The interplay between quantum entanglement, coherence, and convex optimization

Dissertation
zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von
Timo Yannick Simnacher

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen, 2021

Betreuer und erster Gutachter
Prof. Dr. Otfried Gühne
Universität Siegen

Zweiter Gutachter
Prof. Dr. Karol Życzkowski
Uniwersytet Jagielloński

Tag der mündlichen Prüfung
22. Juli 2021

To my family, my love, and my friends.

Abstract

In this thesis, we strive to advance the knowledge of relations between convex optimization and the quantum phenomena entanglement and coherence. The main research areas we explore are rank-constrained semidefinite programming, the quantum pure-state marginal problem and the existence of AME states as well as quantum codes, entanglement detection, and the certification of quantum memories with coherence.

First, we start with real and complex rank-constraint semidefinite optimization problems and rephrase them as an optimization over separable two-copy states. This reformulation allows to approach the problem through a hierarchy of efficiently solvable semidefinite programs that provide better and better certified bounds. We apply the new technique to various problems in quantum information theory and beyond, such as the optimization over pure states or unitary channels and the well-known maximum cut problem. Furthermore, we describe an inherent symmetry in our formulation that significantly improves the performance.

Second, we consider the application of our method to the quantum pure-state marginal problem. In particular, we prove that the existence of n-partite absolutely maximally entangled states with local dimension d is equivalent to the bipartite separability of a certain state of $2n$ particles, and we compute that state explicitly. This application is a striking example of how symmetries can simplify semidefinite programs and we use them to compute high orders of our hierarchy despite the rapidly increasing dimension. Moreover, we rewrite the existence problem of quantum error-correcting codes as a marginal problem making our method also applicable to this area of research.

Third, since entanglement is not only a theoretically interesting phenomenon, but also a vital resource for quantum information protocols, we investigate entanglement detection in practical experiments. We examine scrambled data, a scenario in which the mapping between outcomes and their respective probabilities is lost. Furthermore, we use the joint numerical range of observables to find measurements that allow entanglement detection even when the confidence region due to statistical and systematic errors is large.

Finally, we introduce a quality measure for quantum memories that quantifies the performance based on the memory's ability to preserve coherence. Remarkably, this measure also distinguishes entanglement-breaking channels from genuine quantum memories. For the case of single-qubit channels, we find various theoretical bounds and a simple measurement scheme to approximate our performance measure.

Zusammenfassung

Mit dieser Dissertation wollen wir das Verständnis der Zusammenhänge zwischen konvexer Optimierung und der Quantenphänomene Verschränkung und Kohärenz erweitern. Die Hauptforschungsgebiete, die wir erkunden, sind rangbeschränkte semidefinite Programmierung, das Marginalproblem reiner Quantenzustände und die Existenz von AME-Zuständen sowie Quantencodes, Verschränkungsdetektion und die Zertifizierung von Quantenspeichern mittels Kohärenz.

Als erstes beschäftigen wir uns mit reellen und komplexen rangbeschränkten semidefiniten Optimierungsproblemen und formulieren diese als Optimierung über separierbare Zwei-Kopien-Zustände um. Das erlaubt es, mittels einer Hierarchie effizient lösbarer semidefiniter Programme immer bessere zertifizierte Schranken zu berechnen. Wir wenden die Methode auf verschiedene Probleme in der Quanteninformationstheorie an, wie etwa die Optimierung über reine Zustände oder unitäre Kanäle und das Problem des maximalen Schnitts eines Graphen. Außerdem beschreiben wir eine inhärente Symmetrie unserer Formulierung, die die Komplexität erheblich verringert.

Dann wenden wir unsere Methode auf das Marginalproblem reiner Quantenzustände an. Insbesondere beweisen wir, dass die Existenz n-partiter absolut maximal verschränkter Zustände mit lokaler Dimension d äquivalent zu der bipartiten Separierbarkeit eines bestimmten $2n$-Teilchenzustands ist, den wir explizit berechnen. Das zeigt eindrucksvoll, wie Symmetrien semidefinite Programme vereinfachen können, sodass hohe Ordnungen unserer Hierarchie trotz rasch steigender Dimension berechenbar sind. Ferner formulieren wir das Existenzproblem von Quantenfehlerkorrekturcodes als Marginalproblem, sodass unsere Methode auch hierfür anwendbar wird.

Da Verschränkung nicht nur theoretisch interessant ist, sondern auch eine essentielle Ressource für Quanteninformationsprotokolle, erforschen wir anschließend deren Detektion in der Praxis. Wir untersuchen ein Szenario, bei dem die Zuordnung von Messergebnissen zu den entsprechenden Wahrscheinlichkeiten unklar ist. Außerdem nutzen wir das gemeinsame numerische Bild von Observablen, um Messungen zu finden, die Verschränkungsdetektion selbst dann ermöglichen, wenn die Konfidenzregion aufgrund statistischer und systematischer Fehler relativ groß ist.

Abschließend stellen wir ein Qualitätsmaß für Quantenspeicher vor, das die Leistungsfähigkeit auf Basis von Kohärenzerhaltung misst. Bemerkenswerterweise differenziert das Maß auch zwischen verschränkungszerstörenden Kanälen und echten Quantenspeichern. Für Ein-Qubit-Kanäle beschreiben wir theoretische Schranken und einfache Messungen, um das Maß näherungsweise zu bestimmen.

Table of contents

"Denn wenn man nicht zunächst über die Quantentheorie entsetzt ist, kann man sie doch unmöglich verstanden haben."
Niels Bohr

1 Introduction

Since physical phenomena had first been explained by their quantum nature more than a hundred years ago [1, 2], quantum physics has been established as one of the most well-tested theories, if not the most well-tested theory in science. On the one hand, the discovery of quantum physics advanced our theoretical knowledge of the inner workings of nature, explaining how atoms consisting of positively and negatively charged particles can form a stable system. Nowadays, the quantum field theory known as the Standard Model of particle physics describes three out of the four fundamental forces: the electromagnetic, weak and strong interactions. Only the gravitational force is still resisting its reconciliation with quantum theory, however, recent works suggest that experiments probing the quantum nature of gravity might be in reach in the near future [3], which may help finding a theory of everything. On the other hand, quantum physics enabled unprecedented technological progress through the development of transistors, the semiconductor devices that are the core components of classical computers, and lasers.

Since the notion of quantum information theory had been termed almost fifty years ago [4], many scientific efforts have been made to better understand quantum physics and its relation to information theory. While in the early stages many prominent physicists, among them Albert Einstein, were quite skeptical towards quantum theory as a complete description of nature [5], especially because of its probabilistic character according to the Copenhagen interpretation, Bell's seminal work [6] and its recent loophole-free experimental implementation [7–9] show that it is impossible to find a macrorealistic, local theory explaining the quantum correlations observed in experiments. Although being a common misunderstanding, spatially separated entangled, i.e. quantum correlated, particles do not allow for superluminal communication, but there are also correlations not allowed by quantum theory which still respect special relativity and prevent superluminal information exchange. Thus, both the boundary between classical and quantum physics as well as the confinement restricting quantum correlations are subject of ongoing theoretical research. Furthermore, quantum

information theory is on the verge of fulfilling the promise of an entirely new technology: quantum computation, Feynman's idea to efficiently simulate quantum particles using a well-controlled quantum system - the computer [10]. This idea inspired a wide field of research leading to experiments with highly controllable and manipulable microscopic systems, improved sensing techniques, and powerful ways to speed up computation, most prominently using Shor's algorithm that facilitates breaking the RSA cryptosystem on a digital quantum computer [11]. However, quantum information exchange also provides a cryptosystem whose security relies merely on the validity of quantum mechanics [12–14].

There still remain numerous essential concepts which are not yet fully understood such as entanglement, coherence, and Bell nonlocality, and ground-breaking results are still discovered, such as the recent breakthrough $\mathrm{MIP}^* = \mathrm{RE}$ [15]. This result establishes an intriguing connection between entanglement, one of the fundamental phenomena distinguishing quantum from classical physics, and the expressive power of multiple interactive provers. A classical verifier with limited resources can be convinced of the solution to the Halting Problem for a given program by the all-powerful quantum provers if the verifier is assured that the provers can share quantum correlations but no correlations which are outside the possibilities of quantum physics. Quantum entanglement allows the provers to establish a convincing joint argument while still allowing the verifier to interrogate them independently in a way that prevents them from cheating, i.e., the verifier can exclude the possibility that the provers try to convince them of a false statement. In Chapter 3 of this thesis, we describe another interesting connection between entanglement and a seemingly unrelated topic, namely, rank-constraint semidefinite optimization. With this new method, findings from entanglement theory can be applied to many optimization problems such as pseudo-boolean optimization, the maximum cut problem of graphs, and the optimization over pure quantum states.

Quantum entanglement is one of the key concepts in quantum information theory since it is indispensible for basic building blocks of quantum communication like quantum teleportation, entanglement swapping, and superdense coding as well as other types of quantum correlations such as steering and Bell nonlocality [16]. For systems consisting of only two particles, there exists a - up to local unitary transformations - unique maximally entangled state from which any other state can be reached via local operations and classical communication. Considering more particles, however, there are different notions of maximally entangled states, that cannot be interconverted through local operations and classical communication. One of these notions is the concept of absolutely maximally entangled states, which are pure states

with the property that the reduced state of any at most half of the particles is maximally mixed, i.e., the reduced state contains no information as indicated by its maximal von Neumann entropy. Absolutely maximally entangled states do not exist for every number of particles and local dimension [17]. Based on the method connecting entanglement and rank-constraint semidefinite optimization mentioned above, we describe in Chapter 4 an algorithmic method that decides the existence of absolutely maximally entangled states, making it computationally feasible by heavily utilizing underlying symmetries. Furthermore, we depict in detail how the method can be used to decide the existence of quantum error correcting codes, a vital tool for reliable future quantum computation.

Since entanglement is such an invaluable resource for quantum information processing, it is essential to verify the presence of entanglement in experiments. This is usually not a simple task because it requires precise measurements of highly sensitive microscopic systems. The number of needed measurements to characterize a quantum state scales exponentially with the number of particles, and hence, it is often not a viable approach. Instead, a smartly selected small number of easily measurable local observables is a better strategy. In this work, we consider entanglement detection in the presence of data scrambling, a measurement error that prevents the association of measurement outcomes to outcome probabilities, see Chapter 5, as well as efficient entanglement detection with few measurements such that large confidence regions in experiments allow for statistically significant entanglement verification, see Chapter 6.

Finally, we describe in Chapter 7 how the quality of quantum memories can be characterized in terms of their ability to preserve the coherence of the stored quantum state. As indispensible building blocks of future large-scale quantum computers quantum memories preserve quantum states over an extended period of time protecting it against decoherence through the interaction with the environment. In contrast to classical bits where bit flips are the source of errors, qubits can additionally undergo erroneous phase transformations and, depending on the underlying system, also particle loss is a significant issue. Thus, the validation of a functioning quantum memory is a lot more complex than the validation of its classical counterpart. Our research advances the understanding of valuable quantum memories with respect to the amount of coherence they preserve.

The individual chapters of this thesis are written in a way that they can be read independently. In Chapter 2, the mathematical foundations to understand this thesis are explained. Following the presentation in Ref. [18], we highlight the necessary and helpful prerequisites at the beginning of each subsequent chapter. Each of these main chapters is strongly based on a corresponding scientific publication replicating most

of the text which has been revised several times already, however, augmented with insightful supplements and enlightning connections between the different topics.

2 Mathematical fundamentals

2.1 Introduction

In this chapter, we introduce the mathematical formalism necessary to understand the thesis. Although we explain the fundamental concepts, a thorough treatment of every topic is impossible as some of them fill entire books. However, further references are given that allow the interested reader to acquire advanced knowledge.

We start with the foundations of quantum mechanics, explaining the underlying axioms for pure and mixed states. Special attention is given to the theory of measurements and quantum states of multiple systems. Subsequently, we focus on the simplest quantum system, the two-dimensional qubit and its representation via the Bloch ball.

In addition to the time evolution of closed systems given by the Schrödinger equation, we discuss quantum channels which describe the general open system dynamics. We describe the Kraus representation as well as the Choi-Jamiołkowski isomorphism which allows to transfer properties of quantum states to channels and vice versa. Examples of single-qubit channels illustrate the concept intuitively.

Having laid out the fundamentals, we introduce two of the most important quantum phenomena, namely, coherence and entanglement. Based on the corresponding resource theories, we highlight similarities and differences between the two. In particular, the notion of a maximally resourceful state is examined which breaks down in the scenario of multipartite, i.e., at least tripartite, entanglement. Moreover, we describe different methods for entanglement detection such as entanglement witnesses, positive maps, especially the PPT criterion, and numerical range. In the context of spatial quantum correlations, we also explain the marginal problem as well as quantum error-correcting codes.

Furthermore, we present semidefinite programming as an exceedingly valuable tool for numerical and analytical optimization. We spotlight the apparent connection to quantum physics and describe the Doherty-Parrilo-Spedalieri hierarchy as a striking

example for its application to entanglement theory. Finally, classical entropies and (their relation to) majorization are introduced as fundamental information theoretic concepts.

2.2 Quantum mechanics

At the end of the nineteenth and the beginning of the twentieth century, experiments showed that classical electrodynamics is not sufficient to describe nature at the microscopic level. Most importantly, the photoelectric effect and black-body radiation could only be explained with the invention of quantum mechanics. After a rigorous mathematical foundation was laid out, many more experiments confirmed the newly developed theory. Up until today, quantum theory became one of the most well-tested theories of nature.

Throughout this thesis, we restrict ourselves to finite-dimensional quantum systems. Almost all, if not all, quantum information protocols can be implemented using finite-dimensional systems and hence, this restriction is not essential. In many cases, even though there is an infinite-dimensional quantum system accessible such as the energy levels of an ion or the position of a photon, only a hand full of those energy levels is used for quantum manipulation and computation or just a finite number of possible paths is considered, respectively. Nevertheless, there are interesting consequences of effects when continuous, infinite-dimensional systems are investigated, and it is an intriguing open question whether nature at the fundamental level is indeed discrete or continuous. This question is closely related to the problem of unifying quantum theory and general relativity. The continuous, geometrical character of gravity qualitatively differs drastically from the quanta and uncertainty relations appearing in quantum physics.

2.2.1 The postulates of quantum mechanics

In this introduction, we mainly follow two classic introductions to quantum information theory, namely Ref. [19, 20]. Since quantum information obeys and utilizes the laws of quantum mechanics, we start with explaining the description of physical systems and their behavior in a quantum world. To do so, we repeat the axioms or postulates of quantum mechanics. The first postulate concerns the description of a quantum state. In contrast to classical physics, the state of a quantum system cannot be described by a collection of its properties such as its position in phase space. Instead, a pure quantum state is given by a vector in the, in our case finite-dimensional,

complex Hilbert space $\mathbb{C}^N$, where N is the dimension of the system. More precisely, a pure quantum state corresponds to a whole equivalence class in this space because quantum states are normalized and hence, we restrict ourselves to vectors of length 1, and an overall complex phase $e^{i\varphi}$ is not observable and thus, describes the same physical state. Mathematically speaking, the space of quantum states is the complex projective space $\mathbb{C}\mathbf{P}^{N-1}$. We will usually use the Dirac notation, also know as bra-ket notation, where the state vector is notated as a *ket* vector $|\psi\rangle$. The canonical or standard basis, which is often called computational basis in quantum information theory, is denoted as $|0\rangle, |1\rangle, \ldots, |N-1\rangle$ and we can express $|\psi\rangle = \sum_j \psi_j |j\rangle$, where the ψ_j are complex numbers and normalization requires $\sum_j |\psi_j|^2 = 1$. The dual vector of $|\psi\rangle$ is a *bra* vector $\langle\psi| = \sum_j \psi_j^* \langle j|$, where the coefficients are given by the complex conjugate, and the inner product is given by $\langle\phi|\psi\rangle = \sum_j \phi_j^* \psi_j$.

Quantum states are often not completely known. For instance, the delicate preparation of a quantum state usually introduces errors which might not be negligible or an attacker of a quantum communication protocol has to work with incomplete information. To describe the state of a quantum system in such scenarios, we use the density matrix introduced by Landau [21] and von Neumann [22]. If the system is known to be in the pure state $|\psi_j\rangle$ with probability p_j, we can describe its state by an ensemble of pure states $\{p_j, |\psi_j\rangle\}$. However, it turns out that different ensembles cannot always be distinguished physically. That is why it is sufficient to instead consider the density matrix

$$\rho = \sum_j p_j |\psi_j\rangle \langle\psi_j| . \tag{2.1}$$

We call a state which is not pure, i.e., there is more than one nonzero p_j, a mixed state. By construction, ρ is a positive semidefinite, Hermitian operator of trace one because of the normalization. We denote these properties by first, the conjugate transpose or Hermitian transpose, i.e., for the Hermitian operator ρ is holds that $\rho = \rho^\dagger$. Represented in the computational basis, we have that for $\rho = \sum_{i,j} \rho_{ij} |i\rangle \langle j|$ the conjugate transpose is given by $\rho^\dagger = \sum_{i,j} \rho_{ij}^* |j\rangle \langle i|$. Second, the eigenvalues of ρ, which are all real because of the Hermiticity, are all nonnegative, i.e., ρ is positive semidefinite, denoted as $\rho \geq 0$. Finally, normalization requires $\operatorname{Tr}\rho = 1$. Moreover, any positive semidefinite, Hermitian matrix of trace one can be written in the form of Eq. 2.1 as a convex combination of projectors using its spectral decomposition. Hence, any such matrix describes a quantum state. A pure state is then just the projector onto its one-dimensional subspace $|\psi\rangle \langle\psi|$. Note that we also got rid of the global phase $e^{i\varphi}$ that leads to the same physical state since $|\psi\rangle$ and $e^{i\varphi} |\psi\rangle$ yield the same density matrix. Pure and mixed states can be distinguished mathematically considering the trace of the squared density operator. It holds that $\operatorname{Tr}\rho^2 = 1$ if, and only if, ρ describes a pure

quantum state. Otherwise, we have $\operatorname{Tr}\rho^2 < 1$. This is why $\gamma = \operatorname{Tr}\rho^2$ is also known as the purity of ρ.

The second postulate determines the evolution of a quantum state over time. It states that a closed quantum system transforms via a unitary transformation, i.e.,

$$|\psi(t_2)\rangle = U(t_1, t_2)\,|\psi(t_1)\rangle\,, \tag{2.2}$$

where $|\psi(t_1)\rangle$ and $|\psi(t_2)\rangle$ are the states of the system at times t_1 and t_2, respectively, and $U(t_1, t_2)$ is a unitary operator meaning it holds that $UU^\dagger = U^\dagger U = \mathbb{1}$. The unitarity ensures that the normalization is preserved over time. The density matrix transforms appropriately as

$$\rho(t_2) = U(t_1, t_2)\rho(t_1)U^\dagger(t_1, t_2). \tag{2.3}$$

The continuous evolution of a quantum system is described by the Schrödinger equation

$$i\hbar\frac{d}{dt}\,|\psi\rangle = H\,|\psi\rangle\,, \tag{2.4}$$

where $\hbar$ is Planck's constant and the Hermitian operator H is called the Hamiltonian of the system. Correspondingly, for mixed states we have that

$$i\hbar\frac{d}{dt}\rho = [H, \rho]\,, \tag{2.5}$$

where $[A, B] = AB - BA$ denotes the commutator of A and B. This equation is called the von Neumann or Liouville-von Neumann equation. The time evolution implies that the stationary states of the closed system are exactly the eigenstates $|E\rangle$ of the Hamiltonian, i.e., $H\,|E\rangle = E\,|E\rangle$, since they only aquire a phase $\exp(-iEt/\hbar)$, and statistical mixtures of these pure states. Because the energy is preserved in a closed system, these states are also referred to as enery eigenstates. Correspondingly, the lowest energy is called the ground state energy and the respective eigenstate is called the ground state of the system. For a time-independent Hamiltonian the unitary transformation is given by

$$U(t_1, t_2) = \exp\left[-\frac{H(t_2 - t_1)}{\hbar}\right]. \tag{2.6}$$

In this thesis, however, we will usually study unitary evolutions without considering the underlying Hamiltonian as we will abstract the concrete physical system away from the information theoretical scenario.

When we talk about the state of a classical system, we usually mean a collection of its properties such as its position in phase space. In principle, classical physics allows us to measure each of these properties independently without disturbing the system, or

at least, that they are all well-defined at the same time. However, in quantum physics, the fundamental description of the state of a system is the state vector or density operator. To obtain physical properties such as position or momentum, we have to actively measure the system. In contrast to a passive classical measurement process, this will change the state of the system and will be different from the unitary time evolution described above as the measurement apparatus becomes part of the system for the time of the measurement and hence, the system is not closed anymore. The third postulate delineates how measurements are described in quantum mechanics. An observable is a Hermitian operator with spectral decomposition

$$A = \sum_j a_j \left|a_j\right\rangle \left\langle a_j\right| . \tag{2.7}$$

The eigenvalues are the possible outcomes that can be obtained through a measurement and the probability of obtaining the outcome a_j by measuring the system in the state $|\psi\rangle$ is given by $p_j = |\langle a_j|\psi\rangle|^2$. For a mixed state ρ, the outcome probabilities are hence given by $p_j = \langle a_j|\rho|a_j\rangle$. These so-called von Neumann measurements are, however, not the most general way of obtaining information from a quantum system. Instead, we can first add an ancilla system in a well-defined state, let the joint system evolve unitarily and afterwards measure the state of the ancilla system. Via this process which is specified in Naimark's dilation theorem [23], the possible measurements in quantum mechanics are given by positive operator-valued measures or POVMs. A POVM is a collection of so-called measurement operators $\{M_j\}$, that satisfy the normalization $\sum_j M_j^\dagger M_j = \mathbb{1}$. The operators $E_j = M_j^\dagger M_j$ are called the effects of the measurement. The probability of obtaining measurement outcome j is then given by $p_j = \mathrm{Tr}\,\rho E_j$. The von Neumann measurements characterize the important subclass of projection-valued measures or PVMs, where all effects are projectors.

After the measurement process, the state of the examined quantum system has changed. The post-measurement state, however, depends not only on the effects but on the actual physical implementation of the measurement specified by the measurement operators M_j. For a quantum system in state ρ, it is given by

$$\rho_j = \frac{1}{p_j} M_j \rho M_j^\dagger, \tag{2.8}$$

if outcome j is obtained. In case, one is indifferent to the post-measurement state, it is enough to consider the measurement effects instead of the measurement operators. This is the approach taken throughout this thesis. The intriguing prediction of a discrete set of outcomes often drastically differs from predictions in classical physics. For instance, the Stern-Gerlach experiment [24] reveals the quantization of the spatial

orientation of angular momentum of silver atoms. Similarly, polarization experiments with photons show the same effect and, by measuring the change of light intensity between multiple polarizers, the change of the state through the measurement can be observed.

Finally, the fourth postulate tells us how to describe a composite physical system consisting of multiple smaller systems or particles. In many classic books about quantum mechanics, this feature is hidden somewhere in the mathematical framework. If composite systems are considered, they usually refer to indistinguishable particles, namely fermions and bosons, and their statistics [25, 26]. In quantum information theory, however, also composite systems of particles that can, for example, be reliably distinguished by their spatial distribution is essential. The state space of m systems in Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_m$ is given by its tensor product $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_m$. This means that for two systems, one being in state ρ, the other in state σ, the composite system is in state $\rho \otimes \sigma$, or in the computational basis, $\rho = \sum_{i,j} \rho_{ij} |i\rangle \langle j|$, $\sigma = \sum_{i,j} \sigma_{ij} |i\rangle \langle j|$ and

$$\rho \otimes \sigma = \sum_{i,j,k,l} \rho_{ij}\sigma_{kl} |i\rangle \langle j| \otimes |k\rangle \langle l| = \sum_{i,j,k,l} \rho_{ij}\sigma_{kl} |ik\rangle \langle jl| . \tag{2.9}$$

Correspondingly, for pure states $|\psi\rangle = \sum_j \psi_j |j\rangle$ and $|\phi\rangle = \sum_j \phi_j |j\rangle$, the composite system is in the state $|\psi\rangle \otimes |\phi\rangle = \sum_{i,j} \psi_i \phi_j |ij\rangle$. However, as we will see later when we discuss quantum entanglement, not every state of the composite system can be written as the tensor product of local states.

Interestingly, mixed states can be viewed as parts of a pure state on a composite quantum system. For a state in its spectral decomposition

$$\rho = \sum_{j=1}^{k} p_j |\psi_j\rangle \langle \psi_j| , \tag{2.10}$$

where k is the rank of the density matrix, a possible purification on a composite system with an added k-dimensional ancilla, is given by

$$|\phi\rangle = \sum_j \sqrt{p_j} |\psi_j\rangle |j\rangle . \tag{2.11}$$

Indeed, there are infinitely many different purifications, however, the rank of ρ determines the minimal dimension of the ancilla system. The correct operation to obtain the original state on part of the system is the partial trace $\rho = \mathrm{Tr}_2 |\phi\rangle \langle \phi|$, which as a linear operator can be defined on the computational basis by

$$\mathrm{Tr}_2 (|i\rangle \langle j| \otimes |m\rangle \langle l|) = \delta_{ml} |i\rangle \langle j| , \tag{2.12}$$

where the subscript indicates the part of the system that is traced out and δ_{ml} is the Kronecker delta which is 1 if $m = l$ and 0 otherwise.

To sum up, the postulates of quantum mechanics tell us how to describe the fundamental state of a physical system, how it evolves with time, what measurements we can do to learn about the state and how they change the system, and how we can describe composite systems consisting of multiple particles.

2.2.2 The qubit

The smallest nontrivial quantum system is two-dimensional, i.e., a single qubit. For instance, this can be the spin of an electron, the polarization of a photon, or simply two energy levels of an ion. Qubits are the natural quantum generalization of classical bits and hence, they usually serve as the fundamental building block of quantum computers. Higher-dimensional systems are usually referred to as qudits, where the d indicates the dimension. Sometimes, we also use qutrits, ququarts, or quhex to describe three-, four-, or six-dimensional quantum systems. In the case of a single qubit, the computational basis consists only of the vectors $|0\rangle$ and $|1\rangle$, often also referred to as spin-up and spin-down. Because a global phase is irrelevant physically, normalized, pure single-qubit states can be parameterized using two real parameters as

$$|\psi\rangle = \cos\theta\,|0\rangle + e^{i\varphi}\sin\theta\,|1\rangle\,. \tag{2.13}$$

We can interpret the angles θ and φ as the polar and azimuthal angle in spherical coordinates, respectively. Then, the pure states cover the surface of the unit sphere.

To see that we can map the mixed states to the interior of the unit sphere, we introduce the so-called Pauli matrices:

$$\begin{aligned} \sigma_0 = \mathbb{1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 = Y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.14}$$

The eigenvectors or eigenstates of Z are exactly the computational basis states, the eigenstates of X are often denoted as $|+\rangle$ and $|-\rangle$ and those of Y by $|i_+\rangle$ and $|i_-\rangle$. The Hermitian and unitary Pauli matrices form an orthogonal basis for the Hermitian 2×2-matrices, i.e., $\operatorname{Tr}\sigma_i\sigma_j = 0$ if $i \neq j$. Such a basis does not exist in higher dimensions as one needs to restrict the requirements to either Hermitian or unitary matrices. A

general mixed state can be represented as

$$\rho = \frac{1}{2}\left(\mathbb{1} + \lambda_x X + \lambda_y Y + \lambda_z Z\right) = \frac{1}{2}\left(\mathbb{1} + \boldsymbol{\lambda} \cdot \boldsymbol{\sigma}\right), \tag{2.15}$$

where positive semidefiniteness requires $|\boldsymbol{\lambda}| \leq 1$ and $|\boldsymbol{\lambda}| = 1$ if, and only if, the state is pure. Hence, we can describe the set of single-qubit states geometrically as the three-dimensional unit ball. In quantum information theory, it is also known as Bloch sphere or Bloch ball.

