. Bernád, H. J. Moreno, and G. Alber, “Entanglement in bipartite quantum systems: Euclidean volume ratios and detectability by Bell inequalities,” *J. Phys. A*, vol. 54, no. 49, 2021. [Online]. Available: <http://dx.doi.org/10.1088/1751-8121/ac3469>
- [82] P. A. Parrilo and S. Lall, “Semidefinite programming relaxations and algebraic optimization in control,” *European Journal of Control*, vol. 9, no. 2-3, 2003. [Online]. Available: <https://www.mit.edu/~parrilo/cdc03-workshop/>
- [83] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, “Complete family of separability criteria,” *Phys. Rev. A*, vol. 69, no. 2, 2004. [Online]. Available: <https://doi.org/10.1103/PhysRevA.69.022308>
- [84] M. Navascués, M. Owari, and M. B. Plenio, “Power of symmetric extensions for entanglement detection,” *Phys. Rev. A*, vol. 80, 2009. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.80.052306>
- [85] M. Navascués, S. Pironio, and A. Acín, “Bounding the set of quantum correlations,” *Phys. Rev. Lett.*, vol. 98, 2007. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.98.010401>
- [86] D. Cavalcanti and P. Skrzypczyk, “Quantum steering: a review with focus on semidefinite programming,” *Reports on Progress in Physics*, vol. 80, no. 2, 2016. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1361-6633/80/2/024001>
- [87] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. California Institute of Technology, 2000.
- [88] —, “Semidefinite programming relaxations for semialgebraic problems,” *Mathematical programming*, vol. 96, 2003. [Online]. Available: <https://link.springer.com/content/pdf/10.1007/s10107-003-0387-5.pdf>
- [89] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, “Distinguishing separable and entangled states,” *Phys. Rev. Lett.*, vol. 88, 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.88.187904>
- [90] J. M. Landsberg, *Tensors: geometry and applications: geometry and applications*. American Mathematical Soc., 2011, vol. 128.

- [91] C. Procesi *et al.*, *Lie groups: an approach through invariants and representations*. Springer, 2007, vol. 115.
- [92] L. Guerini and A. Baraviera, “Birkhoff–von Neumann’s theorem, doubly normalized tensors, and joint measurability,” *Linear Multilinear A.*, vol. 70, 2022. [Online]. Available: <https://doi.org/10.1080/03081087.2022.2158164>
- [93] S. Gudder, “Quantum Markov chains,” *J. Math. Phys.*, vol. 49, no. 7, 2008. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/49/7/072105/231936>
- [94] W. B. Arveson, “On subalgebras of C^* -algebras,” *Acta Mathematica*, vol. 123, 1969. [Online]. Available: <https://projecteuclid.org/journals/acta-mathematica/volume-123/issue-none/Subalgebras-of-C-algebras/10.1007/BF02392388.full>
- [95] S. Virmani and M. B. Plenio, “Construction of extremal local positive-operator-valued measures under symmetry,” *Phys. Rev. A*, vol. 67, no. 6, 2003. [Online]. Available: <https://doi.org/10.1103/PhysRevA.67.062308>
- [96] G. M. D’Ariano, P. L. Presti, and P. Perinotti, “Classical randomness in quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 38, no. 26, 2005. [Online]. Available: <http://dx.doi.org/10.1088/0305-4470/38/26/010>
- [97] J.-P. Pellonpää, “Complete characterization of extreme quantum observables in infinite dimensions,” *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 8, 2011. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1751-8113/44/8/085304>
- [98] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*. Springer Science & Business Media, 2011, vol. 1.
- [99] G. Sentís, B. Gendra, S. D. Bartlett, and A. C. Doherty, “Decomposition of any quantum measurement into extremals,” *J. Phys. A: Math. Theor.*, vol. 46, no. 37, 2013. [Online]. Available: <https://doi.org/10.1088/1751-8113/46/37/375302>
- [100] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, “Simulating positive-operator-valued measures with projective measurements,” *Phys. Rev. Lett.*, vol. 119, no. 19, 2017. [Online]. Available: <https://doi.org/10.1103/PhysRevLett.119.190501>
- [101] E. Martínez-Vargas, C. Pineda, and P. Barberis-Blostein, “Quantum measurement optimization by decomposition of measurements into extremals,” *Sci. Rep.*, vol. 10, no. 1, 2020. [Online]. Available: <https://www.nature.com/articles/s41598-020-65934-w>
- [102] G. De les Coves, T. Netzer, and I. Valentiner-Branth, “Magic squares: Latin, semiclassical, and quantum,” *Journal of Mathematical Physics*, vol. 64, no. 2, 2023. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/64/2/022201/2873948/Magic-squares-Latin-semiclassical-and-quantum?redirectedFrom=fulltext>
- [103] F. Buscemi, K. Kobayashi, and S. Minagawa, “A complete and operational resource theory of measurement sharpness,” *Quantum*, vol. 8, 2024. [Online]. Available: <https://quantum-journal.org/papers/q-2024-01-25-1235/>
- [104] E. G. Effros and S. Winkler, “Matrix convexity: operator analogues of the bipolar and hahn–banach theorems,” *Journal of Functional Analysis*, vol. 144, no. 1, 1997. [Online]. Available: <https://core.ac.uk/outputs/82299739/>
- [105] T.-L. Kriel, “An introduction to matrix convex sets and free spectrahedra,” *Complex Analysis and Operator Theory*, vol. 13, 2019. [Online]. Available: <https://link.springer.com/article/10.1007/s11785-019-00937-8>