Since qubits form the smallest and simplest quantum system, they are often of special interest in research and serve as a testbed for more difficult systems. However, many properties are indeed exclusive to qubits and hence, investigating higher-dimensional systems can lead to interesting observations.

2.3 Quantum channels

As we have seen, the class of physically implementable measurements grows significantly if we allow to utilize an ancilla system and a controlled joint evolution with the system of interest, namely, instead of PVMs or von-Neumann measurements we have access to POVMs. Similarly, the time evolution of a closed system is heavily restricted — it is unitary — compared to the time evolution of open systems, i.e., quantum systems that transform as parts of a larger closed system. The general time evolution is a linear map from quantum states of one Hilbert space to quantum states of another Hilbert space $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$, where $L(\mathcal{H})$ denotes the set of density matrices on the Hilbert space $\mathcal{H}$. However, not every such map can be realized in quantum mechanics. The reason is that if only part of a quantum system transforms according to $\mathcal{M}$, the joint system must remain in a valid quantum state. This means that for

$$\sigma = \left(\mathrm{Id}_d \otimes \mathcal{M}\right)(\rho), \tag{2.16}$$

where $\mathrm{Id}_d : \mathcal{H}_d \to \mathcal{H}_d$ is the d-dimensional identity map that maps d-dimensional quantum states to themselves, it must hold that σ is a physical quantum state, i.e., $\sigma \geq 0$ and $\mathrm{Tr}\,\sigma = 1$, for any d and any state ρ. In other words, $\mathcal{M}$ is a trace-preserving and so-called completely positive map, also known as a quantum channel.

Definition 2.1. A quantum channel $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$ is a linear map that is

(i) completely positive, i.e., $\left(\mathrm{Id}_d \otimes \mathcal{M}\right)(\rho) \geq 0$ for any positive semidefinite operator $\rho \in L(\mathcal{H}_d \otimes \mathcal{H}_A)$, $\rho \geq 0$.

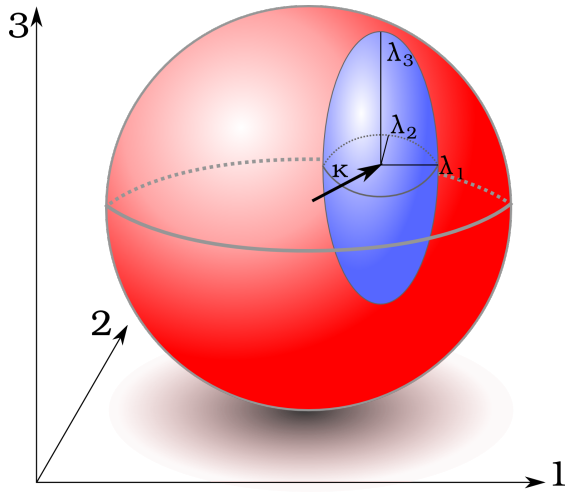


FIGURE 2.1: [29] The image of the Bloch sphere (red) of single-qubit maps is an ellipsoid (blue) with semi-axes λ_i, displaced by $\vec{\kappa}$.

(ii) trace-preserving, i.e., $\mathrm{Tr}\,\mathcal{M}(\rho) = \mathrm{Tr}\,\rho$ for any positive semidefinite operator $\rho \geq 0$.

This is why quantum channels are also known as CPTP maps.

An excellent introduction to quantum channels can be found in Ref. [27]. Indeed, similarly to the relation between POVMs and PVMs, any quantum channel as defined above can be physically implemented using an ancilla system and unitary time evolution. More precisely, for any quantum channel $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$, there exists a unitary operator U such that

$$\mathcal{M}(\rho) = \mathrm{Tr}_E \left[U \left(\rho \otimes |0\rangle \langle 0| \right) U^\dagger \right], \tag{2.17}$$

where the dimension of the environment ancilla Hilbert space is $d_E = d_A d_B$ with d_A and d_B being the dimensions of the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, due to Stinespring's dilation theorem [28].

2.3.1 Single-qubit channels

The action of single-qubit channels can be well understood in the Bloch picture. Let the Bloch decomposition of a qubit state be $\rho = \frac{1}{2}(\mathbb{1} + \vec{v} \cdot \vec{\sigma})$, where $\vec{v} \in \mathbb{R}^3$ is required to have length equal or smaller than 1 in order for ρ to be positive semi-definite, and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ with σ_i being the Pauli matrices. Then, any qubit-qubit quantum channel corresponds to an affine transformation $\vec{v} \mapsto \Lambda\vec{v} + \vec{\kappa}$ with a real matrix Λ and a displacement vector $\vec{\kappa}$ [20], where some restrictions on Λ and $\vec{\kappa}$ apply to ensure complete positivity. Thus, the image of any single-qubit channel $\mathcal{M}$ is given by an

ellipsoid in the Bloch sphere, where the semi-axes are given by the singular values of Λ and the ellipsoid is translated by $\vec{\kappa}$. The surface is given by the image of the pure states under $\mathcal{M}$ because of linearity (see Fig. 2.1).

Quantum channels are, for instance, used to model noise in quantum systems. In the following, we consider important examples of single-qubit channels, i.e., $\mathcal{H}_A$ and $\mathcal{H}_B$ are both two-dimensional. First, the depolarizing channel $\mathcal{D}_p$ is defined by

$$\mathcal{D}_p(\rho) = p\frac{\mathbb{1}}{2} + (1-p)\rho, \tag{2.18}$$

where $0 \leq p \leq 1$ characterizes the strength of the noise. In the Bloch picture, the ball of quantum states is mapped to a ball with smaller radius but the same origin. The states become more mixed, the purity changes as

$$\gamma(\mathcal{D}_p(\rho)) = \frac{p^2}{2} + p(1-p) + (1-p)^2\gamma(\rho) \leq \gamma(\rho). \tag{2.19}$$

Applying the depolarizing channel is also described by mixing with white noise or the so-called maximally mixed state $\mathbb{1}/2$. This channel is also defined for higher-dimensional Hilbert spaces, however, the maximally mixed state is given by $\mathbb{1}/d$ due to normalization. Second, the bit flip channel $\mathcal{B}_p$ is defined by

$$\mathcal{B}_p(\rho) = p\rho + (1-p)X\rho X, \tag{2.20}$$

where again $0 \leq p \leq 1$. A classical bit has only two possible states 0 and 1 and hence, the only error source is the probabilistic flip of the bit, interchanging the states. The bit flip channel emulates this behavior for a quantum bit, i.e., the states $|0\rangle$ and $|1\rangle$ are interchanged with probability $1-p$. The image of the state space under this operation is a deformed sphere, an ellipsoid generated by the contraction of the sphere along the y- and the z-axis. The related phase flip channel $\mathcal{P}_p$ and the bit-phase flip channel are defined similarly, however, using the other Pauli matrices Z and Y instead of X, respectively. They illustrate that errors on quantum computers are more subtle and multifaceted compared to errors on a classical computer. Finally, the amplitude damping channel $\mathcal{A}_p$ is defined by

$$\mathcal{A}_p(\rho) = E_0\rho E_0^\dagger + E_1\rho E_1^\dagger, \tag{2.21}$$

where the operation elements E_0 and E_1 are given by

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}. \tag{2.22}$$

This channel is important to describe energy dissipation or photon loss as it favors the ground or no photon state $|0\rangle$ over the excited or single photon state $|1\rangle$.

2.3.2 Choi-Jamiołkowski isomorphism

Complete positivity of a quantum channel is usually difficult to verify directly because positivity has to be checked for any dimension of the ancilla system. The Choi-Jamiołkowski isomorphism [30–32] provides a technique to avoid this difficulty. It assigns to every CPTP map a quantum state, and to certain quantum states a corresponding quantum channel.

Definition 2.2. For a channel $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$, the Choi state is defined as

$$\eta_{\mathcal{M}} = (\mathrm{Id} \otimes \mathcal{M}) \left(|\phi^+\rangle \langle \phi^+| \right), \tag{2.23}$$

where $\mathrm{Id} : L(\mathcal{H}_A) \to L(\mathcal{H}_A)$, i.e., $\eta_{\mathcal{M}} \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$, and

$$|\phi^+\rangle = \frac{1}{\sqrt{d_A}} \sum_j |j\rangle |j\rangle \tag{2.24}$$

is the so-called maximally entangled state.

Choi's theorem on completely positive maps proves that the positivity of $\eta_{\mathcal{M}}$ is equivalent to the complete positivity of $\mathcal{M}$. Hence, it enables a simple characterization of quantum channels utilizing the positive semidefiniteness of matrices.

On the other hand, given a Choi state $\eta_{\mathcal{M}}$, the action of the corresponding channel can be calculated as

$$\mathcal{M}(\rho) = d_A \, \mathrm{Tr}_1 \left[\eta_{\mathcal{M}} \left(\rho^T \otimes \mathbb{1}_{d_B} \right) \right], \tag{2.25}$$

where ρ^T is the transpose of ρ. From this equation, it is clear that $\mathrm{Tr}\, \mathcal{M}(\rho) = 1$ for all states ρ if and only if $\mathrm{Tr}_B \, \eta_{\mathcal{M}} = \mathbb{1}/d_A$, i.e., the partial state on system A of the joint state $\eta_{\mathcal{M}}$ is maximally mixed. This condition characterizes the set of Choi states. Furthermore, a quantum channel $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$ is called unital if it maps the maximally mixed state of $\mathcal{H}_A$ to the maximally mixed state of $\mathcal{H}_B$ or $\mathcal{M}(\mathbb{1}/d_A) = \mathbb{1}/d_B$. For the Choi state $\eta_{\mathcal{M}}$, we have that $\mathrm{Tr}_A \, \eta_{\mathcal{M}} = 1$ if, and only if, the corresponding channel $\mathcal{M}$ is unital.

While we did define the bit-flip and the amplitude damping channel in a way that can be applied to general matrices as well, this is not the case for our definition of the depolarizing channel since it is not trace-preserving for operators that are not

normalized. Hence, it is harder to see how to apply it to part of the state

$$|\phi^+\rangle\langle\phi^+| = \frac{1}{2}\sum_{i,j}|ii\rangle\langle jj| \quad = \frac{1}{2}\sum_{i,j}|i\rangle\langle j|\otimes|i\rangle\langle j| \tag{2.26}$$

because $\mathrm{Tr}\,|i\rangle\langle j| = \delta_{ij}$. However, with the, for quantum states equivalent, definition

$$\mathcal{D}_p(\rho) = p\,(\mathrm{Tr}\,\rho)\,\frac{\mathbb{1}}{2} + (1-p)\rho, \tag{2.27}$$

we obtain $\eta_{\mathcal{D}} = p\mathbb{1}\otimes\mathbb{1}/4 + (1-p)\,|\phi^+\rangle\langle\phi^+|$ which as the mixture of quantum states is easily seen to be positive semidefinite and hence, complete positivity for $\mathcal{D}_p$ follows.

2.3.3 Kraus representation

The way we defined the depolarizing channel, albeit very intuitive, is quite different from the way we defined the bit-flip and the amplitude damping channel. As we saw with the Choi-Jamiołkowski isomorphism, this can complicate computations and further analysis. Fortunately, there is a standard form for quantum channels called the Kraus representation [27, 33]. The action of any CPTP map $\mathcal{M}:\mathcal{H}_A\to\mathcal{H}_B$ can be written as

$$\mathcal{M}(\rho) = \sum_j K_j\rho K_j^\dagger, \tag{2.28}$$

where the K_j are called Kraus operators. Because $\mathcal{M}$ is trace-preserving, we have that $\sum_j K_j^\dagger K_j = \mathbb{1}$. Complete positivity follows immediately by construction since

$$\langle\psi|\left[(\mathrm{Id}\otimes\mathcal{M})\left(|\phi^+\rangle\langle\phi^+|\right)\right]|\psi\rangle = \sum_j\left[\langle\psi|\left(\mathbb{1}\otimes K_j^\dagger\right)^\dagger\right]|\phi^+\rangle\langle\phi^+|\left[\left(\mathbb{1}\otimes K_j^\dagger\right)|\psi\rangle\right] \geq 0, \tag{2.29}$$

where the inequality follows from the positive semidefiniteness of $|\phi^+\rangle\langle\phi^+|$ as a quantum state. Furthermore, if $\mathcal{M}$ is a unital channel, then it also holds that $\sum_j K_j K_j^\dagger = \mathbb{1}$.

Although the Kraus representation is not unique, it is straightforward to find a decomposition into a minimal number of Kraus operators with the property that they are orthogonal, i.e., $\mathrm{Tr}\,K_i^\dagger K_j \propto \delta_{ij}$. This is done by considering the (unnormalized) spectral decomposition of the corresponding Choi state $\eta_{\mathcal{M}} = \sum_j |\psi_j\rangle\langle\psi_j|$. From the definition of $\eta_{\mathcal{M}}$ in Eq. 2.23, we have that

$$\sum_j |\psi_j\rangle\langle\psi_j| = \sum_j(\mathbb{1}\otimes K_j)\,|\phi^+\rangle\langle\phi^+|\,(\mathbb{1}\otimes K_j^\dagger), \tag{2.30}$$

where we can identify $|\psi_j\rangle = (\mathbb{1}\otimes K_j)\,|\phi^+\rangle$ which is always possible, as we will see in Section 2.5, for states $|\psi_j\rangle$ with $\mathrm{Tr}_1\,|\psi_j\rangle\langle\psi_j| = \mathbb{1}/d$ which is the case for Choi states.

In the case of the depolarizing channel for single-qubit states, we find such a Kraus decomposition as

$$\mathcal{D}_p(\rho) = \sqrt{1-\frac{3p}{4}}\mathbb{1}\rho\sqrt{1-\frac{3p}{4}}\mathbb{1} + \frac{\sqrt{p}}{2}X\rho\frac{\sqrt{p}}{2}X + \frac{\sqrt{p}}{2}Y\rho\frac{\sqrt{p}}{2}Y + \frac{\sqrt{p}}{2}Z\rho\frac{\sqrt{p}}{2}Z. \quad (2.31)$$

2.4 Coherence

Coherence is one of the distinguishing features of quantum physics compared to classical physics. In experiments, or generically in physical systems, there is often a distinct basis given, e.g., by the eigenstates of the underlying Hamiltonian or another observable, that is protected by conservation laws or superselection rules. This basis is also known as the classical basis of the system. Then, it is usually much easier to prepare, manipulate, and measure these basis states and mixtures thereof compared to superpositions. For instance, consider a single-qubit system with distinguished basis $\{|0\rangle, |1\rangle\}$. The mixture

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \quad (2.32)$$

is qualitatively very different from the superposition of the basis states

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, \quad (2.33)$$

even though, when measured in the preferred basis, they both yield outcome 0 and 1 with probability 1/2 each. Measured in a different basis, e.g., the eigenbasis of the Pauli operator X, $\{|+\rangle, |-\rangle\}$, however, the state ρ yields outcomes $+$ and $-$ with equal probability while the state $|+\rangle$ deterministically yields outcome $+$. Such superpositions are important for quantum information protocols such as quantum key distribution [13]. States that are mere mixtures of distinguished basis states are called incoherent, whereas states that contain superpositions of these basis states are called coherent. Hence, a coherent state requires superposition to describe it. That is why superposition and coherence are sometimes used interchangeably.

To quantify the coherence, a resource theory of coherence has been developed that appropriately characterizes sensible coherence measures [34–36]. Quantum resource theories naturally share a common structure [37, 38] that can, and has been, applied to different resources such as coherence, entanglement, and Bell nonlocality [39–43]. This structure consists of, first, free states, i.e., states that are for example easy to prepare in an experiment and hence, are regarded as resourceless. In the context of coherence, the free states are obviously the incoherent states. The set of incoherent states is usually denoted as $\mathcal{I}$. Second, there are free operations that are also easy to

implement in practice and cannot transform a free state into a resourceful state. For a resource theory of coherence, the choice of free operations is not unique, however, they are usually given by the so-called incoherent operations defined as quantum channels $\Phi_{\mathrm{ICPTP}}(\rho) = \sum_j K_j \rho K_j^\dagger$ with Kraus operators satisfying

$$K_j \mathcal{I} K_j^\dagger \subset \mathcal{I}, \tag{2.34}$$

for all j, i.e., all the Kraus operators map incoherent states to (unnormalized) incoherent states.

The advantage of quantum resource theories is that they provide natural criteria for sensible resource measures that quantify the amount of resource present in a quantum state. The criteria for a coherence measure $C : L(\mathcal{H}) \to \mathbb{R}$ are given by

C1 $C(\rho) = 0$ for all $\rho \in \mathcal{I}$.

C2.a Monotonicity under incoherent operations: $C(\Phi_{\mathrm{ICPTP}}(\rho)) \leq C(\rho)$.

C2.b Monotonicity under selective measurements on average: $\sum_j p_j C(\rho_j) \leq C(\rho)$, where $p_j = \mathrm{Tr}\, K_j \rho K_j^\dagger$, $\rho_j = K_j \rho K_j^\dagger / p_j$ and the Kraus operators form an incoherent channel $\Phi_{\mathrm{ICPTP}}(\rho) = \sum_j K_j \rho K_j^\dagger$.

C3 Convexity: $C(\sum_j p_j \rho_j) \leq \sum_j p_j C(\rho_j)$ for any probability distribution $\{p_j\}$ and quantum states ρ_j.

Resource measures are often also called resource monotones. Coherence monotones satisfying the above criteria are, for instance, the robustness of coherence [44], the l_1-norm of coherence [45], and the relative entropy measure [46]. The measures differ from each other in their physical and operational interpretation.

The the l_1-norm of coherence intuitively quantifies the amount of superposition using the off-diagonal elements of the density matrix.

Definition 2.3. The l_1-norm of coherence C_{l_1} of a quantum state $\rho \in L(\mathcal{H})$ is defined by

$$C_{l_1}(\rho) = \sum_{i \neq j} |\rho_{ij}|. \tag{2.35}$$

Note that the analogous l_2-norm of coherence is indeed not a coherence monotone [35]. While easy to compute, the l_1-norm of coherence does not provide a direct physical interpretation in terms of an underlying quantum information protocol whose performance could be measured via this coherence measure. The robustness of coherence, on the other hand, provides such an operational measure of quantum coherence [47].

Definition 2.4. The robustness of coherence C_R of a quantum state $\rho \in L(\mathcal{H})$ is defined by

$$C_R(\rho) = \min_{\sigma \in L(\mathcal{H})} \left\{ t \geq 0 \middle| \frac{\rho + t\sigma}{1+t} \in \mathcal{I} \right\}. \tag{2.36}$$

Note that for a single qubit, i.e., a two-dimensional quantum system, we have that $C_R(\rho) = C_{l_1}(\rho)$. Computing the robustness of coherence can be done efficiently using a semidefinite program as described in Section 2.8 [47]. The operational meaning associated to this measure becomes apparent by examining the following scenario [44]. In a phase discrimination game, someone, let us call her Alice, prepares a quantum state $\rho \in L(\mathcal{H})$ which subsequently undergoes the transformation

$$\rho \mapsto \rho_\varphi = U_\varphi \rho U_\varphi^\dagger, \tag{2.37}$$

where $U_\varphi = \exp(iH\varphi)$. Without loss of generality, we consider a Hamiltonian H that has evenly spaced energy levels, i.e., $H = \sum_n n \left|n\right\rangle \left\langle n\right|$. Suppose that there is a finite set of angles $\varphi_j \in \mathbb{R}$ that are imprinted on the quantum state with probability p_j. Alice, however, is unaware of which angle φ_j has actually been encoded and wants to guess it in an optimal way by measuring the transformed state ρ_φ. The most general protocol is to measure a POVM with effects E_j and returning the outcome of the measurement as the guess for φ. Then, for such a game $\Theta = \{(p_j, \varphi_j)\}$, the optimal success probability is given by

$$p_\Theta^{\text{succ}}(\rho) = \max_{\{E_j\}} \sum_j p_j \operatorname{Tr} E_j \rho_\varphi. \tag{2.38}$$

The Hamiltonian provides a natural choice for the classical basis. Then, incoherent states are indeed invariant under U_φ and hence, the measurement cannot reveal any information about the underlying φ_j. Thus, the best Alice can do is to guess the most likely φ_j, yielding a success probability of $\max_j p_j$. It turns out that the maximal quantum advantage that can be reached using a coherent input state for any game is given by the robustness of coherence [44]

$$\max_\Theta \frac{p_\Theta^{\text{succ}}(\rho)}{\max_j p_j} = 1 + C_R(\rho). \tag{2.39}$$

Thus, this measure has a clear physical interpretation benefiting its usefulness.

In some, but not all, resource theories, there are states that unambiguously contain a maximal amount of that resource. This is the case if they can be compared with any other state under the monotonicity conditions, i.e., using free operations, it is possible to reach any other state from these states. The resource theory of coherence contains

a set of such states given by [35, 45]

$$|\Psi^{\boldsymbol{\alpha}}\rangle = \frac{1}{\sqrt{D}} \sum_{j=0} \exp(i\alpha_j) \, |b_j\rangle \,, \tag{2.40}$$

where $\{|b_j\rangle\}_{j=0}^{D-1}$ is the D-dimensional classical basis. Thus, the maximally coherent states are those that are an equal superposition of all basis states, albeit they might contain different phases for each basis state.

2.5 Entanglement

While coherence is a feature of quantum mechanics that is assigned to a single system, quantum entanglement describes spatial correlations between quantum systems that are not allowed in classical physics. Also, entanglement is in contrast to coherence independent of the choice of a special basis. The physical spatial separation of subsystems provides a natural split of the joint quantum state into reduced substates. It is the most fundamental concept of quantum correlations between spatially separated particles that is essential for other correlations such as EPR steering and Bell nonlocality which are impossible without entanglement. A good introduction to entanglement theory can be found in Ref. [16]. Entanglement between two particles is already quite remarkable as it plays a crucial role in many quantum information and quantum computation protocols, however, entanglement between multiple particles provides a much richer structure that can lead to surprising mechanisms, for instance, the distribution of entanglement using separable states [48]. Since it is such an interesting and important concept, detecting entanglement and measuring its amount in experiments is vital. In the following, we present various detection methods and a resource theory of entanglement that uses physically motivated free operations. Furthermore, we discuss so-called entanglement-breaking channels, which remove any entanglement present in a quantum state.

2.5.1 Entanglement between two particles

As we have seen, the state space of two particles is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of the individual particles' state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. If the states $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$ of the two systems are uncorrelated, the joint state is also given by their tensor product $|\psi_A\rangle \otimes |\psi_B\rangle$, however, there exist states in the joint state space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ that

cannot be written as a tensor product. Most prominently, for the two-qubit state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle \right), \tag{2.41}$$

which is known as the Bell state and is the maximally entangled state of two qubits, it is easy to see that it cannot be written as the tensor product of uncorrelated states, i.e., it is not separable and hence, it is entangled. The following definition introduces bipartite entanglement and separability formally for pure states.

Definition 2.5. A state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is called separable if there exist states $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$ such that

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle . \tag{2.42}$$

It is then also called a product state. Otherwise, $|\psi\rangle$ is called entangled.

For a general state

$$|\psi\rangle = \sum_{i,j} \psi_{ij} |i\rangle |j\rangle , \tag{2.43}$$

we can determine whether or not it is entangled using the so-called Schmidt decomposition [19]. The coefficients ψ_{ij} can be interpreted as a matrix with singular value decomposition $\psi = udv$, where u and v are unitary matrices and d is a diagonal matrix. Then, we have

$$\begin{aligned} |\psi\rangle &= \sum_{i,j,k} u_{ik} d_{kk} v_{kj} |i\rangle |j\rangle \\ &= \sum_k d_{kk} \left(\sum_i u_{ik} |i\rangle \right) \left(\sum_j v_{kj} |j\rangle \right) \\ &= \sum_k \lambda_k |k_A\rangle |k_B\rangle , \end{aligned} \tag{2.44}$$

where we have introduced new bases $|k_A\rangle = \sum_i u_{ik} |i\rangle$ and $|k_B\rangle = \sum_j v_{kj} |j\rangle$ — orthogonality follows from the unitarity of u and v — and identified the so-called Schmidt coefficients $\lambda_k = d_{kk} \geq 0$. The number of nonzero Schmidt coefficients is known as the Schmidt number and measures, in some sense, the amount of entanglement present. In particular, the state $|\psi\rangle$ is a product state if, and only if, its Schmidt number is 1. A very useful property of the Schmidt number is its invariance under local unitary transformations as can be seen from the procedure of the singular value decomposition. Thus, the entanglement of pure states is easy to detect when the state is known.

For mixed states, however, the situation is a lot more involved.

Definition 2.6. A state $\rho \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called separable if there exist states $|\psi_j\rangle \in \mathcal{H}_A$ and $|\phi_j\rangle \in \mathcal{H}_B$ as well as a probability distribution $\{p_j\}$ such that ρ can be written as

$$\rho = \sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| \otimes \left|\phi_j\right\rangle \left\langle\phi_j\right| . \tag{2.45}$$

Otherwise, it is called entangled. Furthermore, ρ is called a product state if there exist states $\rho_A \in L(\mathcal{H}_A)$ and $\rho_B \in L(\mathcal{H}_B)$ such that ρ can be written as

$$\rho = \rho_A \otimes \rho_B. \tag{2.46}$$

While product states are completely uncorrelated, separable states contain classical correlations but no spatial quantum correlations. For a mixed state ρ, there are infinitely many possibilities to write it as an ensemble $\{p_j, |\psi_j\rangle\}$, i.e., as a mixture of pure states

$$\rho = \sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| . \tag{2.47}$$

The Schrödinger-HJW theorem [49–51] tells us how two different ensembles $\{p_j, |\psi_j\rangle\}$ and $\{q_k, |\phi_k\rangle\}$ that represent the same quantum state, and hence, are physically indistinguishable, are connected. Namely, the two ensembles represent the same state ρ, i.e.,

$$\rho = \sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| = \sum_k q_k \left|\phi_k\right\rangle \left\langle\phi_k\right| , \tag{2.48}$$

if, and only if, there exists a unitary matrix u such that

$$\sqrt{p_j} \left|\psi_j\right\rangle = \sum_k u_{jk} \sqrt{q_k} \left|\phi_k\right\rangle , \tag{2.49}$$

where the smaller ensemble is padded with zero-probability states such that u is a square matrix. Hence, apart from the spectral decomposition, all other possible ensembles have to be checked to ensure that the underlying state is indeed entangled. This is the reason why entanglement detection is such a vast field of research and many criteria have been developed to detect it [52].

2.5.2 Multipartite entanglement

As already mentioned, the entanglement structure in multipartite systems is far more complex compared to bipartite systems. With three (or more) particles, there are different notions of entanglement. First, there are states $|\psi_{A|B|C}\rangle$ that are fully separable, i.e.,

$$\left|\psi_{A|B|C}\right\rangle = \left|\psi_A\right\rangle \otimes \left|\psi_B\right\rangle \otimes \left|\psi_C\right\rangle . \tag{2.50}$$

Second, one can consider a fixed bipartition of the three particles and the bipartite entanglement between them, i.e., between particle A and particles BC and so on, often written as $A|BC$, $B|AC$, and $C|AB$. A so-called biseparable state $|\psi_{A|BC}\rangle$ with respect to the bipartition $A|BC$, for instance, can be written as

$$|\psi_{A|BC}\rangle = |\psi_A\rangle \otimes |\psi_{AB}\rangle\,, \tag{2.51}$$

where $|\psi_{AB}\rangle$ might be entangled. Finally, $|\psi\rangle$ is called genuine tripartite entangled if it cannot be written in such a biseparable form. Three-qubit examples for genuine tripartite entangled states are the GHZ [53, 54] and the W state [55, 56]

$$|\mathrm{GHZ}_3\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right), \tag{2.52}$$

$$|\mathrm{W}_3\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right). \tag{2.53}$$

A mixed state $\rho_{A|B|C}$ is called fully separable if it can be written as the mixture of fully separable pure states. Furthermore, the mixture of biseparable pure states which are biseparable w.r.t. a fixed bipartition are biseparable mixed states w.r.t. this partition. For example,

$$\rho_{A|BC} = \sum_j p_j \, |\psi_j^{A|BC}\rangle \, \langle\psi_j^{A|BC}| \tag{2.54}$$

is biseparable w.r.t. the bipartition $A|BC$. Convex combinations of biseparable states w.r.t. different bipartitions are simply called biseparable. Mixed states which are not biseparable are then genuine tripartite entangled.