- [106] I. Klep and T. Štrelj, “Facial structure of matrix convex sets,” *Journal of Functional Analysis*, vol. 283, no. 7, 2022. [Online]. Available: <http://dx.doi.org/10.1016/j.jfa.2022.109601>
- [107] S. Gudder and G. Nagy, “Sequential quantum measurements,” *J. Math. Phys.*, vol. 42, no. 11, 2001. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/42/11/5212/231213>
- [108] A. Jenčová, “Extremal generalized quantum measurements,” *Linear Algebra Appl.*, vol. 439, no. 12, 2013. [Online]. Available: <https://doi.org/10.1016%2Fj.laa.2013.10.006>
- [109] J. Geller and M. Piani, “Quantifying non-classical and beyond-quantum correlations in the unified operator formalism,” *J. Phys. A*, vol. 47, no. 42, 2014. [Online]. Available: <https://dx.doi.org/10.1088/1751-8113/47/42/424030>
- [110] J. I. de Vicente, “On nonlocality as a resource theory and nonlocality measures,” *J. Phys. A*, vol. 47, no. 42, 2014. [Online]. Available: <https://dx.doi.org/10.1088/1751-8113/47/42/424017>
- [111] D. Rosset, D. Schmid, and F. Buscemi, “Type-independent characterization of spacelike separated resources,” *Phys. Rev. Lett.*, vol. 125, 2020. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.125.210402>
- [112] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, “Causal and localizable quantum operations,” *Phys. Rev. A*, vol. 64, 2001. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.64.052309>
- [113] D. Schmid, H. Du, M. Mudassar, G. Coulter-de Wit, D. Rosset, and M. J. Hoban, “Postquantum common-cause channels: the resource theory of local operations and shared entanglement,” *Quantum*, vol. 5, 2021. [Online]. Available: <https://doi.org/10.22331/q-2021-03-23-419>
- [114] D. Leung and W. Matthews, “On the power of PPT-preserving and non-signalling codes,” *IEEE Trans. Inf. Theory*, vol. 61, no. 8, 2015. [Online]. Available: <https://ieeexplore.ieee.org/document/7115934>
- [115] K. Wang, Z. Song, X. Zhao, Z. Wang, and X. Wang, “Detecting and quantifying entanglement on near-term quantum devices,” *npj Quantum Information*, vol. 8, 2022. [Online]. Available: <https://www.nature.com/articles/s41534-022-00556-w>
- [116] S. Morelli, H. Yamasaki, M. Huber, and A. Tavakoli, “Entanglement detection with imprecise measurements,” *Physical Review Letters*, vol. 128, 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.128.250501>
- [117] I. Frérot, F. Baccari, and A. Acín, “Unveiling quantum entanglement in many-body systems from partial information,” *PRX Quantum*, vol. 3, 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PRXQuantum.3.010342>
- [118] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel, “Bell inequalities for graph states,” *Physical Review Letters*, vol. 95, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.95.120405>
- [119] G. Tóth and O. Gühne, “Entanglement detection in the stabilizer formalism,” *Physical Review A*, vol. 72, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.72.022340>
- [120] M. Ghio, D. Malpetti, M. Rossi, D. Bruß, and C. Macchiavello, “Multipartite entanglement detection for hypergraph states,” *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 4, 2017. [Online]. Available: <https://doi.org/10.1088%2F1751-8121%2Faa99c9>

- [121] F. Huber, I. Klep, V. Magron, and J. Volčič, “Dimension-free entanglement detection in multipartite Werner states,” *Communications in Mathematical Physics*, vol. 396, 2022. [Online]. Available: <https://doi.org/10.1007%2Fs00220-022-04485-9>
- [122] M. A. Jafarizadeh, M. Rezaee, and S. K. A. Seyed Yagoobi, “Bell-state diagonal-entanglement witnesses,” *Physical Review A*, vol. 72, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.72.062106>
- [123] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, “Inseparability criterion for continuous variable systems,” *Physical Review Letters*, vol. 84, no. 12, 2000. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.84.2722>
- [124] R. Simon, “Peres-Horodecki separability criterion for continuous variable systems,” *Physical Review Letters*, vol. 84, no. 12, 2000. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.84.2726>
- [125] C. Marconi, A. Aloy, J. Tura, and A. Sanpera, “Entangled symmetric states and copositive matrices,” *Quantum*, vol. 5, 2021. [Online]. Available: <https://doi.org/10.22331%2Fq-2021-10-07-561>
- [126] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, “The randomized measurement toolbox,” *Nature Reviews Physics*, vol. 5, no. 1, 2022. [Online]. Available: <https://doi.org/10.1038%2Fs42254-022-00535-2>
- [127] O. Gühne and N. Lütkenhaus, “Nonlinear entanglement witnesses,” *Physical Review Letters*, vol. 96, 2006. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.96.170502>
- [128] P. Hyllus and J. Eisert, “Optimal entanglement witnesses for continuous-variable systems,” *New Journal of Physics*, vol. 8, no. 4, 2006. [Online]. Available: <https://doi.org/10.1088%2F1367-2630%2F8%2F4%2F051>
- [129] M. Kotowski, M. Kotowski, and M. Kuś, “Universal nonlinear entanglement witnesses,” *Physical Review A*, vol. 81, 2010. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.81.062318>
- [130] M. Gessner, A. Smerzi, and L. Pezzè, “Metrological nonlinear squeezing parameter,” *Physical Review Letters*, vol. 122, 2019. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.122.090503>