Extending the concept of multipartite entanglement naturally to more particles leads to even more involved structures. Importantly, the definition of fully, bi-, tri-, and so on separable mixed states yields convex subsets of the state space. This is why entanglement witnesses are a vital tool in the detection of the different forms of entanglement.

2.5.3 Entanglement witnesses

Entanglement witnesses are Hermitian operators and as such, they are observables that can be measured in experiments. This feature makes them a resource-saving tool for entanglement detection, although they do not give direct access to the amount of entanglement present in a quantum state. Formally, they are defined as follows.

Definition 2.7. An entanglement witness W is an observable that first, yields a non-negative expectation value for separable states, i.e.,

$$\langle W \rangle_{\rho_{\text{sep}}} = \operatorname{Tr} W \rho_{\text{sep}} \geq 0, \tag{2.55}$$

and second, there exists an entangled state ρ with a negative expectation value

$$\langle W \rangle_{\rho} = \operatorname{Tr} W \rho < 0. \tag{2.56}$$

States with a negative expectation value are said to be detected by the corresponding entanglement witness.

In this context, separability can mean full separability or biseparability or triseparability and so on, depending on the underlying quantum system.

Thus, in an experiment, the measurement of a negative expectation value indicates the presence of entanglement. An entanglement witness defines via $\operatorname{Tr} W\rho$ a hyperplane in the space of density matrices with Hilbert-Schmidt inner product $\langle A, B \rangle = \operatorname{Tr} AB$, and hence, also in the space of density matrices. Such a hyperplane geometrically splits the state space into two parts. On the one hand, for entanglement witnesses, separable states can only be found on one side of the corresponding hyperplane. On the other hand, for any entangled state outside the convex set of separable states, there exists a hyperplane separating the entangled state from the separable states, and this hyperplane can be defined by an entanglement witness detecting the entangled state [52].

Obviously, for a given entangled state, there are different entanglement witnesses that detect it. Entanglement witnesses are the better the more states they detect. More formally, an entanglement witness W_1 is called finer than a witness W_2 if it detects at least all the states that are detected by W_2. This is equivalent to the existence of a positive semidefinite operator $P \neq 0$ such that

$$W_2 = W_1 + P. \tag{2.57}$$

An entanglement witness is called optimal if there is no entanglement witness that is finer. However, it is very difficult to check this optimality condition in practice. A weaker property that characterizes the optimality, sometimes called weak optimality, is the condition that there exists a separable state with $\operatorname{Tr} W\rho_{\text{sep}} = 0$. This necessary condition for the optimality of W can be easily interpreted geometrically as a hyperplane that touches the set of separable states [52].

2.5.4 Positive maps and the PPT criterion

We have seen that quantum channels are completely positive, trace-preserving maps such that applying them to part of a quantum system always yields a physical state

$$(\mathrm{Id} \otimes \mathcal{M})(\rho) \geq 0. \tag{2.58}$$

Linear positive maps $\mathcal{M}$, i.e., $\mathcal{M}(\rho) \geq 0$ for all quantum states ρ, that are not completely positive in general only satisfy the above relation if the state that is transformed is separable. This is because for a separable state

$$\rho_{\mathrm{sep}} = \sum_j p_j |\psi_j\rangle\langle\psi_j| \otimes |\phi_j\rangle\langle\phi_j|, \tag{2.59}$$

due to linearity, we have that

$$(\mathrm{Id} \otimes \mathcal{M})(\rho_{\mathrm{sep}}) = \sum_j p_j |\psi_j\rangle\langle\psi_j| \otimes \mathcal{M}(|\phi_j\rangle\langle\phi_j|) \geq 0. \tag{2.60}$$

Thus, applying a positive map to part of a quantum state detects its entanglement if the resulting operator is not positive semidefinite. Indeed, for any entangled state, there exists a positive map detecting it with this procedure [57].

A very prominent example of such a positive map is the partial transpose. Writing a state ρ in the computational basis

$$\rho = \sum_{i,j,k,l} \rho_{ij,kl} |i\rangle\langle j| \otimes |k\rangle\langle l|, \tag{2.61}$$

its partial transpose on the second subsystem is given by

$$\rho^{T_B} = \sum_{i,j,k,l} \rho_{ij,kl} |i\rangle\langle j| \otimes |l\rangle\langle k|. \tag{2.62}$$

Thus, we transpose the operators on the second part of the system. Although, the partial transpose w.r.t. a different basis yields a different state, the positivity, and more generally the spectrum, of the resulting state is independent of the underlying basis. The entanglement criterion $\rho^{T_B} \geq 0$ is known as PPT criterion or Peres-Horodecki criterion, originally introduced in Ref. [58].

Especially for bipartite systems, the PPT criterion is a very powerful tool to detect entanglement. For two-qubit and qubit-qutrit states, it is indeed sufficient for entanglement detection [57], i.e., a two-qubit or qubit-qutrit state is entangled if, and only if, its partial transpose is positive semidefinite. Hence, the separability problem for these

small dimensions is solved with this simple characterization. In higher dimensions, however, this is not the case [59] and the separability problem becomes NP hard to solve [60, 61]. Then, the best we can do is to find a hierarchy that can, in principle, detect all entangled states and detects more states with each level of the hierarchy. In Section 2.8, we introduce such a hierarchy found by Doherty, Parrilo, and Spedalieri [62, 63].

2.5.5 Resource theory of entanglement

So far, we have seen various methods to detect entanglement that is present in some quantum state. To quantify the amount of entanglement, however, so-called entanglement measures are necessary. For a constructive approach to find such measures, it is useful to consider a resource theory of entanglement similar to the resource theory of coherence discussed in Section 2.4. Naturally, the separable states form the set of free states as they do not contain any spatial quantum correlations. The free operations are those that can be done easily in experiments. Considering spatial correlations, local operations are regarded as simple and thus, belong to the free operations. Furthermore, as we are interested in quantum correlations, classical communication, which, of course, can generate classical correlations, is also part of the free operations. The set of LOCC, short for local operations and classical communication, is hence physically well-motivated. However, its mathematical characterization is very involved since an unbounded number of rounds of communication might be necessary for certain entanglement transformations [64].

Analogously to the criteria for a sensible coherence measure, this resource theory of entanglement allows us to do the same for entanglement measures. The criteria for an entanglement measure $E : L(\mathcal{H}) \to \mathbb{R}$ are given by

- E1 $E(\rho) = 0$ for all separable states $\rho \in \mathrm{SEP}$.
- E2 Monotonicity under LOCC operations: $E(\mathcal{M}(\rho)) \leq E(\rho)$ for all quantum states ρ and all LOCC operations $\mathcal{M}$.
- E3 Convexity: $E(\sum_j p_j \rho_j) \leq \sum_j p_j E(\rho_j)$ for any probability distribution $\{p_j\}$ and quantum states ρ_j.

Since entanglement describes the spatial quantum correlations, there are other useful criteria that are sometimes regarded necessary for an entanglement measure [65, 66]. Namely, there are additivity and subadditivity,

E4 Additivity: $E(\rho^{\otimes n}) = nE(\rho)$, i.e., n copies of a quantum state contain n times the amount of entanglement compared to a single copy.

E5 Subadditivity: $E(\rho \otimes \sigma) \leq E(\rho) + E(\sigma)$ for all quantum states ρ and σ, i.e., two uncorrelated states do not contain more entanglement than the individual states together.

as well as a stronger version of $(E1)$,

E1' $E(\rho) = 0$ if, and only if, $\rho \in \mathrm{SEP}$.

which leads, however, to an entanglement measure that is NP-hard to compute as it decides the separability problem. A measure satisfying $(E4)$ is called extensive. There is no common agreement on what criteria are necessary for an entanglement measure or entanglement monotone, for which criterion $(E2)$ is sometimes regarded sufficient. An important feature of entanglement measures, that is sometimes added as an extra criterion, is their invariance under local unitary (LU) operations. LU invariance follows from $(E2)$ since local unitaries are reversible local operations.

As in the resource theory of coherence, there exists a (up to local unitaries) unique bipartite maximally entangled state

$$|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_j |jj\rangle \,, \tag{2.63}$$

where d is the local dimension of the subsystems, i.e., using local operations and classical communication any other quantum state can be reached [67]. Hence, it can be used for any bipartite quantum information protocol that allows for classical communication, independent of which state is actually needed, making $|\phi^+\rangle$ the unambiguously most valuable resource. This state is often referred to as the maximally entangled state. In the multipartite scenario, there is no such concept, i.e., a maximally entangled state does not exist [56]. For instance, in the case of three qubits, there are two states that are incomparable and there is no LOCC operation to reach them from another state. These are the W- and GHZ-state introduced in Eqs. (2.52,2.53).

2.5.6 Entanglement-breaking channels

In Chapter 7, we will discuss quantum memories, which are an essential part of future universal quantum computers. In essence, a quantum memory is a quantum channel

that preserves a quantum state rather well. In contrast, a so-called entanglement-breaking channel $\mathcal{M}$ is a channel that destroys the entanglement between any transformed state and other systems, i.e.,

$$(\mathrm{Id} \otimes \mathcal{M})(\rho) \in \mathrm{SEP}, \tag{2.64}$$

for all states ρ. Fortunately, there is a simple characterization of entanglement-breaking channels. Indeed, they are exactly the so-called measure-and-prepare channels [68, 69].

Definition 2.8. A measure-and-prepare channel $\mathcal{M} : L(\mathcal{H}_A) \to L(\mathcal{H}_B)$ is a CPTP map, i.e. a quantum channel, whose action can be written as

$$\mathcal{M}(\rho) = \sum_j \mathrm{Tr}(E_j \rho) \rho_j, \tag{2.65}$$

where the E_j are the effects of a POVM and the ρ_j are arbitrary quantum states.

Furthermore, a quantum channel is entanglement-breaking if, and only if, the corresponding Choi state is separable [69]. The description as measure-and-prepare channels also allows for a physical interpretation. Entanglement-breaking channels can be implemented by measuring the quantum state and, depending on the measurement outcome, prepare some other state. As such, they provide very bad quantum memories since they do not store any quantum information.

Although the PPT criterion is not a sufficient entanglement criterion in higher dimensions, there might be an interesting connection to separability in the context of entanglement-breaking channels. Quantum channels $\mathcal{M}$ that map all states to states with a positive semidefinite partial transpose are known as PPT channels, i.e.,

$$(\mathcal{M}(\rho))^{T_B} \geq 0, \tag{2.66}$$

for all states ρ. It has been conjectured that applying such a channel twice yields a measure-and-prepare channel, i.e., the composition $\mathcal{M} \circ \mathcal{M}$ is entanglement-breaking [70]. Although it has been proved for some special cases [70–72], a general treatment is still missing.

2.6 Numerical range

The numerical range was originally introduced as the Wertvorrat W of a complex square matrix A defined by [73]

$$W(A) = \{\langle\psi| A |\psi\rangle \mid \langle\psi|\psi\rangle = 1\}, \tag{2.67}$$

which forms a convex set of complex numbers, where convexity follows from the Hausdorff-Toeplitz theorem [73, 74]. Beyond this original formulation, we are interested in the joint (restricted) numerical range L of multiple observables $A_1, A_2, \ldots, A_m$, i.e. Hermitian matrices,

$$L_X(A_1, \ldots, A_m) = \left\{ (\operatorname{Tr}\rho A_1, \ldots, \operatorname{Tr}\rho A_m)^T \;\middle|\; \rho \in X \right\}, \tag{2.68}$$

where X restricts the accessible set of normalized quantum states and is, for instance, the set of all quantum states ALL of given dimension, that of PPT states or that of separable quantum states SEP. If $X = \text{ALL}$, we sometimes simply omit the subscript to improve readability. In the case of two Hermitian observables A_1, A_2 and $X = \text{ALL}$, we indeed recover the Wertvorrat $W(A)$ with $A_1 = (A + A^\dagger)/2$ and $A_2 = -i(A - A^\dagger)/2$ because $L_{\text{ALL}}(A_1, A_2)$ is clearly the convex hull of $W(A)$ due to $\mathbb{C}$ and $\mathbb{R}^2$ being isomorphic and the convexity of $W(A)$ then implies that $L_{\text{ALL}}(A_1, A_2)$ and $W(A)$ describe the same set of numbers.

In Chapter 6, we explicitly compute the separable and general numerical range for various examples. The smallest instance is a single Hermitian observable and the correponding numerical range a line connecting the minimal and maximal eigenvalue, i.e., $L(A) = [\lambda_{\min}(A), \lambda_{\max}(A)]$. In contrast, the separable numerical range is given by the optimization problem

$$L_{\text{SEP}}(A) = \left[\min_{\rho_{\text{sep}}} \operatorname{Tr} A\rho_{\text{sep}}, \max_{\rho_{\text{sep}}} \operatorname{Tr} A\rho_{\text{sep}} \right], \tag{2.69}$$

which is, in general, not easy to solve. For instance, it is not known what the minimal relative one-dimensional volume of L_{SEP} is compared to L_{ALL}.

In quantum information theory, the concept of restricted numerical range is useful to detect features that distinguish quantum from classical physics. In Ref. [75], various applications are considered. For given observables, in the case of ensemble measurements with access only to the expectation value, the numerical range provides a collection of all measurement information for different quantum states corresponding

to different points in the underlying real vector space. Even when projective measurements or POVMs are considered, the (restricted) numerical range of the effects provides full measurement information since the expectation value with an effect yields the related outcome probability. For a small number of observables, it also conveys an intuitive visualization of the measurement information that neglects inaccessible information due to a restricted measurement apparatus.

A different perspective is to view the (restricted) numerical range as an affine projection of the (restricted) state space with Hilbert-Schmidt norm. This point of view can reveal interesting geometric properties of the (restricted) state space investigating manageable small dimensions. The structure of higher-dimensional Hilbert spaces is much richer than the Bloch ball for a single qubit. For example, the boundary is partly flat and contains mixed states which can be visualized using low-dimensional projections.

It is also interesting to examine nonconvex sets of states such as pure or mixed product states, yielding what is referred to as product numerical range. This leads, however, to likewise nonconvex numerical ranges which tend to be difficult to investigate. As the convex hull of the corresponding pure state numerical ranges yields the numerical range of mixed quantum states, it is nevertheless a method to facilitate computation. For instance, for two Hermitian observables A_1 and A_2, there is a known procedure to compute the joint numerical range for $X = \mathrm{ALL}$ [76–78] which we use in Chapter 6. The so-called generating line $C(A_1, A_2)$ is defined via its dual (line) equation

$$\det\left(uA_1 + vA_2 + w\mathbb{1}\right) = 0, \tag{2.70}$$

where $ux + vy + w = 0$ is the equation of a supporting line to $C(A_1, A_2)$ in the x-y-plane, i.e., in the numerical range space. The numerical range itself is then given by the convex hull of its generating line [76, 77]. To obtain an explicit expression for the generating line, a usual procedure is to dehomogenize Eq. (2.70) by setting either $u = 1$ or $v = 1$, and replace w by $w = -ux - y$. In the latter case, an expression for the generating line is then given by the solution to the equations

$$F(u, x, y) = \det\left[uA_1 + A_2 - (ux + y)\mathbb{1}\right] = 0 \tag{2.71}$$

and $\partial F/\partial u = 0$ [78].

2.7 The marginal problem and quantum codes

In the introduction to quantum mechanics, we described how the state of a subsystem is obtained from the joint quantum state via the partial trace. This quantum channel is straightforward to compute if the joint state is known as it contains all information about the investigated physical system. Considering the maximally entangled state $|\phi^+\rangle$, the marginal states, i.e., the states of the subsystems, are maximally mixed and hence, do not contain any information even though the state of the joint system is pure. Hence, the connection between the whole and its parts of a system is especially relevant to quantum physics [18], and it manifests itself in the quantum marginal problem.

The marginal problem is the question whether or not, given a set of marginal states $\rho_{S_1}, \rho_{S_2}, \ldots, \rho_{S_m}$ on subsystems $S_1, S_2, \ldots, S_m \subset S$, there exists a joint state ρ_S such that

$$\mathrm{Tr}_{S \setminus S_j} \rho_S = \rho_{S_j}, \tag{2.72}$$

for all $j = 1, \ldots, m$ where $S \setminus S_j$ denotes the complement of S_j relative to S, i.e., a global state with the desired marginals. In the case of maximally mixed marginals $\rho_A = \mathbb{1}/d$ and $\rho_B = \mathbb{1}/d$, there are many possibilities for a global state such as (one of) the maximally entangled state(s) $\rho_{AB} = |\phi^+\rangle \langle\phi^+|$ or the maximally mixed state $\rho_{AB} = \mathbb{1}/d^2$.

Thus, apart from the existence, it is also insightful to investigate the uniqueness of such a solution. Moreover, restrictions on the global state are usually necessary to make the marginal problem appealing. In particular, forcing the joint state to be pure leads, e.g., to the concept of absolutely maximally entangled (AME) states. We consider this type of marginal problem in Chapters 3 and 4. This restriction makes the marginal problem substantially harder to tackle and previously, it has only been solved in the case of disjoint subsets S_j, i.e., nonoverlapping marginals [79].

2.7.1 AME states

An important class of quantum states whose definition is closely linked to certain marginal problems is that of absolutely maximally entangled, or short AME, states.

Definition 2.9. An AME state of n particles and local dimension d is a pure state $|\mathrm{AME}(n,d)\rangle_S$ where marginals with at most half of the particles are maximally mixed, i.e.,

$$\mathrm{Tr}_T |\mathrm{AME}(n,d)\rangle \langle \mathrm{AME}(n,d)| = \frac{\mathbb{1}}{d^{n-|T|}}, \tag{2.73}$$

for all $T \subset S$ with $|T| \leq \lfloor n/2 \rfloor$.

Due to the Schmidt decomposition it is clear that Eq. (2.73) cannot hold if T contains more than half of the particles as the corresponding marginal state is not of full rank.

From the perspective of the marginal problem, AME states are the natural multipartite extension to the bipartite maximally entangled states. In contrast to bipartite systems, however, it is not possible to reach any other quantum state via LOCC. As we have seen, the concept of a maximally entangled state breaks down already at three particles because the GHZ state, which is an AME state of three qubits, and the W state are incomparable. Furthermore, AME states do not exist for arbitrary number of particles and local dimensions. The existence problem of AME states is indeed an outstanding challenge. A regularly updated summary of known results can be found in Ref. [17].

Since AME states are maximally entangled with respect to any bipartition, they serve as essential resource for various quantum information protocols such as quantum secret sharing and open-destination quantum teleportation [80, 81].

2.7.2 Quantum error-correcting codes

In a classical computer, the only error possible is the bit-flip. If errors on individual bits occur independently, they can be corrected using a so-called repetition code. The simplest version is to copy the bit two times such that a single bit is represented by three bits and the states 0 and 1 are encoded as 000 and 111, respectively. Then, in the most probable erroneous scenario, only one of the bits is affected and the original state can be recovered by majority vote.

In a quantum computer, however, the possible errors not only comprise a continuous set, but also the no-cloning theorem [82] prevents the use of a repetition code, and syndrom measurements that detect an error might destroy the encoded quantum information. Unfortunately, errors are ubiquitous in quantum computing because of decoherence due to interaction with the environment. Surprisingly, quantum error correction is still possible [83, 84].

Mathematically, a set of possible error operators E_μ might be applied to the code state $|\psi\rangle$ through erroneous quantum computation. Single-qubit errors are, for instance, the bit-flip and phase-flip error but also any other rotation on the Bloch sphere. Although the set of error operators is usually continuous, it can be reduced to a finite set of operators for finite-dimensional quantum systems whose successful correction implies also that any linear combination is corrected effectively [19].

Typically, a K-dimensional quantum system is encoded into a K-dimensional subspace Q of n particles with local dimension d via a unitary map. The errors on the individual parties are assumed to be independent due to the spatial separation. Then, it is particularly helpful to assume that errors only affect a limited number of particles at a time which leads to the following definition.

Definition 2.10. An $((n, K, m+1))_d$ error-correcting code encodes a K-dimensional quantum system into a subspace Q of n d-dimensional particles such that all errors of the form

$$E = M_1 \otimes M_2 \otimes \cdots \otimes M_n, \tag{2.74}$$

where the number of $M_j \neq \mathbb{1}$ is at most m, are successfully corrected by the underlying recovery channel. The code is said to have minimum distance $m+1$.

The existence of an $((n, K, m+1))_d$ error-correcting code is equivalent to the existence of a K-dimensional subspace Q such that for all states $|\psi\rangle \in Q$ it holds that

$$\mathrm{Tr}_T |\psi\rangle \langle\psi| = \rho_T, \tag{2.75}$$

for all collections of subsystems T with $n - |T| \leq m$, where ρ_T is independent from $|\psi\rangle$ [85, 86]. Furthermore, a so-called pure $((n, K, m+1))_d$ error-correcting code where distict errors map any code state to orthogonal states exists if, and only if, Eq. (2.75) holds with ρ_T being the maximally mixed state for all T. Thus, the existence of an $AME(n, d)$ state is equivalent to the existence of a pure $((n, 1, \lfloor n/2 \rfloor + 1))_d$ error-correcting code.

An important necessary condition for the existence of quantum codes is the quantum Singleton bound [85, 87].

Theorem 2.11. *If there is an $((n, K, m+1))_d$ error-correcting code, then $K \leq d^{n-2m}$.*

This result limits the distance that can be reached for an otherwise fixed quantum code. Quantum error-correcting codes that meet the Singleton bound are known as maximum-distance separable (MDS) codes.

2.8 Semidefinite programming

Optimization problems are ubiquitous in science and quantum information theory is no exception. Finding optimal states, channels, and quantum information protocols is essential to advance our understanding of quantum physics. Some of the typical

problems such as the separability problem are NP hard to solve [60, 61], meaning that there is most likely no efficient, i.e. polynomial time, algorithm. At the other end of the spectrum, linear programs are optimization problems with a linear objective function subject to linear equality and inequality constraints. Indeed, any linear program can be solved in polynomial time [88].

In quantum information theory, however, a natural constraint that is omnipresent is the positive semidefiniteness of a matrix since this is one of the conditions that define a quantum state. Although such a constraint is apparently not linear, there is a larger class of optimization problems, namely the class of so-called semidefinite programs (SDPs), that allows semidefinite constraints and is still efficiently solvable [89]. A comprehensive review of the theory and applications of semidefinite programs can be found in Ref. [90].

2.8.1 Duality

Formally, any semidefinite program can be written in the standard form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c} \cdot \boldsymbol{x} \\ \text{s.t.} \quad & F_0 + \sum_j x_j F_j \geq 0, \end{aligned} \tag{2.76}$$

where $\boldsymbol{c}$ is a constant vector defining the objective function and the F_j are matrices and hence, the inequality denotes positive semidefiniteness. Linear inequalities can be implemented via blockwise diagonal matrices F_j and equalities through the combination of two inequalities. Further examples such as rewriting convex quadratic constraints in terms of linear matrix inequalities can be found in Ref. [90] The above standard form is also referred to as primal problem. Consequently, the corresponding dual problem is defined by

$$\begin{aligned} \max_{Z} \quad & -\operatorname{Tr} F_0 Z \\ \text{s.t.} \quad & \operatorname{Tr} F_j Z = c_j, \quad \text{for } j > 0, \\ & Z \geq 0. \end{aligned} \tag{2.77}$$

It is easy to see that the solution of the dual problem provides a lower bound to the solution of the primal problem because

$$-\operatorname{Tr} F_0 Z \leq \sum_j x_j \operatorname{Tr} F_j Z = \sum_j x_j c_j = \boldsymbol{c} \cdot \boldsymbol{x}, \tag{2.78}$$

where the inequality follows from the positivity of Z and the constraint of the primal problem. Likewise, this also means that the solution of the primal problem provides

an upper bound to the solution of the dual problem. The difference between the optimal solutions is known as duality gap.

Although the primal and dual SDP are not guaranteed to have the same optimal value, there are sufficient conditions implying what is called strong duality. Most prominently, we have Slater's condition for semidefinite programs [91, 92].

Theorem 2.12. *The optimal solution of the primal and dual problem in Eqs. (2.76,2.77), respectively, coincide if one of them is strictly feasible. That means that either there is a vector* $\boldsymbol{x}$ *such that* $F_0 + \sum_j x_j F_j > 0$ *is positive definite or a positive definite matrix* $Z > 0$ *satisfying* $\operatorname{Tr} F_j Z = c_j$ *for all* $j > 0$.

Note that, unless both the primal and dual problem are feasible, it can happen that one of the problems is strictly feasible but unbounded and the other is infeasible.

To solve an SDP in practice, the dual problem is computed automatically and both problems are numerically solved in parallel using an interior-point method [90]. As intermediate feasible points provide bounds for the optimal value, it can be computed up to the desired accuracy if strong duality holds, which is commonly the case in practical scenarios. Sometimes the exact solution can be obtained analytically. If one finds feasible points to the primal and dual problem with the same objective value, they must be optimal. In other words, they provide a certificate for the optimality of the solution. Even if one does not find optimal solutions, any feasible point provides an analytical upper or lower bound to the optimal solution.

2.8.2 The Doherty-Parrilo-Spedalieri hierarchy

An important application of semidefinite programming in entanglement theory is the Doherty-Parrilo-Spedalieri (DPS) hierarchy for bipartite entanglement detection [62, 63]. By definition, any separable state ρ_{sep} can be written as

$$\rho_{\text{sep}} = \sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| \otimes \left|\phi_j\right\rangle \left\langle\phi_j\right|, \tag{2.79}$$

and hence, it can be mathematically extended to a fully separable state of more parties as

$$\rho_{\text{sep}}^{(n)} = \sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right| \otimes \left|\phi_j\right\rangle \left\langle\phi_j\right|^{\otimes n} \tag{2.80}$$

for any $n > 1$, called a symmetric extension of ρ_{sep} to n copies satisfying first, that the marginal on the first two particles is given by the original state, i.e.,

$$\operatorname{Tr}_{S\setminus\{A,B\}} \rho_{\text{sep}}^{(n)} = \rho_{\text{sep}}, \tag{2.81}$$

and second, that it is is symmetric under permutations of the copies, i.e.,

$$(\mathbb{1}_A \otimes P_\sigma)\, \rho_{\text{sep}}^{(n)}\, (\mathbb{1}_A \otimes P_\sigma) = \rho_{\text{sep}}^{(n)}, \tag{2.82}$$

for any permutation σ of n elements and the corresponding permutation operator $P_\sigma = \sum_{i_1,\ldots,i_n} |\sigma(i_1,\ldots,i_n)\rangle \langle i_1,\ldots,i_n|$. It turns out that not only there exists a symmetric extension for any separable state but also for any entangled state, there is an $n > 1$ such that no extension can be found that satisfies the two conditions [62, 63].

Theorem 2.13. *A bipartite state $\rho \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable if, and only if, for every $n \in dN$ there exists a symmetric extension $\rho^{(n)}$ satisfying the conditions in Eqs. (2.81,2.82).*

These conditions can apparently be checked by an SDP for fixed n. That is why this result is commonly referred to as a hierarchy of semidefinite programs to decide the separability problem. As we noted before, the separability problem is in general NP hard to solve, and hence, the level of the hierarchy n to detect entanglement for some states can be quite large.

To improve entanglement detection via the DPS hierarchy, it is helpful to add further constraints to the SDPs that are satisfied by the natural extension of separable states in Eq. (2.81). For instance, the symmetry constraint in Eq. (2.82) can be replaced by the stronger condition that $\rho^{(n)}$ must live in the corresponding symmetric subspace, i.e., $(\mathbb{1}_A \otimes P_\sigma)\rho_{\text{sep}}^{(n)} = \rho_{\text{sep}}^{(n)}$ for any permutation σ. Furthermore, any linear or semidefinite separability criterion can be added since the natural extension is fully separable. Typically, the extension is required to have a positive partial transpose with respect to all bipartitions. Then, the DPS hierarchy can be viewed as an extension of the PPT criterion.

2.9 Classical entropy and majorization

As the term quantum information theory suggests, it is about the interplay between quantum physics and information theory. Hence, it should come as no surprise that information theoretic concepts such as entropy play an essential role. Entropy quantifies the amount of uncertainty of a random variable before an outcome is obtained and, from a different perspective, it is the average information gained when learning the variables' value [19, 20]. The Shannon entropy defined by [93]

$$S(\{p_j\}) = -\sum_j p_j \log p_j, \tag{2.83}$$

which, if the logarithm is taken with base 2, operationally yields the average number of bits needed to communicate the outcome of the underlying random variable. Note that for $p_j = 0$, the corresponding summand in the Shannon entropy in Eq. (2.83) is set to 0 in agreement with the related limit.

In the context of quantum information theory, the random variable is usually the measurement of a quantum state. For instance, Heisenberg's famous uncertainty principle [94] can be generalized in an entropic formulation. A good survey on entropic uncertainty relations can be found in Ref. [95]. A well-known entropic uncertainty relation is the following result by Maassen and Uffink [96],

$$S(A) + S(B) \geq \log \frac{1}{c}, \tag{2.84}$$

where A and B are observables with eigenvectors $|a\rangle$ and $|b\rangle$, respectively, and $S(A) = S(\{\langle a| \rho |a\rangle\})$ for some state ρ. Finally, c is the maximal overlap between any eigenvectors of A and B, i.e., $c = \max |\langle a|b\rangle|^2$. What makes this inequality so powerful is that it is independent of the quantum state ρ that is considered in the measurements, as long as it is the same for both observables.

Apart from the Shannon entropy, there are other information measures such as the Tsallis-q entropy [97]

$$S^{(q)}(\{p_j\}) = \frac{1}{q-1}\left(1 - \sum_j p_j^q\right), \tag{2.85}$$

and the Rényi-α entropy [98]

$$H_\alpha(\{p_j\}) = \frac{1}{1-\alpha} \log\left(\sum_j p_j^\alpha\right), \tag{2.86}$$

which are monotonic functions of each other for $q = \alpha$. While values related to the Tsallis entropy are easier to compute as they avoid logarithms, Rényi entropies are additive, i.e., $H_\alpha(\boldsymbol{p} \otimes \boldsymbol{q}) = H_\alpha(\boldsymbol{p}) + H_\alpha(\boldsymbol{q})$. In the limits $q, \alpha \to 1$, both converge to the Shannon entropy.

Independent from the concrete choice of information measure, the concept of majorization provides an intuitive way to characterize how chaotic a random variable is and thus, a partial order on probability distributions [20]. A probability vector $\boldsymbol{p}$ is said to be majorized by another probability vector $\boldsymbol{q}$, i.e. $\boldsymbol{p} \prec \boldsymbol{q}$, if

$$\sum_{j=1}^{k} p_j^\downarrow \leq \sum_{j=1}^{k} q_j^\downarrow, \tag{2.87}$$

for all $k = 1, 2, \ldots$, where $\boldsymbol{p}^{\downarrow}$ and $\boldsymbol{q}^{\downarrow}$ are vectors with the same components as $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively, but the components are sorted in decreasing order. In this partial order, there is a (up to permutations) unique maximum and minimum given by the flat and a deterministic distribution. A function f is called Schur concave if $f(\boldsymbol{p}) \geq f(\boldsymbol{q})$ for all $\boldsymbol{p} \prec \boldsymbol{q}$. All the introduced entropies satisfy this relation illustrating the independence of majorization from a concrete information measure.

3 Quantum-inspired hierarchy for rank-constrained optimization

Prerequisites

2.2 Quantum mechanics
2.5 Entanglement
2.7 The marginal problem and quantum codes
2.8 Semidefinite programming
2.9 Classical entropy and majorization

3.1 Introduction

The main parts of this chapter have been published as Publication (D) [99]. The mathematical theory of optimization has become a vital tool in various branches of science. This is not only due to the fact that some central problems (e.g., finding the ground state energy of a given Hamiltonian in condensed matter physics) are by definition optimization problems, where mathematical methods can directly be applied. It also turned out that other physical problems, which are not directly optimizations, can be reformulated as optimization tasks.

Recently, many efforts have been devoted to so-called semidefinite programs (SDPs), which is a class of highly tractable convex optimization problems as described in Section 2.8. In quantum information theory, they have been used to characterize quantum entanglement via the DPS hierarchy [62] and quantum correlations [100]. In condensed matter physics, SDPs are relevant for solving ground-state problems [101]. In conformal field theory, they have been employed for bootstrap problems [102]. In fact, SDPs also found widespread applications in more general topics beyond physics, examples include the Shannon capacity of graphs [103] and global polynomial optimization [104, 105].

In many cases, however, one cannot directly formulate an SDP, as some non-convex constraints remain. Well-known examples are the characterization of quantum correlations for a fixed dimension [106, 107], the determination of the faithfulness of quantum entanglement [108], the ground state energy in spin glasses [109], and compressed sensing tomography [110]. Interestingly, these non-convex optimization problems share a common structure: They can be formulated as SDPs with an extra rank constraint. Apart from these physics examples, rank-constrained optimizations are also widely-used in signal processing, model reduction, and system identification [111]. All these applications demonstrate that to achieve significant progress, it would be highly desirable to develop techniques to deal with rank constraints in SDPs.

In the following, we provide a method to deal with rank constraints based on the theory of quantum entanglement. More precisely, we prove that a large class of rank-constrained SDPs can be written as a convex optimization over separable two-party quantum states. Based on this, a complete hierarchy of SDPs can be constructed. In this way, we demonstrate that quantum information theory does not only benefit from ideas of optimization theory, but the results obtained in this field can also be used to study mathematical problems (like the Max-Cut problem) from a fresh perspective. Notably, unlike widely-used local optimization methods [112, 113], our method can give global bounds for the rank-constrained optimization. This makes our method especially useful for certification problems in quantum information, where global bounds are usually necessary to establish conclusions with certainty.

In order to demonstrate the usefulness of our method, we first show that the optimization over pure quantum states or unitary matrices in quantum information can be naturally written as a rank-constrained optimization. This provides a complete characterization of faithful entanglement [108, 114] and of mixed unitary channels [115, 116]. The second example concerns majorization uncertainty relations [117, 118], and the third dimension-bounded orthonormal representations of graphs [119], which is closely related to the existence of quantum contextuality in a given measurement configuration [120, 121]. Finally, we consider the maximum cut (Max-Cut) problem [122] and quadratic optimization over Boolean vectors [123]. Not only are these problems very important in classical information theory, but they also find various applications in statistical physics [124] and complex networks [125]. Remarkably, solving these optimization problems with noisy intermediate-scale quantum computers has drawn a lot of research interest in recent years [126–129].

To begin with, we explain the core idea of our method, first for matrices with complex entries, then for real matrices. Furthermore, we provide some insight into the quantum de Finetti theorem and the uniqueness of the corresponding decompositions

into multi-copy states. We also discuss how symmetries can be used to simplify the resulting sequence of SDPs. Subsequently, we present several examples, where our methods can be applied. Finally, we discuss more general forms of rank-constrained SDPs, rank-constrained quadratic and higher-order optimization problems, as well as open questions.

3.2 Rank-constrained SDP and quantum entanglement

SDPs are widely used in various branches of science, especially in the quantum regime. One of the reasons is that density matrices are automatically positive semidefinite, so that related optimization problems naturally contain some semidefinite constraints. Another important reason that SDPs have drawn a lot of interest is that there are efficient algorithms for solving them [89], moreover, symmetries can be used to drastically simplify the SDPs [130–132]. In many cases, however, one cannot directly formulate an SDP, as some non-convex constraints remain. This happens, for example, when the underlying quantum states are required to be pure or the quantum system is of bounded dimension. These restrictions will introduce some extra rank constraints, which is the main focus of this chapter.

The prototype optimization problem we consider is given by

$$\begin{aligned} \max_{\rho} \quad & \mathrm{tr}(X\rho) \\ \text{s.t.} \quad & \Lambda(\rho) = Y,\ \mathrm{tr}(\rho) = 1, \\ & \rho \geq 0,\ \mathrm{rank}(\rho) \leq k. \end{aligned} \tag{3.1}$$

Here, ρ and X are $n \times n$ matrices with real ($\mathbb{F} = \mathbb{R}$) or complex ($\mathbb{F} = \mathbb{C}$) entries, which are symmetric (resp. Hermitian). Λ is a map from matrices in $\mathbb{F}^{n\times n}$ to matrices in $\mathbb{F}^{m\times m}$ and consequently $Y \in \mathbb{F}^{m\times m}$. In this way, the constraint $\Lambda(\rho) = Y$ denotes all affine equality constraints. While our main results are formulated for the rank-constrained SDP in Eq. (3.1), we stress that our method can also be extended to more general cases with (semidefinite) inequality constraints $\Lambda(\rho) \leq Y$, without the normalization condition $\mathrm{tr}(\rho) = 1$, or even without the positivity constraint $\rho \geq 0$.

3.2.1 Optimization over complex matrices

We start with $\mathbb{F} = \mathbb{C}$ for the optimization in Eq. (3.1), where we can easily apply the results from quantum information. Let $\mathcal{F}$ be the feasible region of optimization (3.1),

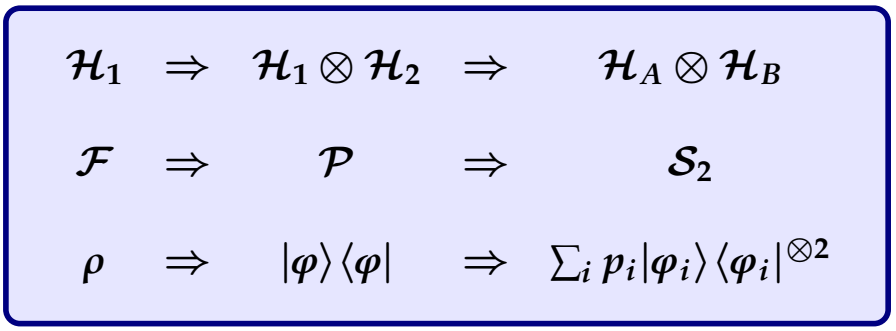


FIGURE 3.1: [99] An illustration of the relations between the feasible region $\mathcal{F}$, the purification $\mathcal{P}$, and the two-party extension $\mathcal{S}_2$. $|\varphi\rangle$ is a purification of ρ, $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}_1 \otimes \mathcal{H}_2$, and $|\varphi_i\rangle$ are states in $\mathcal{P}$.

i.e.,

$$\mathcal{F} = \{\rho \mid \Lambda(\rho) = Y, \mathrm{tr}(\rho) = 1, \rho \geq 0, \mathrm{rank}(\rho) \leq k\}. \tag{3.2}$$

With the terminology in quantum information, $\mathcal{F}$ is a subset of quantum states in the quantum system (or Hilbert space) $\mathbb{C}^n$.

Now, we recall the notion of state purification in quantum information as described in Section 2.2. Let $\mathcal{H}_1 = \mathbb{C}^n$ and $\mathcal{H}_2 = \mathbb{C}^k$ be two quantum systems (Hilbert spaces). Then, a quantum state ρ in $\mathcal{H}_1$ satisfies that $\mathrm{rank}(\rho) \leq k$ if, and only if, there exists a pure state $|\varphi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ such that $\mathrm{tr}_2(|\varphi\rangle\langle\varphi|) = \rho$. Thus, $\mathcal{F} \subset L(\mathcal{H}_1)$ can be written as $\mathrm{tr}_2(\mathcal{P})$, where

$$\mathcal{P} = \{ |\varphi\rangle\langle\varphi| \mid \widetilde{\Lambda}(|\varphi\rangle\langle\varphi|) = Y,\ \langle\varphi|\varphi\rangle = 1\} \subset L(\mathcal{H}_1 \otimes \mathcal{H}_2), \tag{3.3}$$

with $\widetilde{\Lambda}(\cdot) = \Lambda[\mathrm{tr}_2(\cdot)]$. Let $\mathrm{conv}(\mathcal{P})$ be the convex hull of $\mathcal{P}$, i.e., all states of the form $\sum_i p_i |\varphi_i\rangle\langle\varphi_i|$, where the p_i form a probability distribution and $|\varphi_i\rangle\langle\varphi_i| \in \mathcal{P}$. By noting that the maximum value of a linear function can always be achieved at extreme points, the optimization in Eq. (3.1) is equivalent to

$$\max_{\rho \in \mathcal{F}} \mathrm{tr}(X\rho) = \max_{\Phi \in \mathrm{conv}(\mathcal{P})} \mathrm{tr}(\widetilde{X}\Phi), \tag{3.4}$$

where $\widetilde{X} = X \otimes \mathbb{1}_k \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ with $\mathbb{1}_k$ being the identity operator on $\mathcal{H}_2$ as $X \in L(\mathcal{H}_1)$ because ρ and X have the same matrix dimensions.

Equation (3.4) implies that if we can fully characterize $\mathrm{conv}(\mathcal{P})$, the optimization (3.1) is solved. To this end, we utilize the notion of separable states. More specifically, we let $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^n \otimes \mathbb{C}^k$ and define the SEP cone on $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$\mathrm{SEP} = \mathrm{conv}\{M_A \otimes N_B \mid M_A \geq 0,\ N_B \geq 0\}. \tag{3.5}$$

Physically, SEP is the set of all unnormalized separable quantum states (besides the zero matrix). SEP is a proper convex cone, and its dual cone is given by

$$\mathrm{SEP}^* = \left\{ W_{AB} \;\middle|\; \mathrm{tr}(W_{AB}\Phi_{AB}) \geq 0 \quad \forall\, \Phi_{AB} \in \mathrm{SEP} \right\}, \tag{3.6}$$

which, in the language of quantum information, corresponds to the set of entanglement witnesses and positive semidefinite matrices, i.e., the set of block-positive matrices.

Then, we consider the two-party extension of the purified feasible states,

$$\mathcal{S}_2 = \mathrm{conv}\left(\left\{ |\varphi\rangle\langle\varphi|_A \otimes |\varphi\rangle\langle\varphi|_B \;\middle|\; |\varphi\rangle\langle\varphi| \in \mathcal{P} \right\} \right), \tag{3.7}$$

where $|\varphi\rangle_A$ and $|\varphi\rangle_B$ are the same state but belong to $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively; see Fig. 3.1. One can easily check that

$$\mathrm{tr}_B(\mathcal{S}_2) = \mathrm{conv}(\mathcal{P}). \tag{3.8}$$

The benefit of introducing the two-party extension is that we can fully characterize $\mathcal{S}_2$ with the SEP cone, and hence $\mathrm{conv}(\mathcal{P})$ is also fully characterized.

The first necessary condition for $\Phi_{AB} \in \mathcal{S}_2$ is that it is separable with respect to the bipartition $(A|B)$, i.e.,

$$\Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1. \tag{3.9}$$

Second, $\Phi_{AB} \in \mathcal{S}_2$ implies that it is within the symmetric subspace of $\mathcal{H}_A \otimes \mathcal{H}_B$. Mathematically, this can be written as

$$V_{AB}\Phi_{AB} = \Phi_{AB}, \tag{3.10}$$

where V_{AB} is the swap operator between $\mathcal{H}_A$ and $\mathcal{H}_B$, i.e., $V_{AB}|\psi_1\rangle_A|\psi_2\rangle_B = |\psi_2\rangle_A|\psi_1\rangle_B$ for any $|\psi_1\rangle, |\psi_2\rangle \in \mathbb{C}^n \otimes \mathbb{C}^k$. In contrast to Eq. (3.10), the similar intuitive constraint $V_{AB}\Phi_{AB}V_{AB} = \Phi_{AB}$ forces Φ_{AB} to be permutation-invariant but not necessarily in the symmetric subspace. This weaker condition would also allow separable states such as $\Phi_{AB} = \frac{1}{2}|01\rangle\langle 01| + \frac{1}{2}|10\rangle\langle 10|$ which is not of the two-copy form needed in Eq. (3.7). The constraint in Eq. (3.10) can alternatively be formulated as $P_+\Phi_{AB}P_+ = \Phi_{AB}$ with the projector onto the symmetric subspace $P_+ = (\mathbb{1}_{AB} + V_{AB})/2$ as the swap operator V_{AB} has eigenvalues ± 1 only.

The final necessary condition needed arises from Eq. (3.3), i.e., $\widetilde{\Lambda}(|\varphi\rangle\langle\varphi|) = Y$ for $|\varphi\rangle\langle\varphi| \in \mathcal{P}$. Then, Eq. (3.7) implies that

$$\widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = Y \otimes \mathrm{tr}_A(\Phi_{AB}) \tag{3.11}$$

for all $\Phi_{AB} \in \mathcal{S}_2$, where $\widetilde{\Lambda}_A(\cdot)$ is the map $\widetilde{\Lambda}(\cdot) = \Lambda[\mathrm{tr}_2(\cdot)]$ acting on system $\mathcal{H}_A$ only, and Id_B is the identity map on $\mathcal{H}_B$. Hereafter, we will also use a similar convention for matrices, e.g., $\widetilde{X}_A$ denotes the matrix $\widetilde{X}$ on system $\mathcal{H}_A$.

Surprisingly, the conditions in Eqs. (3.9, 3.10, 3.11) are also sufficient for $\Phi_{AB} \in \mathcal{S}_2$. To see this, note that the constraints in Eqs. (3.9, 3.10) imply that Φ_{AB} is a separable state in the symmetric subspace, which always admits the form [133]

$$\Phi_{AB} = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|_A \otimes |\varphi_i\rangle\langle\varphi_i|_B , \tag{3.12}$$

where the p_i form a probability distribution and the $|\varphi_i\rangle$ are normalized. Hereafter, without loss of generality, we assume that all p_i are strictly positive. From Eqs. (3.3, 3.7), to show that $\Phi_{AB} \in \mathcal{S}_2$ we only need to show that

$$\widetilde{\Lambda}(|\varphi_i\rangle\langle\varphi_i|) = Y \tag{3.13}$$

for all $|\varphi_i\rangle$. To this end, we introduce an auxiliary map

$$\mathcal{E}(\cdot) = \widetilde{\Lambda}(\cdot) - \mathrm{tr}(\cdot)Y. \tag{3.14}$$

Thus, the last constraint is equivalent to $\mathcal{E}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = 0$, which implies that

$$\mathcal{E}_A \otimes \mathcal{E}_B^\dagger(\Phi_{AB}) = 0, \tag{3.15}$$

where $\mathcal{E}^\dagger$ is the linear map satisfying $\mathcal{E}^\dagger(X) = [\mathcal{E}(X)]^\dagger$ for any Hermitian operator X, and the subscripts A, B in $\mathcal{E}_A, \mathcal{E}_B^\dagger$ indicate that the maps operate on systems $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Note that $\mathcal{E}^\dagger$ is not the dual map of $\mathcal{E}$. Then, Eqs. (3.12, 3.15) imply that

$$\sum_i p_i E_i \otimes E_i^\dagger = 0, \tag{3.16}$$

where $E_i = \mathcal{E}(|\varphi_i\rangle\langle\varphi_i|)$. Let V be the swap operator acting on the same space as $E_i \otimes E_i^\dagger$, then the relations $\mathrm{tr}[V(E_i \otimes E_i^\dagger)] = \mathrm{tr}(E_i E_i^\dagger)$ imply that

$$\mathrm{tr}\left[V\left(\sum_i p_i E_i \otimes E_i^\dagger\right)\right] = \sum_i p_i \,\mathrm{tr}(E_i E_i^\dagger) = 0. \tag{3.17}$$

Furthermore, as $\mathrm{tr}(E_i E_i^\dagger) > 0$ unless $E_i = 0$, we obtain

$$\mathcal{E}(|\varphi_i\rangle\langle\varphi_i|) = E_i = 0 \tag{3.18}$$

for all $|\varphi_i\rangle$. Then, Eq. (3.13) follows directly from the definition of $\mathcal{E}$ in Eq. (3.14), and hence $\Phi_{AB} \in \mathcal{S}_2$.

With the full characterization of $\mathcal{S}_2$ from Eqs. (3.9, 3.10, 3.11), we can directly rewrite the rank-constrained SDP in Eq. (3.1). The result is a so-called conic program, as one constraint is defined by the cone of separable states.

Theorem 3.1. *For $\mathbb{F} = \mathbb{C}$, the rank-constrained SDP in Eq.* (3.1) *is equivalent to the following conic program*

$$\begin{aligned} \max_{\Phi_{AB}} \quad & \mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_B \Phi_{AB}) \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1,\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = Y \otimes \mathrm{tr}_A(\Phi_{AB}). \end{aligned} \tag{3.19}$$

This conic program cannot be directly solved because the characterization of the SEP cone is still an NP-hard problem [60]. Actually, this is expected, because the rank-constrained SDP is, in general, also NP-hard. However, in quantum information theory many outer relaxations of the SEP cone are known. For example, the PPT criterion provides a pretty good approximation for low-dimensional quantum systems. More generally, inspired by the DPS hierarchy described in Section 2.8, we obtain a complete hierarchy for rank-constrained optimization in Eq. (3.1).

To express the hierarchy, we need to introduce the notion of symmetric subspaces for multiple parties. We label the N parties as $A, B, \ldots, Z$ and $\mathcal{H}_A = \mathcal{H}_B = \cdots = \mathcal{H}_Z = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^n \otimes \mathbb{C}^k$. For any $\mathcal{H}^{\otimes N} := \mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots \otimes \mathcal{H}_Z$, the symmetric subspace is defined as

$$\left\{ |\Psi\rangle \in \mathcal{H}^{\otimes N} \;\middle|\; V_\sigma |\Psi\rangle = |\Psi\rangle \quad \forall\, \sigma \in S_N \right\}, \tag{3.20}$$

where S_N is the permutation group over N symbols and V_σ are the corresponding operators on the N parties $A, B, \ldots, Z$. Let P_N^+ denote the orthogonal projector onto the symmetric subspace of $\mathcal{H}^{\otimes N}$, then P_N^+ can be explicitly written as

$$P_N^+ = \frac{1}{N!} \sum_{\sigma \in S_N} V_\sigma. \tag{3.21}$$

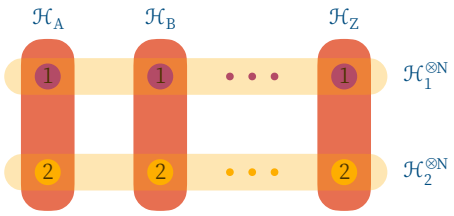


FIGURE 3.2: [99] An illustration of the N-party extension $\Phi_{AB\cdots Z}$. $\mathcal{H}_1 = \mathbb{C}^n$ is the n-dimensional Hilbert space on which the rank-constrained optimization is defined. $\mathcal{H}_2 = \mathbb{C}^k$ is the k-dimension auxiliary Hilbert space that is used for purifying the rank-k (more precisely, rank no larger than k) states in $\mathcal{H}_1 = \mathbb{C}^n$. Sometimes, we also denote $\mathcal{H}_1^{\otimes N}$ as $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1} \otimes \cdots \otimes \mathcal{H}_{Z_1}$ in order to distinguish the Hilbert spaces $\mathcal{H}_1$ for different parties (similarly for $\mathcal{H}_2^{\otimes N} = \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \cdots \otimes \mathcal{H}_{Z_2}$).

Hereafter, without ambiguity, we will also use P_N^+ to denote the corresponding symmetric subspace. For example, a state $\Phi_{AB\cdots Z}$ is within the symmetric space, i.e., $\Phi_{AB\cdots Z} = \sum_i p_i |\Psi_i\rangle \langle\Psi_i|$ for $|\Psi_i\rangle \in P_N^+$, if and only if $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$.

Now we are ready to state the complete hierarchy for rank-constrained optimization.

Theorem 3.2. *For $\mathbb{F} = \mathbb{C}$, let ξ be the solution of the rank-constrained SDP in Eq. (3.1). Then, for any N, ξ is upper bounded by the solution ξ_N of the following SDP hierarchy*

$$\begin{aligned} \max_{\Phi_{AB\cdots Z}} \ & \operatorname{tr}(\widetilde{X}_A \otimes \mathbb{1}_{B\cdots Z} \Phi_{AB\cdots Z}) \\ \text{s.t.} \ & \Phi_{AB\cdots Z} \geq 0, \ \operatorname{tr}(\Phi_{AB\cdots Z}) = 1, \\ & P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}, \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_{B\cdots Z}(\Phi_{AB\cdots Z}) = Y \otimes \operatorname{tr}_A(\Phi_{AB\cdots Z}). \end{aligned} \tag{3.22}$$

Furthermore, the SDP hierarchy is complete in the sense that $\xi_{N+1} \leq \xi_N$ and $\lim_{N\to+\infty} \xi_N = \xi$.

The proof is similar to the proof of Theorem 2 in Ref. [134]. For completeness, we also present it here. To prove Theorem 3.2, we take advantage of the following lemma, which can be viewed as a special case of the quantum de Finetti theorem [135]; see also related results in Refs. [136, 137].

Lemma 3.3. *Let ρ_N be an N-party quantum state in the symmetric subspace P_N^+, then for all $\ell < N$ there exists an ℓ-party quantum state*

$$\sigma_\ell = \sum_\mu p_\mu |\varphi_\mu\rangle \langle\varphi_\mu|^{\otimes \ell}, \tag{3.23}$$

i.e., a fully separable state in P_ℓ^+, such that

$$\|\mathrm{tr}_{N-\ell}(\rho_N) - \sigma_\ell\| \leq \frac{4\ell D}{N}, \tag{3.24}$$

where $\|\cdot\|$ is the trace norm and D is the local dimension.

The part that ξ is upper bounded by ξ_N for any N is obvious. Hence, we only need to prove that $\xi_{N+1} \leq \xi_N$ and $\lim_{N\to+\infty} \xi_N = \xi$.

We first show that $\xi_{N+1} \leq \xi_N$. This follows from the fact that if a multi-party quantum state is within the symmetric subspace, so are the reduced states. Mathematically, we have the relation

$$(P_N^+ \otimes \mathbb{1}_{nk}) P_{N+1}^+ = P_{N+1}^+. \tag{3.25}$$

Suppose that there exists an $(N+1)$-party extension $\Phi_{AB\cdots ZZ'}$ satisfying all constraints that achieves the maximum ξ_{N+1} in Theorem 3.2. Then, the constraint

$$P_{N+1}^+ \Phi_{AB\cdots ZZ'} P_{N+1}^+ = \Phi_{AB\cdots ZZ'} \tag{3.26}$$

implies that

$$\begin{aligned}(P_N^+ \otimes \mathbb{1}_{nk})\Phi_{AB\cdots ZZ'}(P_N^+ \otimes \mathbb{1}_{nk}) &= P_N^+ \otimes \mathbb{1}_{nk} P_{N+1} \Phi_{AB\cdots ZZ'} P_{N+1} P_N^+ \otimes \mathbb{1}_{nk} \\ &= P_{N+1}\Phi_{AB\cdots ZZ'}P_{N+1} \\ &= \Phi_{AB\cdots ZZ'}.\end{aligned} \tag{3.27}$$

Thus, one can easily verify that the reduced state $\mathrm{tr}_{Z'}(\Phi_{AB\cdots ZZ'})$ is an N-party extension satisfying all the constraints in Theorem 3.2 with objective value ξ_{N+1}. From this, the result $\xi_{N+1} \leq \xi_N$ follows.