- [131] V. Saggio and P. Walther, “A perspective on few-copy entanglement detection in experiments,” *arXiv: 2201.02641*, 2022. [Online]. Available: <https://arxiv.org/abs/2201.02641>
- [132] P. Liu, Z. Liu, S. Chen, and X. Ma, “Fundamental limitation on the detectability of entanglement,” *Physical Review Letters*, vol. 129, 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.129.230503>
- [133] H. Yamasaki, S. Morelli, M. Miethlinger, J. Bavaresco, N. Friis, and M. Huber, “Activation of genuine multipartite entanglement: Beyond the single-copy paradigm of entanglement characterisation,” *Quantum*, vol. 6, 2022. [Online]. Available: <https://doi.org/10.22331/q-2022-04-25-695>
- [134] R. Trényi, Á. Lukács, P. Horodecki, R. Horodecki, T. Vértesi, and G. Tóth, “Activation of metrologically useful genuine multipartite entanglement,” *arXiv: 2203.05538*, 2022. [Online]. Available: <https://arxiv.org/abs/2203.05538>

- [135] P. Horodecki, “From limits of quantum operations to multicopy entanglement witnesses and state-spectrum estimation,” *Physical Review A*, vol. 68, 2003. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.68.052101>
- [136] K. Chen and H.-K. Lo, “Multipartite quantum cryptographic protocols with noisy GHZ states,” *Quantum Information and Computation*, vol. 7, 2004. [Online]. Available: <https://www.rintonpress.com/journals/doi/QIC7.8-1.html>
- [137] L. S. Bishop, L. Tornberg, D. Price, E. Ginossar, A. Nunnenkamp, A. A. Houck, J. M. Gambetta, J. Koch, G. Johansson, S. M. Girvin, and R. J. Schoelkopf, “Proposal for generating and detecting multi-qubit GHZ states in circuit QED,” *New Journal of Physics*, vol. 11, no. 7, 2009. [Online]. Available: <https://dx.doi.org/10.1088/1367-2630/11/7/073040>
- [138] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, “Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer,” *Journal of Physics Communications*, vol. 5, no. 9, 2021. [Online]. Available: <https://dx.doi.org/10.1088/2399-6528/ac1df7>
- [139] O. Gühne and M. Seevinck, “Separability criteria for genuine multiparticle entanglement,” *New Journal of Physics*, vol. 12, no. 5, 2010. [Online]. Available: <https://dx.doi.org/10.1088/1367-2630/12/5/053002>
- [140] A. Gabriel, B. C. Hiesmayr, and M. Huber, “Criterion for k-separability in mixed multipartite states,” *Quantum Info. Comput.*, vol. 10, no. 9, 2010. [Online]. Available: <https://dl.acm.org/doi/abs/10.5555/2011464.2011472>
- [141] K. M. R. Audenaert, “A digest on representation theory of the symmetric group,” 2006. [Online]. Available: <https://api.semanticscholar.org/CorpusID:9997410>
- [142] R. Merris, *Multilinear Algebra*, ser. Algebra, logic, and applications. Taylor & Francis, 1997.
- [143] M. Marcus and H. Minc, “Generalized matrix functions,” *Transactions of the American Mathematical Society*, vol. 116, 1965. [Online]. Available: <http://www.jstor.org/stable/1994120>
- [144] H. Maassen and B. Kümmerer, “Entanglement of symmetric Werner states,” *Workshop: Mathematics of Quantum Information Theory*, <http://www.bjadres.nl/MathQuantWorkshop/Slides/SymmWernerHandout.pdf>, 2019.
- [145] F. Huber, “Positive maps and trace polynomials from the symmetric group,” *Journal of Mathematical Physics*, vol. 62, no. 2, 2021. [Online]. Available: <https://doi.org/10.1063%2F5.0028856>
- [146] F. Huber and H. Maassen, “Matrix forms of immanant inequalities,” *arXiv: 2103.04317*, 2021. [Online]. Available: <https://arxiv.org/abs/2103.04317>
- [147] I. Klep, V. Magron, and J. Volčič, “Optimization over trace polynomials,” *Annales Henri Poincaré*, vol. 23, no. 1, 2021. [Online]. Available: <https://doi.org/10.1007%2Fs00023-021-01095-4>
- [148] Z. Liu, P. Zeng, Y. Zhou, and M. Gu, “Characterizing correlation within multipartite quantum systems via local randomized measurements,” *Phys. Rev. A*, vol. 105, 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.105.022407>
- [149] H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” *Nature Physics*, vol. 16, no. 10, 2020. [Online]. Available: <https://doi.org/10.1038%2Fs41567-020-0932-7>

- [150] R. Kueng, “Quantum and classical information processing with tensors,” *Caltech CMS Lecture Notes*, 2019. [Online]. Available: <https://resolver.caltech.edu/CaltechAUTHORS:20220412-220843410>
- [151] M. Keyl and R. F. Werner, “Estimating the spectrum of a density operator,” *Physical Review A*, vol. 64, 2001. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.64.052311>
- [152] M. Christandl and G. Mitchison, “The spectra of quantum states and the Kronecker coefficients of the symmetric group,” *Communications in Mathematical Physics*, vol. 261, no. 3, 2005. [Online]. Available: <https://doi.org/10.1007%2Fs00220-005-1435-1>
- [153] T. Eggeling and R. F. Werner, “Hiding classical data in multipartite quantum states,” *Phys. Rev. Lett.*, vol. 89, 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.89.097905>
- [154] J. Siewert and C. Eltschka, “Quantifying tripartite entanglement of three-qubit generalized Werner states,” *Phys. Rev. Lett.*, vol. 108, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.108.230502>
- [155] A. Vesperini, G. Bel-Hadj-Aissa, and R. Franzosi, “Entanglement and quantum correlation measures for quantum multipartite mixed states,” *Scientific Reports*, vol. 13, no. 1, 2023. [Online]. Available: <https://doi.org/10.1038%2Fs41598-023-29438-7>