Next, we prove the convergence part, i.e., $\lim_{N\to+\infty} \xi_N = \xi$. Suppose that the solution ξ_N of the N-party extension in Theorem 3.2 is achieved by the quantum state $\Phi_{AB\cdots Z}$. Let $\Phi_{AB}^N = \mathrm{tr}_{C\cdots Z}(\Phi_{ABC\cdots Z})$, then Φ_{AB}^N satisfies that

$$\begin{aligned}&\mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_B \Phi_{AB}^N) = \xi_N,\ \mathrm{tr}(\Phi_{AB}^N) = 1, \\ &\widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}^N) = Y \otimes \mathrm{tr}_A(\Phi_{AB}^N).\end{aligned} \tag{3.28}$$

Further, Lemma 3.3 implies that there exist separable states $\widetilde{\Phi}_{AB}^N$ such that

$$V_{AB}\widetilde{\Phi}_{AB}^N = \widetilde{\Phi}_{AB}^N, \tag{3.29}$$

$$\|\Phi_{AB}^N - \widetilde{\Phi}_{AB}^N\| \leq \frac{8nk}{N}. \tag{3.30}$$

As the set of quantum states for any fixed dimension is compact, we can choose a convergent subsequence $\Phi_{AB}^{N_i}$ of the sequence Φ_{AB}^{N}. Thus, Eq. (3.30) implies that

$$\Phi_{AB} := \lim_{i\to+\infty} \Phi_{AB}^{N_i} = \lim_{i\to+\infty} \widetilde{\Phi}_{AB}^{N_i}. \tag{3.31}$$

As all $\widetilde{\Phi}_{AB}^{N_i}$ are separable and the set of separable states is closed, we have that $\Phi_{AB} = \lim_{i\to+\infty} \widetilde{\Phi}_{AB}^{N_i}$ is separable. Further, as all the functions on Φ_{AB}^{N} or $\widetilde{\Phi}_{AB}^{N}$ in Eqs. (3.28, 3.29) are continuous, Eq. (3.31) implies that Φ_{AB} satisfies all the constraints in Eq. (3.19). In other words, Φ_{AB} is a feasible point of program (3.19), thus $\mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_B \Phi_{AB}) = \lim_{N\to+\infty} \xi_N \leq \xi$. Together with the fact that $\xi_N \geq \xi$, we then have $\lim_{N\to+\infty} \xi_N = \xi$.

In addition, any criterion for the full separability of $\Phi_{AB\cdots Z}$ can be added to the optimization in Eq. (3.22), which can give a better upper bound for the optimization in Eq. (3.1). For example, the PPT criterion, more precisely, PPT with respect to all bipartitions, can also be added as additional constraints, which can give better upper bounds ξ_N^T, i.e., $\xi \leq \xi_{N+1}^T \leq \xi_N^T \leq \xi_N$ and $\lim_{N\to+\infty} \xi_N^T = \xi$. Furthermore, it is sometimes convenient to denote the solution of the SDP by relaxing the rank constraint in Eq. (3.1) as ξ_1, then we have $\xi_2 \leq \xi_1$.

Let us estimate the complexity of the SDP hierarchy in Theorem 3.2. For the N-th level of the hierarchy, the dimension of the matrix reads $\dim(\mathcal{H}^{\otimes N}) = (nk)^N$, but it can be further reduced by taking advantage of the fact that $\Phi_{AB\cdots Z}$ is within the symmetric subspace, which has the dimension

$$\dim\left(P_N^+\right) = \binom{nk+N-1}{N} = \binom{nk+N-1}{nk-1}. \tag{3.32}$$

By noticing that $k \leq n$, Eq. (3.32) implies that for fixed dimension n the complexity of the SDP grows polynomially with the level of the hierarchy N, and for fixed level of hierarchy N the complexity of the SDP also grows polynomially with the dimension n. Similar results also hold when considering the PPT criterion, because the partial transpose of $\Phi_{AB\cdots Z}$ with respect to any bipartition is within the tensor product of two symmetric subspaces $P_k^+ \otimes P_{N-k}^+$ for some k [63].

3.2.2 Don't let de Finetti be misunderstood

Before proceeding to the optimization over real matrices, we want to add a few remarks in the context of the quantum de Finetti theorem. While Lemma 3.3 provides a quantitative statement about the distance between separable states and marginals of symmetric quantum states, the qualitative statement that any exchangeable state can

be written as the mixture of multi-copy states is more widely known [138]. More precisely, a state $\Phi_{AB\cdots Z}$ is called exchangeable if it is permutation-invariant, i.e., $V_\sigma \Phi_{AB\cdots Z} V_\sigma = \Phi_{AB\cdots Z}$ for any permutation σ, and there exist permutation-invariant extensions $\Phi_{AB\cdots Z\alpha\cdots\Omega}$ for any number of extra parties M, such that $\mathrm{Tr}_{\alpha\cdots\Omega}\, \Phi_{AB\cdots Z\alpha\cdots\Omega} = \Phi_{AB\cdots Z}$. An exchangeable state $\Phi_{AB\cdots Z}$ can then be written as

$$\Phi_{AB\cdots Z} = \int P(\rho)\rho^{\otimes N} d\rho, \tag{3.33}$$

where $P(\rho)$ is a probability distribution over all quantum states $\rho \in \mathcal{H}_A = \cdots = \mathcal{H}_\Omega$.

The authors of the highly cited Ref. [138], however, falsely claim that the probability distribution P would be always unique. Indeed, this is generally only the case if all the extensions $\Phi_{AB\cdots Z\alpha\cdots\Omega}$ are fixed. We want to clear up this misunderstanding by first, providing a very simple explicit counterexample. Namely, we find the following two-qubit state originating from different ensembles of two-copy mixtures,

$$\begin{aligned}
\Phi_{AB} &= \frac{1}{4}\left\{\left[\frac{1}{2}(\mathbb{1}+X)\right]^{\otimes 2} + \left[\frac{1}{2}(\mathbb{1}-X)\right]^{\otimes 2} + \left[\frac{1}{2}(\mathbb{1}+Z)\right]^{\otimes 2} + \left[\frac{1}{2}(\mathbb{1}-Z)\right]^{\otimes 2}\right\} \\
&= \frac{1}{4}\sum_{\pm_1,\pm_2}\left[\frac{1}{2}\left(\mathbb{1} \pm_1 \frac{1}{\sqrt{2}}X \pm_2 \frac{1}{\sqrt{2}}Z\right)\right]^{\otimes 2} \\
&= \frac{1}{4}\left(\mathbb{1}\otimes\mathbb{1} + \frac{1}{2}X\otimes X + \frac{1}{2}Z\otimes Z\right).
\end{aligned} \tag{3.34}$$

Second, we show that such counterexamples exist for any number of parties N and local dimension d.

Observation 3.4. For any number of parties N and local dimension d, the normalized projector onto the symmetric subspace P_N^+ has different decompositions $\{p_\mu, \rho_\mu^{\otimes N}\}$.

Proof. We have that

$$P_N^+ \propto \int dU\, U^{\otimes N} |0\rangle\langle 0|^{\otimes N} \left(U^\dagger\right)^{\otimes N}, \tag{3.35}$$

where the integral is taken w.r.t. the Haar measure, as well as

$$P_N^+ \propto \sum_{k=1}^{K} U_k^{\otimes N} |0\rangle\langle 0|^{\otimes N} \left(U_k^\dagger\right)^{\otimes N}, \tag{3.36}$$

where the U_k form a so-called unitary N-design, that transforms the integral to a finite sum, which exists for all N and d [139, 140]. □

On the other hand, there are also states with a unique decomposition without fixing any of the permutation-invariant extensions. In Lemma 4.2, we will show that indeed

all multi-copy states $\rho^{\otimes N}$ are extreme points in the space generated by mixtures of multi-copy states and hence, their decomposition is unique. Furthermore, there exist nonextremal states with a unique decomposition, too. Namely, consider the multi-qubit states

$$\Phi_{AB\cdots Z} = \frac{1}{2}\left[\frac{1}{2}(\mathbb{1} + Z)\right]^{\otimes N} + \frac{1}{2}\left[\frac{1}{2}(\mathbb{1} - Z)\right]^{\otimes N}, \tag{3.37}$$

for $N \geq 2$. Let us assume, that there is a different decomposition $\{p_\mu, \rho_\mu^{\otimes N}\}$ with single-qubit states $\rho_\mu = (\mathbb{1} + \boldsymbol{\lambda}^\mu \cdot \boldsymbol{\sigma})/2$. Then, the decomposition of $\rho_\mu^{\otimes N}$ in the Pauli basis is a sum of terms containing

$$\frac{1}{2^N}\left(\lambda_z^\mu\right)^2 Z \otimes Z \otimes \mathbb{1}^{\otimes(N-2)}. \tag{3.38}$$

Since $\Phi_{AB\cdots Z}$ contains this term with maximal weight, we have that $(\lambda_z^\mu)^2 = 1$ must hold for all μ, which leaves only the two multi-copy states that already appear in the decomposition in Eq. (3.37). The weights of these two states are then fixed by any term containing an odd number of Z, implying that the decomposition in Eq. (3.37) is indeed unique. Thus, the space generated by mixtures of multi-copy states has partly the geometry of a simplex. This feature should become more prevalent with increasing N.

3.2.3 Optimization over real matrices

We move on to consider the $\mathbb{F} = \mathbb{R}$ case, which is more important in classical information theory. One can easily verify that Theorem 3.1 can be directly generalized to the $\mathbb{F} = \mathbb{R}$ case, if the decomposition in Eq. (3.12) satisfies that $|\varphi_i\rangle\langle\varphi_i| \in \mathbb{R}^{nk\times nk}$. The obvious way to guarantee this is to define the set of separable states over $\mathbb{R}$. However, this hinders the application of known separability criteria developed in entanglement theory.

Thus, we employ a different method. We still use the separability cone SEP with respect to the complex numbers, more precisely,

$$\Phi_{AB} \in \mathrm{SEP} \cap \mathbb{R}^{nk\times nk}, \tag{3.39}$$

where SEP is still defined as in Eq. (3.5). Equations (3.12, 3.39) are not sufficient for guaranteeing that $|\varphi_i\rangle\langle\varphi_i| \in \mathbb{R}^{nk\times nk}$. An explicit counterexample is given by the (unnormalized) state

$$\Phi_{AB} = P_2^+ \propto \int dU\, U \otimes U\, |00\rangle\langle 00|\, U^\dagger \otimes U^\dagger. \tag{3.40}$$

This state is obviously in the symmetric subspace, real, and separable. However, it cannot be expressed as a mixture of real two-copy pure states, which can be seen by applying P_2^+ to a complex two-copy pure state $|\psi\rangle|\psi\rangle$. Apparently, it holds that $P_2^+|\psi\rangle|\psi\rangle = |\psi\rangle|\psi\rangle$ but also $P_2^+|\psi\psi\rangle\langle\psi\psi|P_2^+$ would be a mixture of real two-copy pure states which leads to a contradiction since the state $|\psi\psi\rangle\langle\psi\psi|$ is an extremal point in the state space.

Still, only a small modification to Eq. (3.39) is needed. For pure states, one has $|\varphi_i\rangle\langle\varphi_i|^T = |\varphi_i^*\rangle\langle\varphi_i^*|$, where $(\cdot)^T$ denotes the matrix transpose and $|(\cdot)^*\rangle$ the complex conjugation with respect to the same fixed basis. Hence, a necessary condition for $|\varphi_i\rangle\langle\varphi_i| \in \mathbb{R}^{nk\times nk}$ is

$$\Phi_{AB}^{T_A} = \Phi_{AB}, \tag{3.41}$$

meaning that the state Φ_{AB} is invariant under partial transposition.

Interestingly, due to the symmetry and separability of Φ_{AB}, Eq. (3.41) is also sufficient for guaranteeing that $|\varphi_i\rangle\langle\varphi_i| \in \mathbb{R}^{nk\times nk}$. From the form of Φ_{AB} in Eq. (3.12), we obtain

$$\Phi_{AB}^{T_A} = \sum_i p_i |\varphi_i^*\rangle\langle\varphi_i^*|_A \otimes |\varphi_i\rangle\langle\varphi_i|_B, \tag{3.42}$$

where $|\varphi_i^*\rangle$ denote the complex conjugate of $|\varphi_i\rangle$. Then, the fact that $\Phi_{AB}^{T_A} = \Phi_{AB}$ is a separable state within the symmetric subspace implies that

$$|\varphi_i\rangle\langle\varphi_i| = |\varphi_i^*\rangle\langle\varphi_i^*|, \tag{3.43}$$

i.e., $|\varphi_i\rangle\langle\varphi_i| \in \mathbb{R}^{nk\times nk}$ for all i. This proves Theorem 3.5. Notably, this argument can be directly generalized to multi-party states, which provides a simple proof for the result in Ref. [141].

Hence, we arrive at the following theorem for rank-constrained optimization over real matrices.

Theorem 3.5. *For $\mathbb{F} = \mathbb{R}$, the rank-constrained SDP in Eq. (3.1) is equivalent to the following conic program*

$$\begin{aligned} \max_{\Phi_{AB}} \quad & \mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_B \Phi_{AB}) \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1, \\ & V_{AB}\Phi_{AB} = \Phi_{AB},\ \Phi_{AB}^{T_A} = \Phi_{AB}, \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = Y \otimes \mathrm{tr}_A(\Phi_{AB}). \end{aligned} \tag{3.44}$$

Similarly to Theorem 3.2, we can also construct a complete hierarchy with the multi-party extension method for the real case:

Theorem 3.6. *For $\mathbb{F} = \mathbb{R}$, let ξ be the solution of the rank-constrained SDP in Eq. (3.1). Then, for any N, ξ is upper bounded by the solution ξ_N of the following SDP hierarchy*

$$\begin{aligned} \max_{\Phi_{AB\cdots Z}} \ & \mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_{B\cdots Z} \Phi_{AB\cdots Z}) \\ \text{s.t.} \ & \Phi_{AB\cdots Z} \geq 0, \ \mathrm{tr}(\Phi_{AB\cdots Z}) = 1, \\ & P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}, \ \Phi_{AB\cdots Z}^{T_A} = \Phi_{AB\cdots Z}, \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_{B\cdots Z}(\Phi_{AB\cdots Z}) = Y \otimes \mathrm{tr}_A(\Phi_{AB\cdots Z}). \end{aligned} \tag{3.45}$$

Furthermore, the SDP hierarchy is complete, i.e., $\xi_{N+1} \leq \xi_N$ and $\lim_{N\to+\infty} \xi_N = \xi$.

We emphasize that all variables involved in Eqs. (3.44) and (3.45) are taken as real matrices. In addition, due to the permutation symmetry induced by $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$, $\Phi_{AB\cdots Z}^{T_A} = \Phi_{AB\cdots Z}$ already ensures the partial-transpose-invariance with respect to all bipartitions. This also makes the PPT criterion as an additional separability condition redundant for the hierarchy in Eq. (3.45).

3.2.4 Inherent symmetry for the hierarchy

Before proceeding further, we briefly describe inherent symmetries emerging from the ancilla introduced for purification in Eqs. (3.19, 3.22, 3.44, 3.45), which is particularly useful for practical implementations. In a convex optimization problem, if a group action G does not change the objective function and feasible region, then the variables can be assumed to be G-invariant. Specifically, if the SDP, $\max_{\Phi\in\mathcal{S}} \mathrm{tr}(\Phi X)$, satisfies that $g\mathcal{S}g^\dagger \subset \mathcal{S}$ and $gXg^\dagger = X$ for all $g \in G$, then we can add an extra G-invariant constraint that $g\Phi g^\dagger = \Phi$ for all $g \in G$.

For the hierarchy in the complex case in Theorems 3.1 and 3.2, regardless of the actual form of X, Λ, and Y, there is an inherent $U^{\otimes N}$ symmetry on $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \cdots \otimes \mathcal{H}_{Z_2} = (\mathbb{C}^k)^{\otimes N}$ for all unitary matrices $U \in SU(k)$, i.e., on the N auxiliary Hilbert spaces $\mathcal{H}_2^{\otimes N}$ shown in Fig. 3.2. Hence, $\Phi_{AB\cdots Z}$ can be restricted to those satisfying

$$(\mathbb{1}_n \otimes U)^{\otimes N} \Phi_{AB\cdots Z} (\mathbb{1}_n \otimes U^\dagger)^{\otimes N} = \Phi_{AB\cdots Z}. \tag{3.46}$$

This implies that $\Phi_{AB\cdots Z}$ is generated by the symmetric group S_N in $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \cdots \otimes \mathcal{H}_{Z_2} = (\mathbb{C}^k)^{\otimes N}$ by the Schur-Weyl duality [142].

We take the case $N = 2$ as an example to illustrate this point. Under the restriction in Eq. (3.46), Φ_{AB} admits the form

$$\Phi_{AB} = \Phi_I \otimes \mathbb{1}_{A_2B_2} + \Phi_V \otimes V_{A_2B_2}, \tag{3.47}$$

where Φ_I and Φ_V are operators on $\mathcal{H}_{A_1B_1}$, and $\mathbb{1}_{A_2B_2}$ and $V_{A_2B_2}$ are the identity and swap operators on $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} = \mathbb{C}^k \otimes \mathbb{C}^k$, respectively. By taking advantage of the relations

$$\begin{aligned} &V_{AB} = V_{A_1B_1} \otimes V_{A_2B_2},\ V_{A_2B_2}^{T_{A_2}} = k\left|\phi_k^+\right\rangle\left\langle\phi_k^+\right|, \\ &\mathrm{tr}_{A_2}(\mathbb{1}_{A_2B_2}) = k\mathbb{1}_{B_2},\ \mathrm{tr}_{A_2}(V_{A_2B_2}) = \mathbb{1}_{B_2}, \end{aligned} \tag{3.48}$$

where $\left|\phi_k^+\right\rangle = \frac{1}{\sqrt{k}}\sum_{\alpha=1}^{k}\left|\alpha\right\rangle_{A_2}\left|\alpha\right\rangle_{B_2}$ is a maximally entangled state, ξ_2^T can be simplified to

$$\begin{aligned} \max_{\Phi_I,\Phi_V}\quad & \mathrm{tr}\left[X_{A_1} \otimes \mathbb{1}_{B_1}(k^2\Phi_I + k\Phi_V)\right] \\ \text{s.t.}\quad & \Phi_V = V_{A_1B_1}\Phi_I,\ \Phi_I + \Phi_V \geq 0, \\ & \Phi_I - \Phi_V \geq 0,\ \Phi_I^{T_{A_1}} + k\Phi_V^{T_{A_1}} \geq 0, \\ & \Phi_I^{T_{A_1}} \geq 0,\ k^2\,\mathrm{tr}(\Phi_I) + k\,\mathrm{tr}(\Phi_V) = 1, \\ & \Lambda_{A_1} \otimes \mathrm{Id}_{B_1}(k\Phi_I + \Phi_V) = Y \otimes \mathrm{tr}_{A_1}(k\Phi_I + \Phi_V). \end{aligned} \tag{3.49}$$

A significant improvement in Eq. (3.49) is that the dimension of the variables is $\mathbb{C}^{n^2 \otimes n^2}$, which no longer depends on the rank k.

For the hierarchy in the real case in Theorems 3.5 and 3.6, we consider the symmetry $Q^{\otimes N}$ for orthogonal matrices $Q \in O(k)$, which would also simplify the structure of $\Phi_{AB\cdots Z}$ in $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \cdots \otimes \mathcal{H}_{Z_2} = (\mathbb{R}^k)^{\otimes N}$. The $O(k)$ symmetry can reduce $\Phi_{AB\cdots Z}$ to the Brauer algebra $B_N(k)$ in $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \cdots \otimes \mathcal{H}_{Z_2} = (\mathbb{R}^k)^{\otimes N}$ [142], which is more complicated than the $SU(k)$ symmetry.

For $N = 2$, the Brauer algebra $B_2(k)$ is the linear span of $\{\mathbb{1}_{A_2B_2}, V_{A_2B_2}, k\left|\phi_k^+\right\rangle\left\langle\phi_k^+\right|\}$, which implies that the symmetrized Φ_{AB} is of the form

$$\Phi_{AB} = \Phi_I \otimes \mathbb{1}_{A_2B_2} + \Phi_V \otimes V_{A_2B_2} + \Phi_\phi \otimes k\left|\phi_k^+\right\rangle\left\langle\phi_k^+\right|, \tag{3.50}$$

where Φ_I, Φ_V, and Φ_ϕ are operators on $\mathcal{H}_{A_1B_1}$. Correspondingly, ξ_2 can be simplified to

$$\begin{aligned} \max_{\Phi_I,\Phi_V,\Phi_\phi}\ & \mathrm{tr}\left[X_{A_1} \otimes \mathbb{1}_{B_1}(k^2\Phi_I + k\Phi_V + k\Phi_\phi)\right] \\ \text{s.t.}\ & \Phi_V = V_{A_1B_1}\Phi_I,\ \Phi_\phi = \Phi_V^{T_{A_1}},\ \Phi_I^{T_{A_1}} = \Phi_I, \\ & V_{A_1B_1}\Phi_\phi = \Phi_\phi,\ \Phi_I + \Phi_V \geq 0, \\ & \Phi_I - \Phi_V \geq 0,\ \Phi_I + \Phi_V + k\Phi_\phi \geq 0, \\ & k^2\,\mathrm{tr}(\Phi_I) + k\,\mathrm{tr}(\Phi_V) + k\,\mathrm{tr}(\Phi_\phi) = 1, \\ & \Lambda_{A_1} \otimes \mathrm{Id}_{B_1}(k\Phi_I + \Phi_V + \Phi_\phi) = Y \otimes \mathrm{tr}_{A_1}(k\Phi_I + \Phi_V + \Phi_\phi). \end{aligned} \tag{3.51}$$

Curiously, in the SDPs in Eqs. (3.49) and (3.51), the rank constraint k appears as a parameter that, in principle, can take on non-integer values. Indeed, k can, in some sense, be considered a continuous rank, which is useful for handling numerical errors.

Observation 3.7. A feasible point $\Phi_{A_1B_1} = k^2\Phi_I + k\Phi_V$ of the SDP in Eq. (3.49) with parameter $k \geq 1$ is also a feasible point $\Phi_{A_1B_1} = k'^2\Phi'_I + k'\Phi'_V$ of the SDP with parameter $k' \geq k$.

Proof. The observation is trivial when $k' = k$. In the following, we assume that $k' > k \geq 1$. From the relations $\Phi_V = V\Phi_I$, $\Phi'_V = V\Phi'_I$, and $\Phi_{A_1B_1} = k^2\Phi_I + k\Phi_V = k'^2\Phi'_I + k'\Phi'_V$, we obtain

$$\begin{aligned} k'^2\Phi'_I + k'\Phi'_V &= k^2\Phi_I + k\Phi_V, \\ k'\Phi'_I + k'^2\Phi'_V &= k\Phi_I + k^2\Phi_V, \end{aligned} \tag{3.52}$$

which further imply that

$$\Phi'_I = \frac{k(kk'-1)}{k'(k'^2-1)}\Phi_I + \frac{k(k'-k)}{k'(k'^2-1)}\Phi_V. \tag{3.53}$$

Thus, we can express Φ'_I and Φ'_V in terms of Φ_I and Φ_V. The feasibility follows from the feasibility of $\Phi_{A_1B_1} = k^2\Phi_I + k\Phi_V$ and

$$\begin{aligned} \Phi'_I \pm \Phi'_V &= \frac{k(k\pm 1)}{k'(k'\pm 1)}\Big(\Phi_I \pm \Phi_V\Big), \\ \big(\Phi'_I\big)^{T_{A_1}} &= \frac{k'-k}{k'(k'^2-1)}\left(\Phi_I^{T_{A_1}} + k\Phi_V^{T_{A_1}}\right) + \frac{k^2-1}{k'^2-1}\Phi_I^{T_{A_1}}, \\ \big(\Phi'_I\big)^{T_{A_1}} + k'\big(\Phi'_V\big)^{T_{A_1}} &= \frac{k}{k'}\left(\Phi_I^{T_{A_1}} + k\Phi_V^{T_{A_1}}\right), \end{aligned} \tag{3.54}$$

since all coefficients are nonnegative. The linear constraints are obviously satisfied as we consider $\Phi'_{A_1B_1} = \Phi_{A_1B_1}$. □

A similar statement also holds in the real case.

Observation 3.8. A feasible point $\Phi_{A_1B_1} = k^2\Phi_I + k\Phi_V + k\Phi_\phi$ of the SDP in Eq. (3.51) with parameter $k \geq 1$ is also a feasible point $\Phi_{A_1B_1} = k'^2\Phi'_I + k'\Phi'_V + k'\Phi'_\phi$ of the SDP with parameter $k' \geq k$.

Proof. In this case, from $\Phi'_V = V\Phi'_I$, $\Phi'_\phi = (\Phi'_V)^{T_{A_1}}$, and $k'^2\Phi'_I + k'\Phi'_V + k'\Phi'_\phi = k^2\Phi_I + k\Phi_V + k\Phi_\phi$, we obtain

$$\Phi'_I = \frac{k(kk'+k-2)}{k'(k'+2)(k'-1)}\Phi_I + \frac{k(k'-k)}{k'(k'+2)(k'-1)}\Phi_V + \frac{k(k'-k)}{k'(k'+2)(k'-1)}\Phi_\phi. \tag{3.55}$$

Analogous to the proof of Observation 3.7, it is straightforward to verify that the coefficients in the following equalities are all nonnegative,

$$\Phi_I' + \Phi_V' = \frac{(k+2)(k-1)}{(k'+2)(k'-1)}\Big(\Phi_I + \Phi_V\Big) + \frac{2(k'-k)}{k'(k'+2)(k'-1)}\Big(\Phi_I + \Phi_V + k\Phi_\phi\Big) \tag{3.56}$$

$$\Phi_I' - \Phi_V' = \frac{k(k-1)}{k'(k'-1)}\Big(\Phi_I - \Phi_V\Big), \tag{3.57}$$

$$\Phi_I' + \Phi_V' + k'\Phi_\phi' = \frac{k}{k'}\Big(\Phi_I + \Phi_V + k\Phi_\phi\Big). \tag{3.58}$$

It is also obvious that $(\Phi_I')^{T_{A_1}} = \Phi_I'$ and $V\Phi_\phi' = \Phi_\phi'$. Hence, the feasibility follows. □

Thus, the set of feasible points grows monotonically with continuous k, and hence, the same is true for the objective value. Apart from the interpretation as a continuous rank, this also helps in preventing invalid conclusions because of numerical errors, since parameters k can be sampled in a region around the considered rank.

3.3 Examples

In this section, we show that our method can be widely used in quantum and classical information theory. As illustrations, we investigate examples of optimization over pure states and unitary channels, the characterization of faithful entanglement, majorization uncertainty relations, and quantum contextuality as problems in quantum information theory. Concerning classical information theory, we study the Max-Cut problem, pseudo-Boolean optimization, and the minimum dimension of the orthonormal representation of graphs.

3.3.1 Optimization over pure quantum states and unitary channels

A direct application of our method in quantum information theory is the optimization over pure states. For example, we consider the optimization problem from incomplete measurement information

$$\begin{aligned} \max_{|\varphi\rangle} \quad & \langle\varphi| X |\varphi\rangle \\ \text{s.t.} \quad & \langle\varphi| M_i |\varphi\rangle = m_i, \end{aligned} \tag{3.59}$$

where the M_i are the performed measurements and the m_i are the corresponding measurement results. This can be viewed as a refined problem of compressed sensing tomography [110], in which the feasibility problem is considered. The optimization in

Eq. (3.59) is obviously a rank-constrained SDP,

$$\begin{aligned} \max_{\rho} \quad & \mathrm{tr}(X\rho) \\ \text{s.t.} \quad & \mathrm{tr}(M_i\rho) = m_i,\ \mathrm{tr}(\rho) = 1, \\ & \rho \geq 0,\ \mathrm{rank}(\rho) = 1. \end{aligned} \tag{3.60}$$

Thus, Theorem 3.1 gives the following equivalent conic program

$$\begin{aligned} \max_{\Phi_{AB}} \quad & \mathrm{tr}(X_A \otimes \mathbb{1}_B \Phi_{AB}) \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1,\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ & \mathrm{tr}_A(M_i \otimes \mathbb{1}_B \Phi_{AB}) = m_i \,\mathrm{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.61}$$

from which a complete SDP hierarchy can be constructed using Theorem 3.2. Similarly, we can also consider the optimization over low-rank quantum states.

Due to the Choi-Jamiołkowski duality described in Section 2.3, the result in Eqs. (3.60, 3.61) can also be used for the optimization over unitary (and low-Kraus-rank) channels. As an example, we show that our method provides a complete characterization of the mixed-unitary channels, which was recently proved to be an NP-hard problem [116].

A channel Λ is called mixed-unitary if there exists a positive integer m, a probability distribution $(p_1, p_2, \ldots, p_m)$, and unitary operators $U_1, U_2, \ldots, U_m$ such that

$$\Lambda(\rho) = \sum_{i=1}^{m} p_i U_i \rho U_i^\dagger. \tag{3.62}$$

According to the Choi-Jamiołkowski duality, a channel is mixed-unitary if, and only if, the corresponding Choi state η_Λ is a mixture of maximally entangled states. Thus, characterizing the mixed-unitary channels is equivalent to characterizing the mixture of maximally entangled states,

$$\mathcal{M} = \mathrm{conv}\left\{ |\phi\rangle\langle\phi| \;\middle|\; \mathrm{tr}_1(|\phi\rangle\langle\phi|) = \frac{\mathbb{1}_n}{n} \right\}. \tag{3.63}$$

According to Eq. (3.8), Λ being mixed-unitary, i.e., $\eta_\Lambda \in \mathcal{M}$, is equivalent to the following feasibility problem

$$\begin{aligned} \text{find} \quad & \Phi_{AB} \in \text{SEP} \\ \text{s.t.} \quad & \text{tr}_B(\Phi_{AB}) = \eta_\Lambda,\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ & \text{tr}_{A_1} \otimes \text{Id}_{A_2} \otimes \text{Id}_B(\Phi_{AB}) = \frac{\mathbb{1}_n}{n} \otimes \text{tr}_A(\Phi_{AB}), \\ & \text{Id}_{A_1} \otimes \text{tr}_{A_2} \otimes \text{Id}_B(\Phi_{AB}) = \frac{\mathbb{1}_n}{n} \otimes \text{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.64}$$

where the last constraint follows from $\text{tr}_2(|\phi\rangle\langle\phi|) = \mathbb{1}_n/n$ according to Eq. (3.63). This constraint is redundant for Eq. (3.64), but it may help when semidefinite relaxations are considered.

A further application comes from entanglement theory. Following Ref. [108], the optimization over $\mathcal{M}$ also provides a complete characterization of faithful entanglement [114], i.e., the entangled states that are detectable by fidelity-based witnesses. In Ref. [108], the authors prove that a state $\rho \in \mathbb{C}^n \otimes \mathbb{C}^n$ is faithful if, and only if, $\xi := \max_{\sigma\in\mathcal{M}} \text{tr}(\sigma\rho) > 1/n$. According to Theorem 3.1, the solution ξ also equals the following conic program

$$\begin{aligned} \text{max} \quad & \text{tr}(\rho_A \otimes \mathbb{1}_B \Phi_{AB}) \\ \text{s.t.} \quad & \Phi_{AB} \in \text{SEP},\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ & \text{tr}_{A_1} \otimes \text{Id}_{A_2} \otimes \text{Id}_B(\Phi_{AB}) = \frac{\mathbb{1}_n}{n} \otimes \text{tr}_A(\Phi_{AB}), \\ & \text{Id}_{A_1} \otimes \text{tr}_{A_2} \otimes \text{Id}_B(\Phi_{AB}) = \frac{\mathbb{1}_n}{n} \otimes \text{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.65}$$

where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^n \otimes \mathbb{C}^n$. By taking advantage of the complete hierarchy, if for some N there is $\xi_N \leq 1/n$ or $\xi_N^T \leq 1/n$, then ρ is unfaithful. In practice, it is already enough to take ξ_2^T for verifying the unfaithfulness of some states that are not detectable by any of the known methods [108, 114]. An explicit example for $n = 4$ is

$$\rho = \frac{p}{16}\mathbb{1}_4 \otimes \mathbb{1}_4 + \frac{1-p}{2}|x\rangle\langle x| + \frac{1-p}{2}|y\rangle\langle y|, \tag{3.66}$$

where

$$\begin{aligned} |x\rangle &= \frac{1}{\sqrt{10}} \sum_{\alpha=1}^{4} \sqrt{\alpha}\,|\alpha\alpha\rangle, \\ |y\rangle &= \frac{1}{\sqrt{10}} \sum_{\alpha=1}^{4} \beta_4^\alpha \sqrt{5-\alpha}\,|\alpha\alpha\rangle, \end{aligned} \tag{3.67}$$

with $p = 23/40$ and $\beta_4 = (1+\mathrm{i})/\sqrt{2}$. For this state, the SDP relaxation of Eq. (3.65) gives the upper bound $\xi_2^T = 0.24888 < 1/4$, which matches the lower bound from

gradient search and is strictly better than the best known upper bound $\xi_1 = 0.25063 > 1/4$ from Ref. [108].

3.3.2 Majorization uncertainty relations

As described in Section 2.9, majorization relations provide a way to compare how chaotic probability distributions are independent from a specific entropy. Let us consider two POVM measurements with effects $\{E_j\}$ and $\{F_j\}$. For each state ρ, they give rise to two probability vectors $p_j = \operatorname{Tr}\rho E_j$ and $q_j = \operatorname{Tr}\rho F_j$. We try to find a minimal $\boldsymbol{\omega}$ such that $\boldsymbol{p}\otimes\boldsymbol{q}$ is majorized by $\boldsymbol{\omega}$, i.e. $\boldsymbol{p}\otimes\boldsymbol{q} \prec \boldsymbol{\omega}$, for all states ρ. Minimality means that, for any other $\boldsymbol{\omega}'$ that satisfies this condition, it holds that $\boldsymbol{\omega} \prec \boldsymbol{\omega}'$, . Then, such an $\boldsymbol{\omega}$ gives rise to a family of entropic uncertainty relations $S(\boldsymbol{p}\otimes\boldsymbol{q}) = S(\boldsymbol{p}) + S(\boldsymbol{q}) \geq S(\boldsymbol{\omega})$ for any additive entropy S.

Via the definition of majorization, the search can be reduced to the optimization problem [117, 118]

$$\omega_k = \max_{T\subset[n]\times[m],|S|=k} \max_{\rho} \sum_{(i,j)\in T} \operatorname{Tr}(\rho E_i)\operatorname{Tr}(\rho F_j) \tag{3.68}$$

$$= \max_{T\subset[n]\times[m],|S|=k} \max_{\rho} \sum_{(i,j)\in T} \operatorname{Tr}[(\rho\otimes\rho)(E_i\otimes F_j)], \tag{3.69}$$

where $[n] = \{1,\ldots,N\}$ with N being the number of effects E_j and similar for $[m]$ and we have $k = 1,2,\ldots,\min(N,M)-1$ because for larger k, ω_k is obviously 1. Since the outer maximization is over a finite set, we can focus on the inner optimization which is of the two-copy type central to Chapters 3 and 4.

The easily computable general approximations of ω_k in Refs. [117, 118] are indeed exact for $k = 1,2$. It turns out that the approximations found in Refs. [117, 118] are indeed exact for $k = 1,2$. For $k \geq 3$, however, we can apply the hierarchy in Theorem 3.2 to obtain better and better bounds. An interesting example is to consider ω_3 for measurements in the computational and the Fourier basis in dimension $d = 4$. In this case, for $T = \{(i,j),(i,k),(i,l)\}$, the method in Refs. [117, 118] gives a maximum value of $(7+4\sqrt{3})/16$ which is realized by the state $|\psi\rangle = \sqrt{a/3}\,|0\rangle + \sqrt{a/3}\,|1\rangle + \sqrt{a/3}\,|2\rangle + \sqrt{1-a}\,|3\rangle$ with $a = (2+\sqrt{3})/4$. Due to the symmetry between the bases, the same result is obtained for $T = \{(j,i),(k,i),(l,i)\}$. Finally, for $T = \{(i,j),(i,k),(l,j)\}$, their method yields a value of 1 implying overall only a trivial bound for ω_3. Our extension method, on the other hand, gives a certified numerical value of 0.84038 for this T, probably realized by the state $|\psi\rangle = \sqrt{a}\,|0\rangle + \sqrt{(1-a-b)/2}\,|1\rangle + \sqrt{b}\,|2\rangle + \sqrt{(1-a-b)/2}\,|3\rangle$ with $a = (33 - 5\cdot 3^{1/3} + 9\cdot 3^{2/3})/68$ and $b = (27 + 33\cdot 3^{1/3} - 5\cdot$

$3^{2/3})/204$ yielding a value of $(121 + 27 \cdot 3^{1/3} + 33 \cdot 3^{2/3})/272$. Although the semidefinite program proves this value only up to numerical precision, at least without being able to guess the analytical expression for the optimum of the dual problem, together with the analytical result for the $T = \{(i,j),(i,k),(i,l)\}$ and $T = \{(j,i),(k,i),(l,i)\}$, this proves that indeed $\omega_3 = (7 + 4\sqrt{3})/16$. Thus, we established the minimal

$$\boldsymbol{\omega} = \left[\frac{9}{16}, \frac{1}{8}\left(3 + 2\sqrt{2}\right), \frac{1}{16}\left(7 + 4\sqrt{3}\right), 1\right] \tag{3.70}$$

for measurements in the computational and the Fourier basis.

Interestingly, further investigations indicate that there might always be an optimal state which is pure for two measurements but presumably not for more POVMs. It is worthwhile to examine this observation in more detail and also consider the special case of projective measurements.

3.3.3 Gram matrix and orthonormal representation

Let $|a_i\rangle \in \mathbb{F}^k$ ($\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$) for $i = 1, 2, \ldots, n$ be a sequence of vectors, then the Gram matrix defined as $\Gamma = [\langle a_i|a_j\rangle]_{i,j=1}^n$ satisfies $\Gamma \geq 0$ and $\text{rank}(\Gamma) \leq k$. The converse is also true in the sense that if an $n \times n$ matrix in $\mathbb{F}^{n\times n}$ satisfies $\Gamma \geq 0$ and $\text{rank}(\Gamma) \leq k$, then there exist $|a_i\rangle \in \mathbb{F}^k$ for $i = 1, 2, \ldots, n$, such that $\Gamma_{ij} = \langle a_i|a_j\rangle$ [119]. This correspondence can trigger many applications of the rank-constrained optimization. For example, it can be used to bound the minimum dimension of the orthonormal representation of graphs.

In graph theory, a graph G is denoted by a pair (V, E), where V is the set of vertices, and E is the set of edges connecting pairs of vertices. For a graph $G = (V, E)$, an orthonormal representation is a set of normalized vectors $\left\{ |a_i\rangle \in \mathbb{F}^k \;\middle|\; i \in V \right\}$, such that $\langle a_i|a_j\rangle = 0$ if $\{i,j\} \notin E$ [119]. The minimum dimension problem is to find the smallest number k such that an orthonormal representation exists. This is not only an important quantity in classical information theory [119], but also widely used in quantum information theory. For example, it is a crucial quantity in quantum contextuality theory [120, 121], and can be directly used for contextuality-based dimension witness [143]. Note that in quantum contextuality, the definition of orthonormal representations is slightly different, where the adjacent instead of the nonadjacent vertices are required to be orthogonal to each other, i.e., $\langle a_i|a_j\rangle = 0$ if $\{i,j\} \in E$. In the following, we will use the standard definition in graph theory. All results can be trivially adapted to the alternative definition by considering the complement graph.

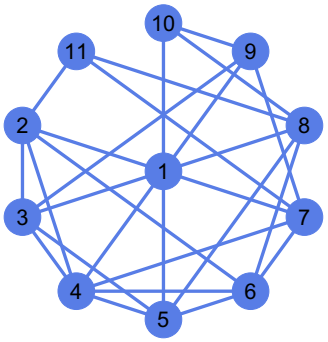


FIGURE 3.3: [99] For this 11-vertex graph, one obtains that $\vartheta(G) = 4$ (up to a numerical error smaller than 10^{-100}) using the standard primal and dual problem of the Lovász ϑ-function's SDP characterization [103] and hence, a lower bound of 4 for the minimal dimension. In contrast, our PPT relaxation of Eq. (3.73) can already exclude both real and complex orthonormal representations in dimension 4.

By taking advantage of the Gram matrix, the problem of the minimum dimension of the orthonormal representation [119] can be expressed as

$$\begin{aligned} \min_{\Gamma} \quad & k \\ \text{s.t.} \quad & \Delta(\Gamma) = \mathbb{1}_n,\ \Gamma_{ij} = 0 \ \ \forall\ \{i,j\} \notin E, \\ & \Gamma \geq 0,\ \operatorname{rank}(\Gamma) \leq k, \end{aligned} \tag{3.71}$$

where $G = (V, E)$ is a graph with $|V| = n$ vertices, E is the set of edges, $\Delta(\cdot)$ denotes the map of eliminating all off-diagonal elements of a matrix (completely dephasing map), and $\Gamma \in \mathbb{R}^{n\times n}$ or $\Gamma \in \mathbb{C}^{n\times n}$ corresponds to the real or complex representation. Let W be the adjacency matrix of G, i.e., $W_{ij} = 1$ if $\{i,j\} \in E$ and $W_{ij} = 0$ otherwise, then the first two constraints in Eq. (3.71) can also be written as $(1 - W_{ij})\Gamma_{ij} = \delta_{ij}$, i.e.,

$$\Lambda(\Gamma) := (\mathbb{J}_n - W) \odot \Gamma = \mathbb{1}_n, \tag{3.72}$$

where $\mathbb{J}_n$ is the $n \times n$ matrix with all elements being one and $[X \odot Y]_{ij} = X_{ij}Y_{ij}$ is the Hadamard product of matrices. Then, the existence of a k-dimensional orthonormal representation is equivalent to the following feasibility problem

$$\begin{aligned} \text{find} \quad & \Phi_{AB} \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \operatorname{tr}(\Phi_{AB}) = n, \\ & V_{AB}\Phi_{AB} = \Phi_{AB},\ \left(\Phi_{AB}^{T_A} = \Phi_{AB}\right), \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = \frac{1}{n}\mathbb{1}_n \otimes \operatorname{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.73}$$

where $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^n \otimes \mathbb{C}^k$ $(\mathbb{R}^n \otimes \mathbb{R}^k)$, $\widetilde{\Lambda}(\cdot) = \Lambda[\operatorname{tr}_2(\cdot)]$, and the extra constraint $\Phi_{AB}^{T_A} = \Phi_{AB}$ is for the case that $\mathbb{F} = \mathbb{R}$ only. Note that the inherent symmetry

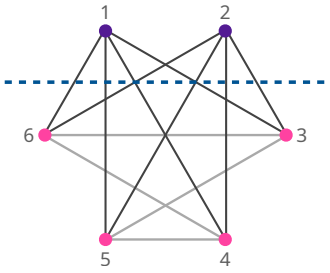


FIGURE 3.4: [99] The Max-Cut problem for a graph is to find a cut, i.e., a bipartition, such that the number of edges that cross the cut is maximized. For the graph shown in the figure, the Max-Cut is 8 (achieved by the cut $1, 2$ versus $3, 4, 5, 6$), which matches our SDP relaxation $\xi_2 = 8$, while the Goemans-Williamson method yields only an upper bound of $\xi_1 = 9$.

presented in Section 3.2.4 can be used to simplify the SDP relaxations.

The Lovász ϑ-function defined by

$$\vartheta(G) = \min_{\{|a_i\rangle\}_{i\in V}, |c\rangle} \max_{i\in V} \frac{1}{|\langle c|a_i\rangle|^2}, \tag{3.74}$$

where the $|a_i\rangle$ form an orthonormal representation and $|c\rangle$ is a unit vector, is probably the best-known way to obtain a lower bound on the minimal dimension of orthonormal representations. Note that the value of the Lovász ϑ-function is independent of whether the orthonormal representation is real or complex [119]. For any k-dimensional orthonormal representation $|a_i\rangle$, also $|a_i\rangle \otimes |a_i^*\rangle$ form an orthonormal representation and, with $|c\rangle = \frac{1}{\sqrt{k}} \sum_{\alpha=1}^{k} |\alpha\rangle \otimes |\alpha\rangle$, the bound $k \geq \vartheta(G)$ is readily obtained from Eq. (3.74). Our method can provide a better bound even for small graphs; see Fig. 3.3.

3.3.4 Max-Cut problem

The Max-Cut problem is among the best-known rank-constrained optimization problems [122] and also draws a lot of interest in quantum computing [144, 145]. Given a graph $G = (V, E)$, the Max-Cut problem is to find a cut, i.e., a bipartition of the vertices (S, S^c), where $S^c = V \setminus S$, that maximizes the number of edges between S and S^c; see Fig. 3.4. A significant breakthrough for the Max-Cut problem was the work by Goemans and Williamson [122], in which they showed that the Max-Cut problem can be written as the following rank-constrained optimization

$$\begin{aligned} \max_{\rho} \quad & \frac{1}{4} \operatorname{tr}[W(\mathbb{J}_n - \rho)] \\ \text{s.t.} \quad & \Delta(\rho) = \mathbb{1}_n,\ \rho \geq 0,\ \operatorname{rank}(\rho) = 1, \end{aligned} \tag{3.75}$$

where $n = |V|$, $\rho \in \mathbb{R}^{n\times n}$, $\mathbb{J}_n$ is the $n \times n$ matrix with all elements being one, and W is the adjacency matrix of G. To see why the Max-Cut problem is equivalent to Eq. (3.75), we denote a cut with the binary vector $\boldsymbol{x} \in \{-1, 1\}^n$ such that $x_i = 1$ if $i \in S$ and $x_i = -1$ if $i \in S^c$ and let $\rho = \boldsymbol{x}\boldsymbol{x}^T$, then the number of edges between S and S^c is $\frac{1}{4}\sum_{(i,j)\in E}(1 - x_i x_j)$, which is equal to the objective function in Eq. (3.75). Furthermore, the set of all cuts $\rho = \boldsymbol{x}\boldsymbol{x}^T$ can be fully characterized by the constraints in Eq. (3.75). The idea of the Goemans-Williamson approximation is to remove the rank constraint in Eq. (3.75) and solve the resulting SDP relaxation, which gives an upper bound ξ_1 for the Max-Cut problem.

In the following, we show how our method can give a better estimate compared to the Goemans-Williamson approximation. By noting that we can add a redundant constraint $\mathrm{tr}(\rho) = n$, Theorem 3.5 implies that the Max-Cut problem is equivalent to the following conic program

$$\begin{aligned} \max_{\Phi_{AB}} \quad & \frac{1}{4}\,\mathrm{tr}[W\mathbb{J}_n] - \frac{1}{4}\,\mathrm{tr}[W_A \otimes \mathbb{1}_B \Phi_{AB}] \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = n, \\ & V_{AB}\Phi_{AB} = \Phi_{AB},\ \Phi_{AB}^{T_A} = \Phi_{AB}, \\ & \Delta_A \otimes \mathrm{Id}_B(\Phi_{AB}) = \frac{1}{n}\mathbb{1}_n \otimes \mathrm{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.76}$$

where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{R}^n$. Consequently, a complete hierarchy of SDPs can be constructed from Theorem 3.6.

We have tested the SDP relaxation ξ_2 (replacing $\Phi_{AB} \in \mathrm{SEP}$ by $\Phi_{AB} \geq 0$) with some random graphs (randomly generated adjacency matrices). Let us discuss the largest two graphs that we have tested. For a 64-vertex graph with 419 edges, the Goemans-Williamson method gives the upper bound $\lfloor \xi_1 \rfloor = 299$, instead $\xi_2 = 287$. For the 72-vertex graph with 475 edges, the Goemans-Williamson method gives the upper bound $\lfloor \xi_1 \rfloor = 335$, instead $\xi_2 = 321$. Furthermore, the optimal Φ_{AB} also shows that the upper bound ξ_2 in these two cases are achievable. Hence, ξ_2 gives the exact solution to the Max-Cut problem in these examples. Actually, for all the graphs that we have tested, ξ_2 already gives the exact solution of the Max-Cut problem. At last, we would like to mention that, although our method gives a much better bound, it is more costly than the Goemans-Williamson method. For example, the size of the matrix grows quadratically on the number of vertices for ξ_2, compared to only growing linearly for the Goemans-Williamson method.

3.3.5 Pseudo-Boolean optimization

Similar to the Max-Cut problem, we can apply the method to general optimization of a real-valued function over Boolean variables. These so-called pseudo-Boolean optimization problems find wide applications in, for example, statistical mechanics, computer science, discrete mathematics, and economics (see [123] and references therein). As a demonstration, we consider the quadratic pseudo-Boolean optimization

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & x_i = \pm 1, \end{aligned} \tag{3.77}$$

where $Q \in \mathbb{R}^{(n-1)\times(n-1)}$, $\boldsymbol{c} \in \mathbb{R}^{n-1}$, and $\boldsymbol{x}^T = [x_1, x_2, \ldots, x_{n-1}]$; higher-order cases can be obtained by reducing to quadratic forms [123] or applying the results of Section 3.5.4. Notably, performing quadratic pseudo-Boolean optimization problems with noisy intermediate-scale quantum computers has drawn a lot of research interest [126–129]. So, the following method may be used to characterize benchmarks of such devices.

The quadratic pseudo-Boolean optimization problem can also be written as a rank-constrained optimization. The basic idea is to write ρ as an $n \times n$ matrix

$$\rho = \begin{bmatrix} \boldsymbol{x}\boldsymbol{x}^T & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^T & 1 \end{bmatrix}. \tag{3.78}$$

Further, we define L as

$$L = \begin{bmatrix} Q & \frac{1}{2}\boldsymbol{c} \\ \frac{1}{2}\boldsymbol{c}^T & 0 \end{bmatrix}. \tag{3.79}$$

Then, the optimization problem in Eq. (3.77) can be written as the following rank-constrained SDP

$$\begin{aligned} \max_{\rho} \quad & \mathrm{tr}(L\rho) \\ \text{s.t.} \quad & \Delta(\rho) = \mathbb{1}_n,\ \rho \geq 0,\ \mathrm{rank}(\rho) = 1, \end{aligned} \tag{3.80}$$

which is of a similar form as in Eq. (3.75). By Theorem 3.5, the quadratic pseudo-Boolean optimization problem is equivalent to the conic program in Eq. (3.76) with the objective function replaced by $\mathrm{tr}(L_A \otimes \mathbb{1}_B \Phi_{AB})$.

To illustrate the performance of our method, we consider the Boolean least squares optimization, i.e.,

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 \\ \text{s.t.} \quad & x_i = \pm 1. \end{aligned} \tag{3.81}$$

We have tested our SDP relaxation ξ_2, compared to the widely-used SDP relaxation ξ_1 [146], for 1000 random matrices $A \in \mathbb{R}^{40\times 30}$ and vectors $\boldsymbol{b} \in \mathbb{R}^{40}$ with elements independently normally distributed. For this size, the optimal value ξ can still be obtained by brute force. In most cases, the optimum is reached by ξ_2 while there is a significant gap between the optimal value ξ and ξ_1. More precisely, for the 1000 random samples, we obtain an average ratio of $\langle \xi_2/\xi \rangle = 99.93\%$ in contrast to $\langle \xi_1/\xi \rangle = 49.32\%$. Note that in Eq. (3.81), a the minimization is considered which means that the ξ_N provide lower bounds for ξ instead of upper bounds.

In passing, we would like to mention that Lasserre's hierarchy for polynomial optimization can also be used for the pseudo-Boolean optimization [104, 105], however, the construction of the SDP hierarchy is much easier with our method. Moreover, our method also makes it more convenient to utilize the symmetry of the optimization problem, which usually plays a crucial role when solving large-scale problems.

3.4 Arbitrary-precision certified semidefinite programming

Reading Section 3.3.3, one might stumble over the remarkably high precision obtained for the Lovász ϑ-function at the graph given in Fig. 3.3, namely $\vartheta(G) = 4 \pm 10^{-100}$. As indicated in the text, we find feasible points of the standard primal and dual problem of the Lovász ϑ-function's SDP characterization [103] which provide a certified lower and upper bound to $\vartheta(G)$, respectively. To do so, we use the arbitrary-precision SDP solver SDPA-GMP [147–149] to obtain highly-accurate numerical solutions to the SDPs.

Then, using fractions, we find exact feasible points close to the numerical solution. To ensure positive semidefiniteness, we examine the characteristic polynomial $p(\lambda) = \det(\lambda \mathbb{1} - \Phi)$ whose roots are the eigenvalues of the $K \times K$-matrix Φ. An efficient way to compute the coefficients c_k in the decomposition

$$p(\lambda) = \sum_{k=0}^{K} c_k \lambda^k \tag{3.82}$$

analytically is to use the Faddeev–LeVerrier algorithm [150, 151]. Using Descartes' rule of signs [152], the positivity of Φ is determined by the signs of the coefficients c_k [153]. More precisely, we consider the characteristic polynomial of $-\Phi$ with coefficients $\tilde{c}_k$ for convenience. If $\tilde{c}_k \geq 0$ for all k, then $\sum_{k=0}^{K} \tilde{c}_k \lambda^k > 0$ for $\lambda > 0$ and hence, $-\Phi$ is negative semidefinite, i.e., Φ is positive semidefinite. On the other hand, writing $\tilde{p}(\lambda)$

as (note that $\tilde{c}_K = 1$)

$$\tilde{p}(\lambda) = \prod_{j=0}^{K} (\lambda - \tilde{\lambda}_j) = \sum_{k=0}^{K} \tilde{c}_k \lambda^k, \tag{3.83}$$

where $\tilde{\lambda}_j$ are the eigenvalues of $-\Phi$, negative semidefiniteness of $-\Phi$ implies that $\tilde{c}_k \geq 0$ by the expansion of the product because $-\tilde{\lambda}_j \geq 0$. Thus, Φ is positive semidefinite if, and only if, all the $\tilde{c}_k$ are nonnegative.

One possibility to find exact feasible points from approximate solutions uses the standard form for SDPs in Eq. (2.76). In this form, the constraints are incorporated into the basis F_j. Then, finding a linear combination $\sum_{j>0} \mu_j F_j > 0$, we can obtain an exact feasible point from an almost semidefinite approximate solution by adding a small multiple of this linear combination and hence, a certified upper of lower bound for the optimal value. Sometimes, the exact solution can even be guessed from the numerical result and verified as described. In this way, we were able to certify that indeed $\vartheta(G) \geq 4$, however, we could not find a certificate for $\vartheta(G) \leq 4$ and hence, a small uncertainty remains.

3.5 More general results on rank-constrained optimization

In this section, we consider extensions of the problem in Eq. (3.1) and general cases of rank-constrained optimization. For simplicity, we only consider the optimization over complex matrices. All results can be similarly applied to the optimization over real matrices by adding the partial-transpose-invariance constraint $\Phi_{AB}^{T_A} = \Phi_{AB}$ or $\Phi_{AB\cdots Z}^{T_A} = \Phi_{AB\cdots Z}$.

3.5.1 Inequality constraints

Starting from the following rank-constrained SDP with inequality constraints

$$\begin{aligned} \max_{\rho} \quad & \operatorname{tr}(X\rho) \\ \text{s.t.} \quad & \Lambda(\rho) \leq Y,\ \operatorname{tr}(\rho) = 1, \\ & \rho \geq 0,\ \operatorname{rank}(\rho) \leq k, \end{aligned} \tag{3.84}$$

where Λ is a Hermiticity-preserving map [154], we can still define the feasible region $\mathcal{F}$ in $\mathbb{C}^{n\times n}$ and its purification $\mathcal{P}$ in $\mathbb{C}^{nk\times nk}$, similar to Eqs. (3.2, 3.3), as

$$\mathcal{F} = \{\rho \mid \Lambda(\rho) \leq Y, \operatorname{tr}(\rho) = 1, \rho \geq 0, \operatorname{rank}(\rho) \leq k\}, \tag{3.85}$$

$$\mathcal{P} = \{ |\varphi\rangle\langle\varphi| \mid \widetilde{\Lambda}(|\varphi\rangle\langle\varphi|) \leq Y,\ \langle\varphi|\varphi\rangle = 1\}, \tag{3.86}$$

where $\widetilde{\Lambda}(\cdot) = \Lambda[\mathrm{tr}_2(\cdot)]$. Again, we denote the solution of Eq. (3.84) as ξ. In this case, the proof of Theorem 3.1 does not work, because although the constraints

$$\begin{aligned} &\Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1,\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ &\widetilde{\Lambda} \otimes \mathrm{Id}_B(\Phi_{AB}) \leq Y \otimes \mathrm{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.87}$$

still provide a necessary condition for $\mathrm{tr}_B(\Phi_{AB}) \in \mathcal{S} := \mathrm{conv}(\mathcal{P})$, they are no longer sufficient. This is because contrary to the equality case, the pure states in the decomposition of Φ_{AB} can no longer be guaranteed to be in $\mathcal{P}$ for the inequality case. However, the complete hierarchy analogously to Eq. (3.22) still provides the exact solution ξ.