- [156] B. Collins and P. Śniady, “Integration with respect to the Haar measure on unitary, orthogonal and symplectic group,” *Communications in Mathematical Physics*, vol. 264, no. 3, 2006. [Online]. Available: <https://doi.org/10.1007%2Fs00220-006-1554-3>
- [157] C. Procesi, “A note on the formanek weingarten function,” *Note di Matematica*, vol. 41, no. 1, pp. 69–110, 2021. [Online]. Available: <http://siba-ese.unisalento.it/index.php/notemat/article/view/24154>
- [158] A. Müller-Hermes, D. Reeb, and M. Wolf, “Positivity of linear maps under tensor powers,” *Journal of Mathematical Physics*, vol. 57, 2015. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/57/1/015202/910460>
- [159] M. van der Eyden, T. Netzer, and G. D. las Cuevas, “Halos and undecidability of tensor stable positive maps,” *Journal of Physics A: Mathematical and Theoretical*, vol. 55, no. 26, 2022. [Online]. Available: <https://doi.org/10.1088%2F1751-8121%2F726e>
- [160] N. Wyderka, A. Ketterer, S. Imai, J. L. Bönsel, D. E. Jones, B. T. Kirby, X.-D. Yu, and O. Gühne, “Complete characterization of quantum correlations by randomized measurements,” *arXiv: 2212.07894*, 2022. [Online]. Available: <https://arxiv.org/abs/2212.07894>
- [161] J. S. Lew, “The generalized cayley-hamilton theorem in n dimensions,” *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 17, 1966. [Online]. Available: <https://link.springer.com/article/10.1007/BF01597249#citeas>
- [162] H. Wolkowicz, R. Saigal, and L. Vandenberghe, *Handbook of semidefinite programming: theory, algorithms, and applications*. Springer Science & Business Media, 2012, vol. 27. [Online]. Available: <https://link.springer.com/book/10.1007/978-1-4615-4381-7>
- [163] X. Wang, W. Xie, and R. Duan, “Semidefinite programming strong converse bounds for classical capacity,” *IEEE Trans. Inf. Theory*, vol. 64, no. 1, 2018. [Online]. Available: <https://ieeexplore.ieee.org/document/8012535>

- [164] A. B. Watts, A. Chowdhury, A. Epperly, J. W. Helton, and I. Klep, “Relaxations and Exact Solutions to Quantum Max Cut via the Algebraic Structure of Swap Operators,” *arXiv: 2307.15661*, 2023. [Online]. Available: <https://arxiv.org/abs/2307.15661>
- [165] J. Takahashi, C. Rayudu, C. Zhou, R. King, K. Thompson, and O. Parekh, “An $SU(2)$ -symmetric Semidefinite Programming Hierarchy for Quantum Max Cut,” *arXiv: 2307.15688*, 2023. [Online]. Available: <https://arxiv.org/abs/2307.15688>
- [166] J. Eisert, “A note on lower bounds to variational problems with guarantees,” *arXiv: 2301.06142*, 2023. [Online]. Available: <https://arxiv.org/abs/2301.06142>
- [167] R. Allerstorfer, M. Christandl, D. Grinko, I. Nechita, M. Ozols, D. Rochette, and P. V. Lunel, “Monogamy of highly symmetric states,” *arXiv: 2309.16655*, 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2309.16655>
- [168] K. Gatermann and P. A. Parrilo, “Symmetry groups, semidefinite programs, and sums of squares,” *Journal of Pure and Applied Algebra*, vol. 192, no. 1-3, 2004. [Online]. Available: [url={http://dx.doi.org/10.1016/j.jpaa.2003.12.011}](http://dx.doi.org/10.1016/j.jpaa.2003.12.011),
- [169] C. Procesi, “Some special bases of the 2-swap algebras,” *arXiv: 2406.18687*, 2024. [Online]. Available: <https://arxiv.org/abs/2406.18687>
- [170] N. J. A. Sloane, *A handbook of integer sequences*. Academic Press, 2014.
- [171] T. Scrimshaw, *Symmetric Group Representation*, 2023, SageMath 10.4 package. [Online]. Available: https://doc.sagemath.org/html/en/reference/combinat/sage/combinat/symmetric_group_representations.html
- [172] S. Olver *Numerical Representation Theory*, 2024, Julia package. [Online]. Available: <https://github.com/dlfivefifty/NumericalRepresentationTheory.jl>
- [173] S. Olver, “Representations of the symmetric group are decomposable in polynomial time,” *arXiv: 2211.12592*, 2024. [Online]. Available: <https://arxiv.org/abs/2211.12592>
- [174] O. P. Sevag Gharibian, Felix Huber and K. Thompson, Private communication.
- [175] A. A. Klyachko, *Quantum marginal problem and N -representability*. IOP Publishing, 2006, vol. 36, no. 1.
- [176] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, “Matrix product states and projected entangled pair states: Concepts, symmetries, theorems,” *Rev. Mod. Phys.*, vol. 93, 2021. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.93.045003>
- [177] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018. [Online]. Available: [10.1017/9781316848142](https://doi.org/10.1017/9781316848142)
- [178] S. P. Gudder, *Quantum Probability*. Academic Press, 2014.
- [179] J. R. Norris, *Markov Chains*. Cambridge university press, 1998, no. 2.
- [180] G. Birkhoff, “Tres observaciones sobre el algebra lineal,” *Univ. Nac. Tucuman, Ser. A*, vol. 5, 1946. [Online]. Available: <https://www.scienceopen.com/document?vid=fa429ed1-6420-401f-9c52-ddff427af671>
- [181] A. Ostrowski, “Sur quelques applications des fonctions convexes et concaves au sens de i. schur,” *J. Math. Pures Appl*, vol. 31, no. 9, 1952. [Online]. Available: <https://api.semanticscholar.org/CorpusID:123761542>

- [182] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*. Springer, 1979.
- [183] R. Bhatia, *Matrix Analysis*. Springer Science & Business Media, 1997, vol. 169.