Theorem 3.9. *For $\mathbb{F} = \mathbb{C}$, let ξ be the solution of the rank-constrained SDP in Eq. (3.84). Then, for any N, ξ is upper bounded by the solution ξ_N of the following SDP hierarchy*

$$\begin{aligned} \max_{\Phi_{AB\cdots Z}}\ & \mathrm{tr}(\widetilde{X}_A \otimes \mathbb{1}_{B\cdots Z}\Phi_{AB\cdots Z}) \\ \text{s.t.}\ & \Phi_{AB\cdots Z} \geq 0,\ \mathrm{tr}(\Phi_{AB\cdots Z}) = 1, \\ & P_N^+\Phi_{AB\cdots Z}P_N^+ = \Phi_{AB\cdots Z}, \\ & \widetilde{\Lambda} \otimes \mathrm{Id}_{B\cdots Z}(\Phi_{AB\cdots Z}) \leq Y \otimes \mathrm{tr}_A(\Phi_{AB\cdots Z}). \end{aligned} \tag{3.88}$$

Furthermore, the SDP hierarchy is complete, i.e., $\xi_{N+1} \leq \xi_N$ and $\lim_{N\to+\infty} \xi_N = \xi$.

Similarly, any criterion for the full separability of $\Phi_{AB\cdots Z}$ or the unnormalized state $Y \otimes \mathrm{tr}_A(\Phi_{AB\cdots Z}) - \widetilde{\Lambda} \otimes \mathrm{Id}_{B\cdots Z}(\Phi_{AB\cdots Z})$, such as the PPT criterion, can be added to the optimization in Eq. (3.88), which can give a better upper bound for the optimization in Eq. (3.84).

For simplicity, we only present the intuition of the proof of Theorem 3.9 here; see [99] for a rigorous proof. The property $\xi_{N+1} \leq \xi_N$ follows from the hierarchical property that if $\Phi_{AB\cdots ZZ'}$ is within the feasible region of level $N+1$, then $\Phi_{AB\cdots Z} = \mathrm{tr}_{Z'}(\Phi_{AB\cdots ZZ'})$ is within the feasible region of level N.

For the convergence property, we consider a separable variant of the optimization in Eq. (3.88) by replacing $\Phi_{AB\cdots Z} \geq 0$ with $\Phi_{AB\cdots Z} \in \mathrm{SEP}$, and denote the corresponding solutions as $\widetilde{\xi}_N$, i.e., add a tilde to distinguish the solution with the separability constraint from the original ξ_N. Then, the quantum de Finetti theorem [135, 138] implies that

$$\lim_{N\to+\infty} \widetilde{\xi}_N = \lim_{N\to+\infty} \xi_N. \tag{3.89}$$

Now, we assume that the $\widetilde{\xi}_N$ are achieved by the separable states

$$\widetilde{\Phi}_{AB\cdots Z} = \int_\psi f_N(\psi) \left|\psi\right\rangle \left\langle\psi\right|^{\otimes N} \mathrm{d}\psi, \tag{3.90}$$

where the $f_N(\psi)\mathrm{d}\psi$ are N-dependent probability distributions, and $\mathrm{d}\psi$ denotes the normalized uniform distribution. As the set of probability distributions on a compact set is also compact in the weak topology [155], we can take $f_\infty(\psi)\mathrm{d}\psi$ as a limit point of $f_N(\psi)\mathrm{d}\psi$. Thus, we get an N-independent probability distribution $f_\infty(\psi)\mathrm{d}\psi$. Let

$$\widetilde{\Phi}^{\infty}_{AB\cdots Z} = \int_{\psi} f_\infty(\psi)\, |\psi\rangle\langle\psi|^{\otimes N}\, \mathrm{d}\psi, \tag{3.91}$$

which satisfies all the constraints in Eq. (3.88) for arbitrary N by the hierarchical property, and moreover

$$\lim_{N\to+\infty} \widetilde{\xi}_N = \lim_{N\to+\infty} \mathrm{tr}\left(\widetilde{X}_A \widetilde{\Phi}^N_A\right) = \mathrm{tr}\left(\widetilde{X}_A \widetilde{\Phi}^\infty_A\right), \tag{3.92}$$

where

$$\widetilde{\Phi}^N_A = \mathrm{tr}_{B\cdots Z}\left(\widetilde{\Phi}_{AB\cdots Z}\right) = \int_{\psi} f_N(\psi)\, |\psi\rangle\langle\psi|\, \mathrm{d}\psi, \tag{3.93}$$

$$\widetilde{\Phi}^\infty_A = \mathrm{tr}_{B\cdots Z}\left(\widetilde{\Phi}^\infty_{AB\cdots Z}\right) = \int_{\psi} f_\infty(\psi)\, |\psi\rangle\langle\psi|\, \mathrm{d}\psi. \tag{3.94}$$

By Eq. (3.89), to prove that $\lim_{N\to+\infty}\xi_N = \xi$, we only need to show that $\widetilde{\Phi}^\infty_A \in \mathrm{conv}(\mathcal{P})$. To this end, it is sufficient to show that $Y_\varphi := \widetilde{\Lambda}(|\varphi\rangle\langle\varphi|) \leq Y$ whenever $f_\infty(\varphi) \neq 0$. By plugging Eq. (3.91) into the last constraint in Eq. (3.88), we get that for arbitrary N

$$\int_{\psi} f_\infty(\psi)(Y - Y_\psi) \otimes |\psi\rangle\langle\psi|^{\otimes N}\, \mathrm{d}\psi \geq 0, \tag{3.95}$$

which implies that

$$\frac{\int_\psi f_\infty(\psi)(Y - Y_\psi)|\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi}{\int_\psi |\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi} \geq 0 \tag{3.96}$$

for any $|\varphi\rangle$ and N. Note that for any $\varepsilon > 0$, the integral over the complement of the ε-ball $B^c_\varphi(\varepsilon) := \left\{ |\psi\rangle\langle\psi| \;\middle|\; |\langle\varphi|\psi\rangle|^2 \leq 1-\varepsilon \right\}$ is

$$\lim_{N\to+\infty} \frac{\int_{B^c_\varphi(\varepsilon)} |\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi}{\int_\psi |\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi} = 0, \tag{3.97}$$

because while the numerator decreases to zero exponentially with N, the denominator $\int_\psi |\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi = 1/\dim(P^+_N)$ decreases only polynomially according to Eq. (3.32) and the relation $\int_\psi |\psi\rangle\langle\psi|^{\otimes N}\,\mathrm{d}\psi = P^+_N/\dim(P^+_N)$ [154]. Hence,

$$\lim_{N\to+\infty} \frac{|\langle\varphi|\psi\rangle|^{2N}}{\int_\psi |\langle\varphi|\psi\rangle|^{2N}\mathrm{d}\psi} = \delta(\psi - \varphi), \tag{3.98}$$

where $\delta(\cdot)$ is the Dirac-delta function. Then, in the limit $N \to +\infty$, Eq. (3.96) gives that

$$\int_{\psi} f_{\infty}(\psi)(Y - Y_{\psi})\delta(\psi - \varphi)\mathrm{d}\psi = f_{\infty}(\varphi)(Y - Y_{\varphi}) \geq 0, \tag{3.99}$$

and hence, $Y_{\varphi} \leq Y$ when $f_{\infty}(\varphi) \neq 0$.

3.5.2 Non-positive-semidefinite variables

Second, we study the rank-constrained optimization for non-positive-semidefinite and even non-square matrices. Consider the rank-constrained optimization

$$\begin{aligned} \max_{\omega} \quad & \mathrm{tr}(X\omega) + \mathrm{tr}(X^{\dagger}\omega^{\dagger}) \\ \text{s.t.} \quad & \Lambda(\omega) = Y,\ \mathrm{rank}(\omega) \leq k, \end{aligned} \tag{3.100}$$

where $\omega \in \mathbb{C}^{m\times n}$, and the form of the objective function is chosen such that it is real-valued. Here, we impose an extra assumption that the optimal value can be attained on bounded ω, i.e., we consider the optimization

$$\begin{aligned} \max_{\omega} \quad & \mathrm{tr}(X\omega) + \mathrm{tr}(X^{\dagger}\omega^{\dagger}) \\ \text{s.t.} \quad & \Lambda(\omega) = Y,\ \|\omega\| \leq R,\ \mathrm{rank}(\omega) \leq k, \end{aligned} \tag{3.101}$$

where $\|\omega\| = \mathrm{tr}(\sqrt{\omega\omega^{\dagger}})$ is the trace norm of ω, and R is a suitably chosen bound depending on the actual problem. Especially, by taking $R \to +\infty$, Eq. (3.101) turns to Eq. (3.100). The key observation for solving Eq. (3.101) is the following lemma.

Lemma 3.10. *A matrix $\omega \in \mathbb{F}^{m\times n}$ ($\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$) satisfies that $\mathrm{rank}(\omega) \leq k$ and $\|\omega\| \leq R$ if, and only if, there exists $A \in \mathbb{F}^{m\times m}$ and $B \in \mathbb{F}^{n\times n}$ such that*

$$\Omega := \begin{bmatrix} A & \omega \\ \omega^{\dagger} & B \end{bmatrix} \tag{3.102}$$

satisfies that $\Omega \geq 0$, $\mathrm{tr}(\Omega) = 2R$, and $\mathrm{rank}(\Omega) \leq k$.

Proof. We take advantage of the following observations from elementary algebra:
Observation (i): For any $a, b, x \geq 0$ satisfying $ab \geq x^2$, we have that $a + b \geq 2x$.
Observation (ii): For any $y \geq x \geq 0$, there exist $a, b \geq 0$ such that $ab = x^2$ and $a + b = 2y$.

We first prove the sufficiency part. The rank statement is obvious because the rank of a submatrix is no larger than that of the whole matrix, i.e., $\mathrm{rank}(\omega) \leq \mathrm{rank}(\Omega) \leq k$.

Now, we show that $\Omega \geq 0$ and $\mathrm{tr}(\Omega) = 2R$ imply that $\|\omega\| \leq R$. Consider the singular value decomposition of ω

$$\omega = U^\dagger D V, \tag{3.103}$$

where U and V are unitary matrices, $D_{ii} \geq 0$, and $D_{ij} = 0$ for $i \neq j$. Furthermore, we have $\|\omega\| = \sum_i D_{ii}$. Let

$$\widetilde{\Omega} = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \Omega \begin{bmatrix} U^\dagger & 0 \\ 0 & V^\dagger \end{bmatrix} = \begin{bmatrix} UAU^\dagger & D \\ D^T & VBV^\dagger \end{bmatrix}. \tag{3.104}$$

Then, $\Omega \geq 0$ implies that

$$(UAU^\dagger)_{ii}(VBV^\dagger)_{ii} \geq D_{ii}^2. \tag{3.105}$$

Thus, Observation (i) implies that

$$(UAU^\dagger)_{ii} + (VBV^\dagger)_{ii} \geq 2D_{ii}, \tag{3.106}$$

whose summation gives

$$\mathrm{tr}(\Omega) = \sum_i \Big[(UAU^\dagger)_{ii} + (VBV^\dagger)_{ii}\Big] \geq 2\sum_i D_{ii} \tag{3.107}$$

Hence, $\mathrm{tr}(\Omega) = 2R$ implies $\|\omega\| = \sum_i D_{ii} \leq R$.

To prove the necessity part, we again consider the decomposition in Eq. (3.104). Then, $\mathrm{rank}(\omega) \leq k$ implies that $D_{ii} = 0$ when $i > k$. One can easily verify that Ω satisfies that $\Omega \geq 0$ and $\mathrm{rank}(\Omega) = \mathrm{rank}(\widetilde{\Omega}) \leq k$ when

$$\begin{aligned} &(UAU^\dagger)_{ij} = 0 \quad \text{for } i \neq j \text{ and } i = j > k, \\ &(VBV^\dagger)_{ij} = 0 \quad \text{for } i \neq j \text{ and } i = j > k, \\ &(UAU^\dagger)_{ii} \geq 0,\ (VBV^\dagger)_{ii} \geq 0 \quad \text{for } i = 1, 2, \ldots k, \\ &(UAU^\dagger)_{ii}(VBV^\dagger)_{ii} = D_{ii}^2 \quad \text{for } i = 1, 2, \ldots k. \end{aligned} \tag{3.108}$$

Then, Observation (ii) and the bound constraint $\|\omega\| = \sum_{i=1}^k D_{ii} \leq R$ imply that we can choose suitable $(UAU^\dagger)_{ii}$ and $(VBV^\dagger)_{ii}$ for $i = 1, 2, \ldots, k$ such that $\mathrm{tr}(\Omega) = \mathrm{tr}(\widetilde{\Omega}) = \sum_{i=1}^k [(UAU^\dagger)_{ii} + (VBV^\dagger)_{ii}] = 2R$. □

By taking advantage of Lemma 3.10, the optimization in Eq. (3.101) can be written as

$$\begin{aligned} \max_{\Omega} \quad & \mathrm{tr}(L\Omega) \\ \text{s.t.} \quad & \Lambda \circ P(\Omega) = Y,\ \mathrm{tr}(\Omega) = 2R, \\ & \Omega \geq 0,\ \mathrm{rank}(\Omega) \leq k, \end{aligned} \tag{3.109}$$

where

$$\Omega = \begin{bmatrix} A & \omega \\ \omega^\dagger & B \end{bmatrix},\ L = \begin{bmatrix} 0 & X^\dagger \\ X & 0 \end{bmatrix},\ P(\Omega) = \omega. \tag{3.110}$$

Then, after normalization, Eq. (3.109) is of the simple form given in Eq. (3.1). Thus, all the methods developed in Section 3.2 are directly applicable.

Furthermore, by applying the technique from Section 3.5.1, it is also possible to consider element-wise inequality constraints of the form $\Lambda(\omega) \preceq Y$ for the optimization in Eq. (3.100), where $\preceq$ denotes the element-wise comparison.

3.5.3 Unnormalized variables

Third, we consider rank-constrained semidefinite optimization without normalization constraint. Formally, we consider the general rank-constrained SDP

$$\begin{aligned} \max_{\rho} \quad & \mathrm{tr}(X\rho) \\ \text{s.t.} \quad & \Lambda(\rho) = Y,\ M(\rho) \leq Z, \\ & \rho \geq 0,\ \mathrm{rank}(\rho) \leq k, \end{aligned} \tag{3.111}$$

which contains both an equality constraint $\Lambda(\rho) = Y$ and an inequality constraint $M(\rho) \leq Z$.

The first method we can try is to find a matrix C such that $W := \Lambda^*(C) > 0$, where Λ^* is the dual/adjoint map of Λ [154]. If this is possible, we can add a redundant normalization-like constraint

$$\mathrm{tr}(W\rho) = w, \tag{3.112}$$

where $w = \mathrm{tr}(CY)$, which follows from $\Lambda(\rho) = Y$. The strictly-positive-definite property of W implies that $w > 0$, otherwise the problem is trivial ($\rho = 0$). Then, by applying the transformation $\widetilde{\rho} = w^{-1}\sqrt{W}\rho\sqrt{W}$, the general rank-constrained SDP is transformed to a form with normalization condition for $\widetilde{\rho}$. Thus, the methods in Section 3.2 and 3.5.1 are directly applicable.

In general, we can combine the techniques of the inequality constraint and the non-positive-semidefinite variable to tackle the problem. Again, we impose an extra assumption that the optimization can be attained on bounded ρ, i.e., we consider the optimization

$$\begin{aligned} \max_{\rho} \quad & \mathrm{tr}(X\rho) \\ \text{s.t.} \quad & \Lambda(\rho) = Y,\ M(\rho) \leq Z,\ \mathrm{tr}(\rho) \leq R, \\ & \rho \geq 0,\ \mathrm{rank}(\rho) \leq k, \end{aligned} \tag{3.113}$$

where R is a suitably chosen bound depending on the actual problem. By taking advantage of Lemma 3.10, the optimization in Eq. (3.113) can be written as

$$\begin{aligned} \max_{\Omega} \quad & \operatorname{tr}(L\Omega) \\ \text{s.t.} \quad & \Lambda \circ P(\Omega) = Y,\ M \circ P(\Omega) \leq Z,\ P(\Omega) \geq 0, \\ & \operatorname{tr}(\Omega) = 2R,\ \Omega \geq 0,\ \operatorname{rank}(\Omega) \leq k, \end{aligned} \tag{3.114}$$

where

$$\Omega = \begin{bmatrix} A & \rho \\ \rho & B \end{bmatrix},\ L = \frac{1}{2}\begin{bmatrix} 0 & X \\ X & 0 \end{bmatrix},\ P(\Omega) = \rho. \tag{3.115}$$

Eq. (3.114) is a rank-constrained SDP with normalization constraint. Thus, by applying the methods from Section 3.2 and Section 3.5.1, a complete SDP hierarchy can be constructed.

3.5.4 Quadratic optimization and beyond

Finally, we show that our method can also be used for (rank-constrained) quadratic and higher-order optimization. The key observation is that quadratic functions over ρ can be written as linear functions over $\rho \otimes \rho$. For example, we can rewrite

$$\begin{aligned} \operatorname{tr}(X\rho Y\rho) &= \frac{1}{2}\operatorname{tr}[\{V, X \otimes Y\}(\rho \otimes \rho)], \\ \operatorname{tr}(X\rho)\operatorname{tr}(Y\rho) &= \operatorname{tr}[(X \otimes Y)(\rho \otimes \rho)], \end{aligned} \tag{3.116}$$

where V is the swap operator, and the anti-commutator $\{\cdot, \cdot\}$ is taken to ensure Hermiticity. Thus, without loss of generality, we consider the following rank-constrained quadratic optimization

$$\begin{aligned} \max_{\rho} \quad & \operatorname{tr}[X_{A_1B_1}(\rho_{A_1} \otimes \rho_{B_1})] \\ \text{s.t.} \quad & \Lambda(\rho) = Y,\ \operatorname{tr}(\rho) = 1, \\ & \rho \geq 0,\ \operatorname{rank}(\rho) \leq k, \end{aligned} \tag{3.117}$$

where $\mathcal{H}_{A_1} = \mathcal{H}_{B_1} = \mathbb{C}^n$, ρ_{A_1} and ρ_{B_1} denote the same state ρ on $\mathcal{H}_{A_1}$ and $\mathcal{H}_{B_1}$, respectively, and $X_{A_1B_1}$ is some Hermitian operator on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$. The generalization to the general cases as in the previous subsections is obvious.

To solve Eq. (3.117), we consider $\mathcal{F}_2 := \{\rho \otimes \rho \mid \rho \in \mathcal{F}\}$ and $\mathcal{P}_2 := \{|\varphi\rangle\langle\varphi| \otimes |\varphi\rangle\langle\varphi| \mid |\varphi\rangle\langle\varphi| \in \mathcal{P}\}$; see Fig. 3.5. From the definitions in Eqs. (3.2, 3.3, 3.7), we have

$$\operatorname{tr}_{A_2B_2}(\mathcal{S}_2) = \operatorname{conv}\left[\operatorname{tr}_{A_2B_2}(\mathcal{P}_2)\right] = \operatorname{conv}(\mathcal{F}_2). \tag{3.118}$$

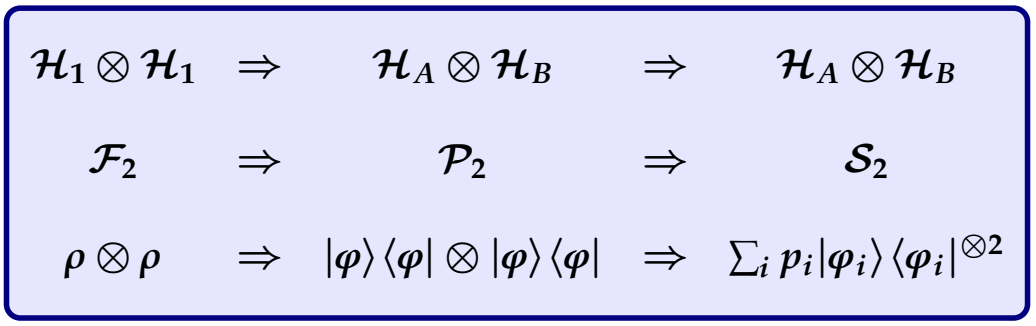


FIGURE 3.5: [99] An illustration of the relations between the two-party feasible region $\mathcal{F}_2$, the two-party purification $\mathcal{P}_2$, and the two-party extension $\mathcal{S}_2$.

As the set $\mathcal{S}_2$ is already fully characterized by Theorem 3.1, the rank-constrained quadratic optimization in Eq. (3.117) is equivalent to the conic program

$$\begin{aligned} \max_{\Phi_{AB}} \quad & \mathrm{tr}[\widetilde{X}_{AB}\Phi_{AB}] \\ \text{s.t.} \quad & \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1,\ V_{AB}\Phi_{AB} = \Phi_{AB}, \\ & \widetilde{\Lambda}_A \otimes \mathrm{Id}_B(\Phi_{AB}) = Y \otimes \mathrm{tr}_A(\Phi_{AB}), \end{aligned} \tag{3.119}$$

where $\widetilde{X}_{AB} = X_{A_1B_1} \otimes \mathbb{1}_{A_2B_2}$. Accordingly, a complete hierarchy can be constructed similarly as in Theorem 3.2.

We conclude this section with a few remarks. First, taking $k = n$ (i.e., taking the rank bound to be the dimension of ρ) corresponds to the quadratic program without rank constraint. Second, this method can be used for various uncertainty relations in quantum information, in which the minimization of the variance is automatically a quadratic program. Finally, the above procedure can be easily generalized to higher-order programming. The main idea is that all the results in Section 3.2 can be directly generalized to fully characterize

$$\mathcal{S}_N := \mathrm{conv}\left(\left\{ |\varphi\rangle\langle\varphi|^{\otimes N} \;\middle|\; |\varphi\rangle\langle\varphi| \in \mathcal{P} \right\}\right), \tag{3.120}$$

and $\mathcal{S}_N$ satisfies that $\mathrm{tr}_{A_2B_2\cdots Z_2}(\mathcal{S}_N) = \mathrm{conv}(\mathcal{F}_N)$, where $\mathcal{F}_N := \{\rho^{\otimes N} \mid \rho \in \mathcal{F}\}$. More precisely, recall that $\mathcal{P}$ is defined as

$$\mathcal{P} = \{ |\varphi\rangle\langle\varphi| \mid \widetilde{\Lambda}(|\varphi\rangle\langle\varphi|) = Y,\ \langle\varphi|\varphi\rangle = 1 \}. \tag{3.121}$$

Then, we show that $\Phi_{ABC\cdots Z} \in \mathcal{S}_N$ if, and only if,

$$\Phi_{ABC\cdots Z} \in \text{SEP},\ \text{tr}(\Phi_{ABC\cdots Z}) = 1, \tag{3.122}$$

$$P_N^+ \Phi_{ABC\cdots Z} P_N^+ = \Phi_{ABC\cdots Z}, \tag{3.123}$$

$$\widetilde{\Lambda}_A \otimes \text{Id}_{BC\cdots Z}(\Phi_{ABC\cdots Z}) = Y \otimes \text{tr}_A(\Phi_{ABC\cdots Z}). \tag{3.124}$$

Similarly to the case of $\mathcal{S}_2$, the constraints in Eqs. (3.122) and (3.123) imply that $\Phi_{ABC\cdots Z}$ is a separable state in the symmetric subspace, which always admits the form [133]

$$\Phi_{ABC\cdots Z} = \sum_i p_i \left|\varphi_i\right\rangle \left\langle\varphi_i\right|^{\otimes N}, \tag{3.125}$$

where the p_i form a probability distribution and the $|\varphi_i\rangle$ are normalized. Thus, Eq. (3.124) implies that

$$\mathcal{E}_A \otimes \mathcal{E}_B^\dagger \otimes \text{tr}_{C\cdots Z}(\Phi_{ABC\cdots Z}) = \sum_i p_i E_i \otimes E_i^\dagger = 0, \tag{3.126}$$

where $E_i = \mathcal{E}(|\varphi_i\rangle \langle\varphi_i|)$. Then, $|\varphi_i\rangle \langle\varphi_i| \in \mathcal{P}$ follows from Eqs. (3.17, 3.18), which prove that $\Phi_{ABC\cdots Z} \in \mathcal{S}_N$. Similarly, in the case of $\mathbb{F} = \mathbb{R}$, we only need to add the partial-transpose-invariant constraint

$$\Phi_{ABC\cdots Z}^{T_A} = \Phi_{ABC\cdots Z}. \tag{3.127}$$

Thus, (rank-constrained) higher-order optimizations over $\rho^{\otimes N}$ are fully characterizable with $\mathcal{S}_N$.

3.6 Conclusion

We have introduced a method to map SDPs with rank constraints to optimizations over separable quantum states. This result establishes a new connection between the theory of quantum entanglement, convex optimization, and rank-constrained semidefinite programming. While the DPS hierarchy characterizes entanglement via a hierarchy of semidefinite programs, we reformulate rank-constrained SDPs as conic optimizations using the cone of separable matrices, which again can be solved through a SDP hierarchy. Furthermore, we studied various examples and demonstrated the practical viability of our approach. In particular, we show how certified bounds with arbitrary precision can be obtained from the SDP relaxations. Since the quantum de Finetti theorem is indispensable for the completeness of our hierarchy, we commented on

the uniqueness of multi-copy decompositions, which is a common misunderstanding. Finally, we discussed several extensions to more general problems.

For further research, there are several interesting directions. First, concerning the presented method, a careful study of possible large-scale implementations, including the exploitation of possible symmetries, is desirable. This may finally shed new light on some of the examples presented here. Second, another promising method for solving the convex optimization problems in Theorems 3.1 and 3.5 is to consider the dual conic programs, which correspond to the optimization over entanglement witnesses. The benefit of this method will be that any feasible witness operator can provide a certified upper bound for the optimization problem. Third, on a broader perspective, it would be interesting to study other SDPs with additional constraints. An example is conditions in a product form, which frequently occur in quantum information due to the tensor product structure of the underlying Hilbert spaces. Finding SDP hierarchies for such problems will be very useful for the progress of this field.

4 A complete hierarchy for the pure-state marginal problem in quantum mechanics

Prerequisites

2.2 Quantum mechanics
2.5 Entanglement
2.7 The marginal problem and quantum codes
2.8 Semidefinite programming

4.1 Introduction

The main parts of this chapter have been published as Publication (C) [134]. Clarifying the relation between the whole and its parts is crucial for many problems in science. In quantum mechanics, this question manifests itself in the quantum marginal problem. For a given multiparticle quantum state $|\varphi\rangle$ it is straightforward to compute its marginals or reduced density matrices on some subsets of the particles. The reverse question, whether a given set of marginals is compatible with a global pure state, is, however, not easy to decide. Still, it is at the heart of many problems in quantum physics. Already in the early days it was a key motivation for Schrödinger to study entanglement [156], and it was recognized as a central problem in quantum chemistry [157]. There, often additional constraints play a role, e.g., if one considers fermionic systems. Then, the anti-symmetry leads to additional constraints on the marginals, generalizing the Pauli principle [158, 159]. A variation of the marginal problem is the question whether or not the marginals determine the global state uniquely or not [160–162]. This is relevant in condensed matter physics, where one may ask whether a state is the unique ground state of a local Hamiltonian [163, 164]. Many other cases, such as marginal problems for Gaussian and symmetric states [165, 166] and applications in quantum correlations [167], quantum causality [168], and interacting quantum many-body systems [169, 170] have been studied.