- [184] E. Chitambar and G. Gour, “Quantum resource theories,” *Rev. Mod. Phys.*, vol. 91, 2019. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.91.025001>
- [185] S. Gudder, “Properties of sequential products,” *arXiv: 2307.16327*, 2023. [Online]. Available: <https://arxiv.org/abs/2307.16327>
- [186] G. De les Coves, T. Drescher, and T. Netzer, “Quantum magic squares: Dilations and their limitations,” *J. Math. Phys.*, vol. 61, no. 11, 2020. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/61/11/111704/234732/>
- [187] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2003. [Online]. Available: <https://doi.org/10.1017/CBO9780511546631>
- [188] E. B. Davies and J. T. Lewis, “An operational approach to quantum probability,” *Communications in Mathematical Physics*, vol. 17, no. 3, 1970. [Online]. Available: <https://link.springer.com/article/10.1007/BF01647093>
- [189] K. Życzkowski, “Quartic quantum theory: an extension of the standard quantum mechanics,” *J. Phys. A: Math. Theor.*, vol. 41, no. 35, 2008. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1751-8113/41/35/355302>
- [190] B. Regula and R. Takagi, “Fundamental limitations on distillation of quantum channel resources,” *Nat. Commun.*, vol. 12, no. 1, 2021. [Online]. Available: <https://www.nature.com/articles/s41467-021-24699-0>
- [191] T. Benoist and I. Nechita, “On bipartite unitary matrices generating subalgebra-preserving quantum operations,” *Linear Algebra Appl.*, vol. 521, 2017. [Online]. Available: <https://doi.org/10.1016%2Fj.laa.2017.01.020>
- [192] A. Bluhm, I. Nechita, and S. Schmidt, “Polytope compatibility—from quantum measurements to magic squares,” *Journal of Mathematical Physics*, vol. 64, no. 12, 2023. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/64/12/122201/2930519>
- [193] A. Epperly, E. Evert, J. W. Helton, and I. Klep, “Matrix extreme points and free extreme points of free spectrahedra,” *arXiv: 2212.00748*, 2022. [Online]. Available: <https://arxiv.org/abs/2212.00748>
- [194] G. Hofer and I. G. Todorov, “Quantum hypergraph homomorphisms and non-local games,” *arXiv: 2211.04851*, 2022. [Online]. Available: <https://arxiv.org/abs/2211.04851>
- [195] E. Evert, “Free extreme points of free spectrahedra and generalized free spectrahedra,” *arXiv: 2302.07382*, 2023. [Online]. Available: <https://arxiv.org/abs/2302.07382>
- [196] M. Brannan, S. J. Harris, I. G. Todorov, and L. Turowska, “Quantum no-signalling bicorrelations,” *arXiv: 2302.04268*, 2023. [Online]. Available: <https://arxiv.org/abs/2302.04268>
- [197] M. S. Leifer, *Conditional Density Operators and the Subjectivity of Quantum Operations*. AIP Conference Proceedings, 2007, vol. 889.

- [198] M. S. Leifer and R. W. Spekkens, “Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference,” *Physical Review A*, vol. 88, 2013. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.88.052130>
- [199] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, “Colloquium: Incompatible measurements in quantum information science,” *Rev. Mod. Phys.*, vol. 95, 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.95.011003>
- [200] H. S. Karthik, A. R. U. Devi, and A. K. Rajagopal, “Unsharp measurements and joint measurability,” *Curr. Sci.*, vol. 109, no. 11, 2015. [Online]. Available: <http://www.jstor.org/stable/24906704>
- [201] R. Beneduci, “Joint measurability through naimark’s dilation theorem,” *Rep. Math. Phys.*, vol. 79, no. 2, 2017. [Online]. Available: [https://doi.org/10.1016/S0034-4877\(17\)30035-6](https://doi.org/10.1016/S0034-4877(17)30035-6)
- [202] A. Bluhm and I. Nechita, “Joint measurability of quantum effects and the matrix diamond,” *J. Math. Phys.*, vol. 59, 2018. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/59/11/112202/232581>
- [203] J. Jae, K. Baek, J. Ryu, and J. Lee, “Necessary and sufficient condition for joint measurability,” *Phys. Rev. A*, vol. 100, 2019. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.100.032113>
- [204] A. Mitra and M. Farkas, “Compatibility of quantum instruments,” *Phys. Rev. A*, vol. 105, 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.105.052202>
- [205] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini, “Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication,” *Quantum*, vol. 7, 2023. [Online]. Available: <https://doi.org/10.22331/q-2023-06-07-1035>
- [206] L. Leppäjärvi and M. Sedlák, “Incompatibility of quantum instruments,” *Quantum*, vol. 8, 2024. [Online]. Available: <http://dx.doi.org/10.22331/q-2024-02-12-1246>
- [207] M. A. Nielsen, “Conditions for a class of entanglement transformations,” *Phys. Rev. Lett.*, vol. 83, 1999. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.83.436>
- [208] M. Horodecki and J. Oppenheim, “Fundamental limitations for quantum and nanoscale thermodynamics,” *Nat. Commun.*, vol. 4, no. 1, 2013. [Online]. Available: <https://www.nature.com/articles/ncomms3059>
- [209] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, “Resource theory of quantum states out of thermal equilibrium,” *Phys. Rev. Lett.*, vol. 111, no. 25, 2013. [Online]. Available: <https://doi.org/10.1103/PhysRevLett.111.250404>
- [210] S. Du, Z. Bai, and Y. Guo, “Conditions for coherence transformations under incoherent operations,” *Phys. Rev. A*, vol. 91, 2015. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.91.052120>
- [211] E. Chitambar and G. Gour, “Comparison of incoherent operations and measures of coherence,” *Phys. Rev. A*, vol. 94, 2016. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.94.052336>
- [212] X. Zhan, *Matrix Inequalities*. Springer, 2004.
- [213] H. Weyl, “Inequalities between the two kinds of eigenvalues of a linear transformation,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 35, no. 7, 1949. [Online]. Available: <https://www.pnas.org/doi/abs/10.1073/pnas.35.7.408>

- [214] F. G. S. L. Brandão and G. Gour, “Reversible framework for quantum resource theories,” *Phys. Rev. Lett.*, vol. 115, 2015. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.115.070503>
- [215] M. Oszmaniec and T. Biswas, “Operational relevance of resource theories of quantum measurements,” *Quantum*, vol. 3, 2019. [Online]. Available: <https://quantum-journal.org/papers/q-2019-04-26-133/>
- [216] T. Guff, N. A. McMahon, Y. R. Sanders, and A. Gilchrist, “A resource theory of quantum measurements,” *J. Phys. A: Math. Theor.*, vol. 54, no. 22, 2021. [Online]. Available: <https://doi.org/10.1088%2F1751-8121%2Fabad67>
- [217] L. Tendick, M. Kliesch, H. Kampermann, and D. Bruß, “Distance-based resource quantification for sets of quantum measurements,” *Quantum*, vol. 7, 2023. [Online]. Available: <http://dx.doi.org/10.22331/q-2023-05-15-1003>
- [218] O. Gühne, M. Kleinmann, A. Cabello, J.-A. Larsson, G. Kirchmair, F. Zähringer, R. Gerritsma, and C. F. Roos, “Compatibility and noncontextuality for sequential measurements,” *Phys. Rev. A*, vol. 81, 2010. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.81.022121>
- [219] X.-M. Hu, J.-S. Chen, B.-H. Liu, Y. Guo, Y.-F. Huang, Z.-Q. Zhou, Y.-J. Han, C.-F. Li, and G.-C. Guo, “Experimental test of compatibility-loop-hole-free contextuality with spatially separated entangled qutrits,” *Phys. Rev. Lett.*, vol. 117, 2016. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.117.170403>
- [220] P. Wang, J. Zhang, C.-Y. Luan, M. Um, Y. Wang, M. Qiao, T. Xie, J.-N. Zhang, A. Cabello, and K. Kim, “Significant loop-hole-free test of kochen-specker contextuality using two species of atomic ions,” *Sci. Adv.*, vol. 8, no. 6, 2022. [Online]. Available: <https://www.science.org/doi/abs/10.1126/sciadv.abk1660>
- [221] J. W. Helton, I. Klep, and S. McCullough, “Matrix convex hulls of free semialgebraic sets,” *Trans. Am. Math. Soc.*, vol. 368, no. 5, 2015. [Online]. Available: <http://dx.doi.org/10.1090/tran/6560>
- [222] J. Ginibre, “Statistical ensembles of complex, quaternion, and real matrices,” *J. Math. Phys.*, vol. 6, no. 3, 1965. [Online]. Available: <https://pubs.aip.org/aip/jmp/article-abstract/6/3/440/595508>
- [223] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, “Generating random density matrices,” *J. Math. Phys.*, vol. 52, no. 6, 2011. [Online]. Available: <https://doi.org/10.1063%2F1.3595693>
- [224] R. Bellman, *Some Inequalities for Positive Definite Matrices*. Birkhäuser Basel: Beckenbach, E.F. (eds) General Inequalities 2., 1980. [Online]. Available: https://doi.org/10.1007/978-3-0348-6324-7_8
- [225] S. Wiesner, “Conjugate coding,” *ACM SIGACT News*, vol. 15, no. 1, 1983. [Online]. Available: <https://dl.acm.org/doi/pdf/10.1145/1008908.1008920>
- [226] S. Pironio, A. Acín, S. Massar, A. B. de La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, “Random numbers certified by Bell’s theorem,” *Nature*, vol. 464, no. 7291, 2010. [Online]. Available: <https://doi.org/10.1038/nature09008>

- [227] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” *Theoret. Comput. Sci.*, vol. 560, 1984. [Online]. Available: <https://doi.org/10.1016/j.tcs.2014.05.025>
- [228] J. Barrett, “Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality,” *Phys. Rev. A*, vol. 65, no. 4, 2002. [Online]. Available: <https://doi.org/10.1103/PhysRevA.65.042302>
- [229] M. J. Hoban and A. B. Sainz, “A channel-based framework for steering, non-locality and beyond,” *New J. Phys.*, vol. 20, no. 5, 2018. [Online]. Available: <https://dx.doi.org/10.1088/1367-2630/aabea8>
- [230] D. Rosset, F. Buscemi, and Y.-C. Liang, “Resource theory of quantum memories and their faithful verification with minimal assumptions,” *Phys. Rev. X*, vol. 8, no. 2, 2018. [Online]. Available: <https://doi.org/10.1103/PhysRevX.8.021033>
- [231] J. H. Selby, A. B. Sainz, V. Magron, L. Czekaj, and M. Horodecki, “Correlations constrained by composite measurements,” *Quantum*, vol. 7, 2023. [Online]. Available: <https://doi.org/10.22331/q-2023-08-10-1080>
- [232] S. Massar, S. Pironio, J. Roland, and B. Gisin, “Bell inequalities resistant to detector inefficiency,” *Phys. Rev. A*, vol. 66, 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.66.052112>
- [233] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, “Quantum nonlocality in two three-level systems,” *Phys. Rev. A*, vol. 65, 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.65.052325>
- [234] S. G. A. Brito, B. Amaral, and R. Chaves, “Quantifying Bell nonlocality with the trace distance,” *Phys. Rev. A*, vol. 97, 2018. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.97.022111>
- [235] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, “The resource theory of nonclassicality of channel assemblages,” *Quantum*, vol. 7, 2023. [Online]. Available: <https://doi.org/10.22331/q-2023-10-10-1134>
- [236] Y. Liu and X. Yuan, “Operational resource theory of quantum channels,” *Phys. Rev. Res.*, vol. 2, 2020. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.012035>
- [237] A. Peres, *Quantum Theory: Concepts and Methods*. Kluwer Academic, London, 1998, vol. 72. [Online]. Available: <https://doi.org/10.1007/0-306-47120-5>
- [238] M. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*. Springer, The Frontiers Collection, 2007. [Online]. Available: <https://doi.org/10.1007/978-3-540-35775-9>
- [239] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control*. Cambridge university press, 2009. [Online]. Available: <https://doi.org/10.1017/CBO9780511813948>
- [240] I. Devetak and P. W. Shor, “The capacity of a quantum channel for simultaneous transmission of classical and quantum information,” *Commun. Math. Phys.*, vol. 256, 2005. [Online]. Available: <https://doi.org/10.1007/s00220-005-1317-6>
- [241] A. D’Arrigo, G. Benenti, and G. Falci, “Quantum capacity of dephasing channels with memory,” *New J. Phys.*, vol. 9, no. 9, 2007. [Online]. Available: <https://dx.doi.org/10.1088/1367-2630/9/9/310>

- [242] K. Brádler, P. Hayden, D. Touchette, and M. M. Wilde, “Trade-off capacities of the quantum hadamard channels,” *Phys. Rev. A*, vol. 81, 2010. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.81.062312>
- [243] M. Hasenöhrl and M. C. Caro, “Quantum and classical dynamical semigroups of superchannels and semicausal channels,” *J. Math. Phys.*, vol. 63, no. 7, 2022. [Online]. Available: <https://doi.org/10.1063/5.0070635>
- [244] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, “Quantifying EPR: the resource theory of nonclassicality of common-cause assemblages,” *Quantum*, vol. 7, 2023. [Online]. Available: <https://doi.org/10.22331/q-2023-02-16-926>
- [245] M. Berta, F. Borderi, O. Fawzi, and V. B. Scholz, “Semidefinite programming hierarchies for constrained bilinear optimization,” *Math. Program.*, vol. 194, 2022. [Online]. Available: <https://doi.org/10.1007/s10107-021-01650-1>
- [246] T. Baumgratz, M. Cramer, and M. B. Plenio, “Quantifying coherence,” *Phys. Rev. Lett.*, vol. 113, 2014. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.113.140401>
- [247] A. Fine, “Hidden variables, joint probability, and the Bell inequalities,” *Phys. Rev. Lett.*, vol. 48, 1982. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.48.291>
- [248] D. Rosset, J.-D. Bancal, and N. Gisin, “Classifying 50 years of Bell inequalities,” *J. Phys. A*, vol. 47, no. 42, 2014. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1751-8113/47/42/424022>
- [249] M. Navascués, S. Pironio, and A. Acín, “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations,” *New J. Phys.*, vol. 10, no. 7, 2008. [Online]. Available: <https://doi.org/10.1088/1367-2630/10/7/073013>
- [250] B. S. Tsirel’son, “Quantum analogues of the Bell inequalities. the case of two spatially separated domains,” *J. Sov. Math.*, vol. 36, 1987. [Online]. Available: <https://doi.org/10.1007/BF01663472>
- [251] L. J. Landau, “Empirical two-point correlation functions,” *Found. Phys.*, vol. 18, 1988. [Online]. Available: <https://doi.org/10.1007/BF00732549>
- [252] L. Masanes, “Extremal quantum correlations for n parties with two dichotomic observables per site,” *arXiv: 0512100*, 2005. [Online]. Available: <https://doi.org/10.48550/arXiv.quant-ph/0512100>
- [253] A. Mikos-Nuszkiewicz and J. Kaniewski, “Extremal points of the quantum set in the CHSH scenario: conjectured analytical solution,” *arXiv: 2302.10658*, 2023. [Online]. Available: <https://arxiv.org/abs/2302.10658>
- [254] L. P. Thinh, A. Varvitsiotis, and Y. Cai, “Geometric structure of quantum correlators via semidefinite programming,” *Phys. Rev. A*, vol. 99, 2019. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.99.052108>
- [255] T. P. Le, C. Meroni, B. Sturmfels, R. F. Werner, and T. Ziegler, “Quantum correlations in the minimal scenario,” *Quantum*, vol. 7, 2023. [Online]. Available: <http://dx.doi.org/10.22331/q-2023-03-16-947>
- [256] B. S. Cirel’son, “Quantum generalizations of Bell’s inequality,” *Lett. Math. Phys.*, vol. 4, 1980. [Online]. Available: <https://doi.org/10.1007/BF00417500>

- [257] S. Pironio, “Violations of Bell inequalities as lower bounds on the communication cost of nonlocal correlations,” *Phys. Rev. A*, vol. 68, 2003. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.68.062102>
- [258] W. Van Dam, R. D. Gill, and P. D. Grunwald, “The statistical strength of nonlocality proofs,” *IEEE Trans. Inf. Theory*, vol. 51, no. 8, 2005. [Online]. Available: <https://ieeexplore.ieee.org/document/1468301>
- [259] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, “Nonlocal correlations as an information-theoretic resource,” *Phys. Rev. A*, vol. 71, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.71.022101>
- [260] J. Y. Lee, C.-M. Jian, and C. Xu, “Quantum criticality under decoherence or weak measurement,” *PRX Quantum*, vol. 4, 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/PRXQuantum.4.030317>
- [261] J. Y. Lee, Y.-Z. You, and C. Xu, “Symmetry protected topological phases under decoherence,” *arXiv: 2210.16323*, 2024. [Online]. Available: <https://arxiv.org/abs/2210.16323>
- [262] A. Lyons, “Understanding stabilizer codes under local decoherence through a general statistical mechanics mapping,” *arXiv: 2403.03955*, 2024. [Online]. Available: <https://arxiv.org/abs/2403.03955>
- [263] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines, “Quantum dynamics as a physical resource,” *Phys. Rev. A*, vol. 67, 2003. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.67.052301>
- [264] M. Musz, M. Kuś, and K. Życzkowski, “Unitary quantum gates, perfect entanglers, and unistochastic maps,” *Phys. Rev. A*, vol. 87, 2013. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.87.022111>
- [265] R. Mansuroglu, A. Adil, M. J. Hartmann, Z. Holmes, and A. T. Sornborger, “Quantum tensor-product decomposition from Choi-state tomography,” *PRX Quantum*, vol. 5, 2024. [Online]. Available: <https://link.aps.org/doi/10.1103/PRXQuantum.5.030306>
- [266] A. B. Sainz, L. Aolita, M. Piani, M. J. Hoban, and P. Skrzypczyk, “A formalism for steering with local quantum measurements,” *New J. Phys.*, vol. 20, no. 8, 2018. [Online]. Available: <https://dx.doi.org/10.1088/1367-2630/aad8df>
- [267] T. Eggeling, D. Schlingemann, and R. F. Werner, “Semicausal operations are semilocalizable,” *Europhys. Lett.*, vol. 57, no. 6, 2002. [Online]. Available: <https://iopscience.iop.org/article/10.1209/epl/i2002-00579-4>
- [268] B. Regula, “Convex geometry of quantum resource quantification,” *J. Phys. A*, vol. 51, no. 4, 2017. [Online]. Available: <http://dx.doi.org/10.1088/1751-8121/aa9100>
- [269] B. Schumacher and M. D. Westmoreland, “Relative entropy in quantum information theory,” *Contemp. Math.*, vol. 305, 2002. [Online]. Available: <https://doi.org/10.1090/conm%2F305%2F05225>
- [270] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Rényi entropies: A new generalization and some properties,” *J. Math. Phys.*, vol. 54, no. 12, 2013. [Online]. Available: <http://dx.doi.org/10.1063/1.4838856>
- [271] H. Fawzi and O. Fawzi, “Defining quantum divergences via convex optimization,” *Quantum*, vol. 5, 2021. [Online]. Available: <https://doi.org/10.22331/q-2021-01-26-387>

- [272] C. W. Helstrom, “Quantum detection and estimation theory,” *J. Stat. Phys.*, vol. 1, 1969. [Online]. Available: <http://dx.doi.org/10.1007/BF01007479>
- [273] C. Fuchs and J. van de Graaf, “Cryptographic distinguishability measures for quantum-mechanical states,” *IEEE Trans. Inf. Theory*, vol. 45, no. 4, 1999. [Online]. Available: <http://dx.doi.org/10.1109/18.761271>
- [274] J. Barrett, “Information processing in generalized probabilistic theories,” *Phys. Rev. A*, vol. 75, 2007. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.75.032304>
- [275] N. S. Jones and L. Masanes, “Interconversion of nonlocal correlations,” *Phys. Rev. A*, vol. 72, 2005. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.72.052312>
- [276] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, “Operational framework for nonlocality,” *Phys. Rev. Lett.*, vol. 109, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.109.070401>
- [277] D. Kaszlikowski, P. Gnaniński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, “Violations of local realism by two entangled N -dimensional systems are stronger than for two qubits,” *Phys. Rev. Lett.*, vol. 85, 2000. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.85.4418>
- [278] D. Pérez-García, M. M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge, “Unbounded violation of tripartite Bell inequalities,” *Commun. Math. Phys.*, vol. 279, 2008. [Online]. Available: <https://doi.org/10.1007/s00220-008-0418-4>
- [279] Y. M. Chee, H. Ta, and V. K. Vu, “Efficient approximation of quantum channel fidelity exploiting symmetry,” *arXiv: 2308.15884*, 2024. [Online]. Available: <https://arxiv.org/abs/2308.15884>
- [280] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” *Phys. Rev. Lett.*, vol. 23, 1969. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.23.880>
- [281] F. G. S. L. Brandão, J. Eisert, M. Horodecki, and D. Yang, “Entangled inputs cannot make imperfect quantum channels perfect,” *Phys. Rev. Lett.*, vol. 106, 2011. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.106.230502>
- [282] L. Wang and R. Renner, “One-shot classical-quantum capacity and hypothesis testing,” *Phys. Rev. Lett.*, vol. 108, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.108.200501>
- [283] M. Mosonyi and N. Datta, “Generalized relative entropies and the capacity of classical-quantum channels,” *J. Math. Phys.*, vol. 50, no. 7, 2009. [Online]. Available: <http://dx.doi.org/10.1063/1.3167288>
- [284] J. M. Renes and R. Renner, “Noisy channel coding via privacy amplification and information reconciliation,” *IEEE Trans. Inf. Theory*, vol. 57, no. 11, 2011. [Online]. Available: <http://dx.doi.org/10.1109/TIT.2011.2162226>
- [285] M. Tomamichel, M. M. Wilde, and A. Winter, “Strong converse rates for quantum communication,” *IEEE Trans. Inf. Theory*, vol. 63, no. 1, 2017. [Online]. Available: <http://dx.doi.org/10.1109/TIT.2016.2615847>
- [286] B. F. Toner and D. Bacon, “Communication cost of simulating Bell correlations,” *Phys. Rev. Lett.*, vol. 91, 2003. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.91.187904>

- [287] C. H. Bennett and S. J. Wiesner, “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states,” *Phys. Rev. Lett.*, vol. 69, 1992. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.69.2881>