With the emergence of quantum information processing, various specifications of the marginal problem moved into the center of attention. In entanglement theory a pure two-particle state is maximally entangled, if the one-particle marginals are maximally mixed. Furthermore, absolutely maximally entangled (AME) states are multiparticle states that are maximally entangled for any bipartition. This makes them valuable ingredients for quantum information protocols [80, 81], but it turns out that AME states do not exist for arbitrary dimensions, as not always global states with the desired mixed marginals can be found [86, 171–173]. In fact, also states obeying weaker conditions, where a smaller number of marginals must be maximally mixed, are of fundamental interest, but in general it is open when such states exist [174–176]. More generally, the construction of quantum error correcting codes, which constitute fundamental building blocks in the design of quantum computer architectures [177–179], essentially amounts to the identification of subspaces of the total Hilbert space, where all states in this space obey certain marginal constraints. This establishes a connection to the AME problem, which consequently was announced to be one of the central problems in quantum information theory [180]. Although an AME$(4, 6)$ state, the specific instance which was asked for in Ref. [180], has been found recently [181], the general existence problem still remains unsolved.

In this chapter, we rewrite the marginal problem as an optimization problem over separable states, which can be seen as a special case of the optimization problem considered in Chapter 3. Here and in the following, the term marginal problem usually refers to the pure-state marginal problem in quantum mechanics. This rewriting allows us to transform the nonconvex and thus intractable purity constraint into a complete hierarchy of conditions for a set of marginals to be compatible with a global pure state. Each step is given by a semidefinite program (SDP), the conditions become stronger with each level, and a set of marginals comes from a global state, if, and only if, all steps are passed. There are at least two advantages of writing the marginal problem as an SDP hierarchy: First, the symmetry in the physical problem can be directly incorporated to drastically simplify the optimization (or feasibility) problem. Second, many known efficient and reliable algorithms are known for solving SDPs [89], which is in stark contrast to nonconvex optimization. To show the effectiveness of our method, we consider the existence problem of AME states. By employing the symmetry, we show that an AME state for a given number of particles and dimension exists, if, and only if, a specific two-party quantum state is separable. In fact, this allows us to reproduce nearly all previous results on the AME problem [17] with only few lines of calculation. Finally, we show that our approach can also be extended to study the existence problem of quantum codes.

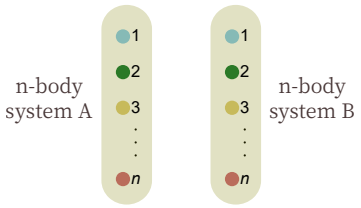


FIGURE 4.1: [134] An illustration of the two-party extension for the marginal problem. In the marginal problem one aims to characterize the pure states $|\varphi\rangle$ on n particles, which are compatible with given marginals. The key idea of our approach is to drop the purity constraint and to consider mixed states ρ with the given marginals. Then, the purity is enforced by considering a two-party extension Φ_{AB}.

4.2 Connecting the marginal problem with the separability problem

The formal definition of the marginal problem is the following: Consider an n-particle Hilbert space $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i$, and let $\mathcal{I} \subset \{I \mid I \subset [n] = \{1, 2, \ldots, n\}\}$ be some subsets of the particles, where the reduced states ρ_I are known marginals. Then, the problem reads

$$\begin{aligned} &\text{find} && |\varphi\rangle \\ &\text{s.t.} && \mathrm{tr}_{I^c}(|\varphi\rangle\langle\varphi|) = \rho_I,\ I \in \mathcal{I}. \end{aligned} \tag{4.1}$$

Here, $I^c = [n] \setminus I$ denotes the complement of the set I. Two facts are worth mentioning: First, if the global state $|\varphi\rangle\langle\varphi|$ is not required to be pure, then the quantum marginal problem without purity constraint is already an SDP. Second, if the given marginals are only one-body marginals, that is $\mathcal{I} = \{\{i\} \mid i \in [n]\}$, the marginals are non-overlapping and the problem in Eq. (4.1) was solved by Klyashko [79]. For overlapping marginals, however, the solution is more complicated, and this is what we want to discuss in this work.

The main idea of our method is to consider, for a given set of marginals, the compatible states and their extensions to two copies. Then, we can formulate the purity constraint using an SDP. We denote the two parties as A and B, and each of them owns an n-body quantum system; see Fig. 4.1.

Theorem 4.1. *There exists a pure quantum state $|\varphi\rangle$ that satisfies $\mathrm{tr}_{I^c}(|\varphi\rangle\langle\varphi|) = \rho_I$ for all $I \in \mathcal{I}$ if, and only if, the solution of the following convex optimization is equal to one,*

$$\max_{\Phi_{AB}} \quad \mathrm{tr}(V_{AB}\Phi_{AB}) = 1 \tag{4.2}$$

$$\text{s.t.} \quad \Phi_{AB} \in \mathrm{SEP},\ \mathrm{tr}(\Phi_{AB}) = 1, \tag{4.3}$$

$$\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \rho_I \otimes \mathrm{tr}_A(\Phi_{AB})\ \forall\ I \in \mathcal{I}. \tag{4.4}$$

where SEP *denotes the set of separable states w.r.t. the bipartition* $(A|B)$*,* A_{I^c} *denotes all subsystems* A_i *for* $i \in I^c$*, and similarly for* B_{I^c}*.*

This result follows directly from Theorem 3.1 for rank-1-constrained optimization by noting that the permuation matrix V_{AB} has eigenvalues ± 1.

Before proceeding further, we would like to add a few remarks. First, in Theorem 4.1 the constraint in Eq. (4.5) can be replaced by a weaker condition

$$\mathrm{tr}_{A_{I^c}B_{I^c}}(\Phi_{AB}) = \rho_I \otimes \rho_I \;\forall\; I \in \mathcal{I}, \tag{4.5}$$

This is because for any separable quantum state Φ_{AB} with $\mathrm{tr}(V_{AB}\Phi_{AB}) = 1$, Eq. (4.5) implies Eq. (4.4). More precisely, in this case, we can write Φ_{AB} as [133]

$$\Phi_{AB} = \sum_{\mu} p_\mu \left|\psi_\mu\right\rangle \left\langle\psi_\mu\right| \otimes \left|\psi_\mu\right\rangle \left\langle\psi_\mu\right|. \tag{4.6}$$

Then, with Eq. (4.5) we have that

$$\mathrm{tr}_{A_{I^c}B_{I^c}}(\Phi_{AB}) = \sum_{\mu} p_\mu \rho_I^{(\mu)} \otimes \rho_I^{(\mu)} = \rho_I \otimes \rho_I. \tag{4.7}$$

Furthermore, the following lemma implies that $\rho_I^{(\mu)} = \rho_I$ for all μ, and hence,

$$\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \rho_I \otimes \sum_{\mu} p_\mu \left|\psi_\mu\right\rangle \left\langle\psi_\mu\right| = \rho_I \otimes \mathrm{tr}_A(\Phi_{AB}). \tag{4.8}$$

Lemma 4.2. *Any state of the form* $\rho \otimes \rho$ *is an extreme point of the convex set* $\mathrm{conv}\{\rho \otimes \rho \mid \rho \geq 0,\ \mathrm{tr}(\rho) = 1\}$.

Proof. Suppose that

$$\rho \otimes \rho = \sum_{\mu} p_\mu \rho_\mu \otimes \rho_\mu, \tag{4.9}$$

for some probability distribution $\{p_\mu\}_\mu$ and quantum states ρ_μ. Without loss of generality, we assume that all p_μ are strictly positive and we want to show that all $\rho_\mu = \rho$. Let X be any Hermitian matrix such that $\mathrm{tr}(X\rho) = 0$, then we have

$$\mathrm{tr}[(X \otimes X)(\rho \otimes \rho)] = \sum_{\mu} p_\mu \,\mathrm{tr}[(X \otimes X)(\rho_\mu \otimes \rho_\mu)] = \sum_{\mu} p_\mu [\mathrm{tr}(X\rho_\mu)]^2. \tag{4.10}$$

Combining Eq. (4.10) with the relations $\mathrm{tr}[(X \otimes X)(\rho \otimes \rho)] = [\mathrm{tr}(X\rho)]^2 = 0$ and $\mathrm{tr}(X\rho_\mu) \in \mathbb{R}$, we get that

$$\mathrm{tr}(X\rho_\mu) = 0, \tag{4.11}$$

for all μ and all X such that $\mathrm{tr}(X\rho) = 0$. This implies that

$$\rho_\mu = c_\mu \rho, \tag{4.12}$$

for some $c_\mu \in \mathbb{C}$. Furthermore, $\mathrm{tr}(\rho) = \mathrm{tr}(\rho_\mu) = 1$ implies that $c_\mu = 1$, i.e., $\rho_\mu = \rho$ for all μ. Thus, we proved that $\rho \otimes \rho$ are extreme points. □

Hence, the replacement of Eq. (4.4) by Eq. (4.5) will lead to an equivalent result as in Theorem 4.1. However, when considering relaxations of the optimization in Eq. (4.2) by replacing the separability constraint in Eq. (4.3) with some entanglement criteria, Eq. (4.4) may be strictly stronger for certain marginal problems.

Second, physically, $\mathrm{tr}(V_{AB}\Phi_{AB}) = 1$ means that Φ_{AB} is a two-party state acting on the symmetric subspace only. Hence, Theorem 4.1 is also equivalent to the feasibility problem

$$\begin{aligned} \text{find} \quad & \Phi_{AB} \in \mathrm{SEP} && (4.13)\\ \text{s.t.} \quad & V_{AB}\Phi_{AB} = \Phi_{AB},\ \mathrm{tr}(\Phi_{AB}) = 1, && (4.14)\\ & \mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \rho_I \otimes \mathrm{tr}_A(\Phi_{AB})\ \forall\ I \in \mathcal{I}. && (4.15) \end{aligned}$$

Furthermore, any feasible state Φ_{AB} can be used to construct the global state $|\varphi\rangle$ with the desired marginals, as the constraints in Theorem 4.1 imply that any pure state in the separable decomposition of Φ_{AB} yields a desired global state.

Third, the separability condition in the optimization Eq. (4.3) is usually not easy to characterize, hence relaxations of the problem need to be considered. The first candidate is the positive partial transpose (PPT) criterion [57, 58], which is an SDP relaxation of the optimization in Eq. (4.2). The PPT relaxation provides a pretty good approximation when the local dimension and the number of parties are small. In the following, inspired by the symmetric extension criterion [62], we propose a multi-party extension method as in Theorem 3.2 and obtain a complete hierarchy for the marginal problem. We denote the N parties as $A, B, \ldots, Z$, and each of them owns an n-body quantum system. For any $\mathcal{H}^{\otimes N} := \mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots \otimes \mathcal{H}_Z$, the symmetric subspace is defined as

$$\left\{ |\Psi\rangle \in \mathcal{H}^{\otimes N} \;\middle|\; V_\Sigma |\Psi\rangle = |\Psi\rangle\ \ \forall\ \Sigma \in S_N \right\}, \tag{4.16}$$

where S_N is the permutation group over N symbols and V_Σ are the corresponding operators on the N parties $A, B, \ldots, Z$; see Fig. 4.2. Let P_N^+ denote the orthogonal

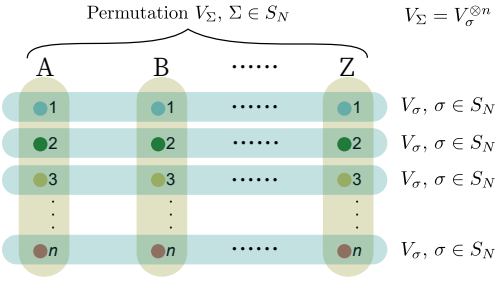


FIGURE 4.2: [134] An illustration of the complete hierarchy for the marginal problem. In order to formulate the hierarchy for the marginal problem, one extends the two copies in Fig. (4.1) to an arbitrary number of copies N. If the marginal problem has a solution $|\varphi\rangle$, then there are multi-party extensions $\Phi_{AB\cdots Z}$ in the symmetric subspace specified by $V_\Sigma = V_\sigma^{\otimes n}$ for any number of copies, obeying the semidefinite constraints in Eqs. (4.19, 4.19).

projector onto the symmetric subspace of $\mathcal{H}^{\otimes N}$. P_N^+ can be explicitly written as

$$P_N^+ = \frac{1}{N!} \sum_{\Sigma \in S_N} V_\Sigma. \tag{4.17}$$

In particular, for two parties we have the well-known relation $P_2^+ = (\mathbb{1}_{AB} + V_{AB})/2$, which implies that $\mathrm{tr}(V_{AB}\Phi_{AB}) = 1$ if, and only if, $\mathrm{tr}(P_2^+\Phi_{AB}) = 1$. Also, $V_{AB}\Phi_{AB} = \Phi_{AB}$ is equivalent to $P_2^+\Phi_{AB}P_2^+ = \Phi_{AB}$. Hereafter, without ambiguity, we will use P_N^+ to denote both the symmetric subspace and the corresponding orthogonal projector.

Then, the SDP hierarchy characterizing the marginal problem is given by the following theorem.

Theorem 4.3. *There exists a pure quantum state $|\varphi\rangle$ that satisfies $\mathrm{tr}_{I^c}(|\varphi\rangle\langle\varphi|) = \rho_I$ for all $I \in \mathcal{I}$ if and only if for all $N \geq 2$ there exists an N-party quantum state $\Phi_{AB\cdots Z}$ such that*

$$P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}, \tag{4.18}$$

$$\Phi_{AB\cdots Z} \geq 0,\ \mathrm{tr}(\Phi_{AB\cdots Z}) = 1, \tag{4.19}$$

$$\mathrm{tr}_{A_{I^c}}(\Phi_{AB\cdots Z}) = \rho_I \otimes \mathrm{tr}_A(\Phi_{AB\cdots Z})\ \forall\ I \in \mathcal{I}. \tag{4.20}$$

Each step of this hierarchy is a semidefinite feasibility problem, and the conditions become more restrictive if N increases.

This result is a direct corollary of Theorem 3.2.

Notably, we can add any criterion of full separability, e.g., the PPT criterion for all bipartitions, as extra constraints to the feasibility problem. Then, Theorem 4.3 still provides a complete hierarchy for the quantum marginal problem. In addition, the quantum marginal problems of practical interest are usually highly symmetric. These symmetries can be utilized to largely simplify the problems in Theorems 4.1 and 4.3.

Indeed, taking advantage of symmetries is usually necessary for practical applications, because the general quantum marginal problem is QMA-complete [182, 183]. Notably, even for non-overlapping marginals, despite recent progress in Refs. [184–186], it is still an open problem whether there exists a polynomial-time algorithm. In the following, we illustrate how symmetry can drastically simplify quantum marginal problems with the existence problem of AME states.

4.3 Absolutely maximally entangled states

For convenience, we recall the definition of AME states introduced in Section 2.7. An n-qudit state $|\psi\rangle$ is called an AME state, denoted as AME(n, d), if it satisfies

$$\mathrm{tr}_{I^c}(|\psi\rangle\langle\psi|) = \frac{\mathbb{1}_{d^r}}{d^r} \;\forall\; I \in \mathcal{I}_r, \tag{4.21}$$

where $\mathcal{I}_r = \{I \subset [n] \mid |I| = r\}$ and $r = \lfloor n/2 \rfloor$. Thus, Eqs. (4.13, 4.14, 4.15) imply that an AME(n, d) exists if, and only if, the following problem is feasible,

$$\text{find} \qquad \Phi_{AB} \in \mathrm{SEP} \tag{4.22}$$

$$\text{s.t.} \qquad \mathrm{tr}(\Phi_{AB}) = 1,\; V_{AB}\Phi_{AB} = \Phi_{AB}, \tag{4.23}$$

$$\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \frac{\mathbb{1}_{d^r}}{d^r} \otimes \mathrm{tr}_A(\Phi_{AB}) \;\forall\; I \in \mathcal{I}_r. \tag{4.24}$$

Direct evaluation of the problem is usually difficult, because the dimension of Φ_{AB} is $d^{2n} \times d^{2n}$, which is already very large for the simplest cases. For instance, for the 4-qubit case, the size of Φ_{AB} is 256×256.

To resolve this size issue, we investigate the symmetries that can be used to simplify the feasibility problem. Let $\mathcal{X}$ denote the set of Φ_{AB} that satisfy the constraints in Eqs. (4.22, 4.23, 4.24). If we find a unitary group G such that for all $g \in G$ and $\Phi_{AB} \in \mathcal{X}$ we have that

$$g\Phi_{AB}g^\dagger \in \mathcal{X}, \tag{4.25}$$

then the convexity of $\mathcal{X}$ implies that we can add a symmetry constraint to the constraints in Eqs. (4.22, 4.23, 4.24), namely,

$$g\Phi_{AB}g^\dagger = \Phi_{AB} \;\forall\; g \in G. \tag{4.26}$$

In the following, we will show that the symmetries of the set of AME states (if they exist for given n and d) are restrictive enough to leave only a single unique candidate for Φ_{AB}, for which separability needs to be checked. The set of AME(n, d) is invariant

under local unitaries and permutations on the n particles, so by Theorem 4.1 (or by direct verification) the following two classes of unitaries satisfy Eq. (4.25),

$$U_1 \otimes \cdots \otimes U_n \otimes U_1 \otimes \cdots \otimes U_n \ \forall\ U_i \in SU(d), \tag{4.27}$$

$$\pi \otimes \pi \ \forall\ \pi \in S_n, \tag{4.28}$$

where $\pi = \pi(A_1, A_2, \ldots, A_n) = \pi(B_1, B_2, \ldots, B_n)$ denotes the permutation operators on $\mathcal{H}_A$ and $\mathcal{H}_B$. Note that the U_i in Eqs. (4.27, 4.28) can be different.

First, let us view V_{AB} and Φ_{AB} as $V_{12\ldots n}$ and $\Phi_{12\ldots n}$, where i labels the subsystems A_iB_i. Hereafter, without ambiguity, we will omit the subscripts of

$$\mathbb{1} := \mathbb{1}_{d^2}, \quad V := V_{A_iB_i}, \tag{4.29}$$

for simplicity. From this perspective, V_{AB} can be written as $V^{\otimes n}$, and the symmetries in Eqs. (4.27, 4.28) can be written as $\bigotimes_{i=1}^n (U_i \otimes U_i)$ for $U_i \in SU(d)$ and $\Pi = \Pi(A_1B_1, A_2B_2, \ldots, A_nB_n)$ for $\Pi \in S_n$, respectively. According to Werner's result [187], a $(U \otimes U)$-invariant Hermitian operator must be of the form $\alpha\mathbb{1} + \beta V$ with $\alpha, \beta \in \mathbb{R}$. This implies that a $[\bigotimes_{i=1}^n (U_i \otimes U_i)]$-invariant state must be a linear combination of operators of the form

$$\bigotimes_{i=1}^n (\alpha_i \mathbb{1} + \beta_i V) \ \forall\ \alpha_i, \beta_i \in \mathbb{R}. \tag{4.30}$$

In addition, we take advantage of the permutation symmetry under $\Pi \in S_n$ to write any invariant Φ_{AB} as

$$\Phi_{AB} = \sum_{i=0}^n x_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\}, \tag{4.31}$$

where $\mathcal{P}$ represents the sum over all possible permutations that give different terms, e.g., $\mathcal{P}\{V \otimes \mathbb{1} \otimes \mathbb{1}\} = V \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes V \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes V$.

Before proving the existence and uniqueness of the symmetrized Φ_{AB}, we show how to simplify the constraints in Eqs. (4.23, 4.24) by taking advantage of Eq. (4.31). The meaning of this simplification is two-fold: first, it gives an intuition about why the symmetrized Φ_{AB} is uniquely determined; second, it can be directly generalized to other marginal problems, such as the m-uniform states and quantum codes, in which the symmetrized Φ_{AB} are no longer uniquely determined.

• **Normalization constraint $\mathrm{tr}(\Phi_{AB}) = 1$:**

$$\mathrm{tr}(\Phi_{AB}) = \mathrm{tr}\left[\sum_{i=0}^n x_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\}\right] = \sum_{i=0}^n \binom{n}{i} d^{2n-i} x_i = 1. \tag{4.32}$$

• **Symmetric subspace constraint $V_{AB}\Phi_{AB} = \Phi_{AB}$:**

$$V_{AB}\Phi_{AB} = V^{\otimes n}\Phi_{AB} = \sum_{i=0}^{n} x_i \mathcal{P}\{V^{\otimes(n-i)} \otimes \mathbb{1}^{\otimes i}\} = \sum_{i=0}^{n} x_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\}, \tag{4.33}$$

which implies that

$$x_i = x_{n-i} \; \forall \; i = 0, 1, \ldots, n-r-1, \tag{4.34}$$

where $r = \lfloor n/2 \rfloor$.

• **Marginal constraints $\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \frac{\mathbb{1}_{d^r}}{d^r} \otimes \mathrm{tr}_A(\Phi_{AB})$:**

Because Φ_{AB} is invariant under permutations $\Pi \in S_n$, it is sufficient to consider $I^c = \{1, 2, \ldots, n-r\}$. Further, as $\frac{\mathbb{1}_{d^r}}{d^r} \otimes \mathrm{tr}_A(\Phi_{AB}) \propto \mathbb{1}_{d^{n+r}}$, it must also hold that $\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) \propto \mathbb{1}_{d^{n+r}}$. Hence, all terms that contain V in $\mathrm{tr}_{A_{I^c}}(\Phi_{AB})$ must be zero. Thus, the marginal constraints $\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \frac{\mathbb{1}_{d^r}}{d^r} \otimes \mathrm{tr}_A(\Phi_{AB})$ are equivalent to

$$\sum_{i=0}^{n-r} \binom{n-r}{i} d^{n-r-i} x_{s+i} = 0 \; \forall \; s = 1, 2, \ldots, r. \tag{4.35}$$

Eqs. (4.32, 4.34, 4.35) provide $n+1$ linear equations, which can uniquely determine the $n+1$ parameters $(x_0, x_1, \ldots, x_n)$ in Φ_{AB}.

To rigorously prove the existence and uniqueness of Φ_{AB}, we also take advantage of the following lemma; for more details about the dual basis see, e.g., Ref. [188].

Lemma 4.4. *Let $\{|x_i\rangle\}_i$ be a basis for a finite-dimensional Hilbert space, which is not required to be orthogonal or normalized. Then, there exists a unique vector $|y\rangle$ satisfying the linear equations $\{\langle x_i|y\rangle = y_i\}_i$ for any $\{y_i\}_i$. Concretely, let $\{|\tilde{x}_i\rangle\}_i$ be the dual basis for $\{|x_i\rangle\}_i$, i.e., $\langle x_i|\tilde{x}_j\rangle = \delta_{ij}$, then $|y\rangle = \sum_i y_i |\tilde{x}_i\rangle$.*

Let $\mathcal{S}$ be the space generated by the linearly independent operators

$$X_i = \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\} \; \forall \; i = 0, 1, \ldots, n, \tag{4.36}$$

and the inner product to be the Hilbert-Schmidt inner product, e.g.,

$$\langle X_i, X_j \rangle = \mathrm{tr}(X_i^\dagger X_j) = \mathrm{tr}(X_i X_j). \tag{4.37}$$

Obviously, $\Phi_{AB} \in \mathcal{S}$ by Eq. (4.31).

By slightly modifying the derivation of Eq. (4.35), it is easy to see that the normalization constraint and the marginal constraints for AME(n, d) are equivalent to

$$\mathrm{tr}_{A_{I^c}B_{I^c}}(\Phi_{AB}) = \frac{\mathbb{1}_{d^r}}{d^r} \otimes \frac{\mathbb{1}_{d^r}}{d^r} \; \forall \; I \in \mathcal{I}_r, \tag{4.38}$$

which implies that

$$\mathrm{tr}(X_i\Phi_{AB}) = \binom{n}{i}\,\mathrm{tr}\left[V^{\otimes i}\frac{\mathbb{1}_{d^i}}{d^i}\otimes\frac{\mathbb{1}_{d^i}}{d^i}\right] = \frac{\binom{n}{i}}{d^i}\;\forall\; i=0,1,\ldots,r. \tag{4.39}$$

The symmetric subspace constraint $V_{AB}\Phi_{AB} = V^{\otimes n}\Phi_{AB} = \Phi_{AB}$ and the relation $X_iV_{AB} = X_iV^{\otimes n} = X_{n-i}$ imply that

$$\mathrm{tr}(X_i\Phi_{AB}) = \mathrm{tr}(X_iV_{AB}\Phi_{AB}) = \mathrm{tr}(X_{n-i}\Phi_{AB})\;\forall\; i=0,1,\ldots,n. \tag{4.40}$$

Thus, we get

$$\langle X_i,\Phi_{AB}\rangle = \mathrm{tr}(X_i\Phi_{AB}) = \frac{\binom{n}{i}}{\min\{d^i,d^{n-i}\}}\;\forall\; i=0,1,\ldots,n. \tag{4.41}$$

which implies the uniqueness by Lemma 4.4.

We want to find the dual basis $\{\widetilde{X}_i\}_{i=0}^n$ for $\{X_i\}_{i=0}^n$ explicitly. To do so, we first compute straightforwardly the dual basis

$$\{\frac{1}{d^2-1}(\mathbb{1}-\frac{1}{d}V),\frac{1}{d^2-1}(V-\frac{1}{d}\mathbb{1})\} \tag{4.42}$$

of $\{\mathbb{1},V\}$ using the definition. Then, for bases $\{|x_i^{(1)}\rangle\}$, $\{|x_i^{(2)}\rangle\}$ and their dual bases $\{|\tilde{x}_i^{(1)}\rangle\}$, $\{|\tilde{x}_i^{(2)}\rangle\}$, respectively, we have that $\{|\tilde{x}_i^{(1)}\rangle\otimes|\tilde{x}_j^{(2)}\rangle\}$ is the dual basis of $\{|x_i^{(1)}\rangle\otimes|x_j^{(2)}\rangle\}$ because

$$\left(\langle x_i^{(1)}|\otimes\langle x_j^{(2)}|\right)\left(|\tilde{x}_k^{(1)}\rangle\otimes|\tilde{x}_l^{(2)}\rangle\right) = \langle x_i^{(1)}|\tilde{x}_k^{(1)}\rangle\,\langle x_j^{(2)}|\tilde{x}_l^{(2)}\rangle = \delta_{ik}\delta_{jl}. \tag{4.43}$$

Hence, the dual basis respects the tensor product structure. Finally, symmetrizing both the primal and dual basis over all permutations ensures that the resulting vectors form bases of the symmetric subspace and remain dual to each other after renormalization by the number of different permutations. Thus, we obtain the dual basis of the $X_i = \mathcal{P}\{V^{\otimes i}\otimes\mathbb{1}^{\otimes(n-i)}\}$ as

$$\widetilde{X}_i = \frac{1}{\binom{n}{i}(d^2-1)^n}\mathcal{P}\left\{(\mathbb{1}-\frac{1}{d}V)^{\otimes i}\otimes(V-\frac{1}{d}\mathbb{1})^{\otimes(n-i)}\right\}\;\forall\; i=0,1,\ldots,n. \tag{4.44}$$

It is straightforward to check that $\mathrm{tr}(\widetilde{X}_i X_j) = \delta_{ij}$. Hence, we can also explicitly compute Φ_{AB} from $x_i = \mathrm{tr}(\widetilde{X}_i \Phi_{AB})$,

$$\begin{aligned} x_i =& \frac{1}{(d^2-1)^n} \mathrm{tr}\left[(\mathbb{1} - \frac{1}{d} V)^{\otimes i} \otimes (V - \frac{1}{d}\mathbb{1})^{\otimes (n-i)} \Phi_{AB} \right] \\ =& \frac{1}{(d^2-1)^n} \sum_{l=0}^{n} \sum_{k=0}^{l} \frac{(-1)^{i+l}}{d^{i+l-2k}} \binom{i}{k} \binom{n-i}{l-k} \mathrm{tr}\left[V^{\otimes (n-l)} \otimes \mathbb{1}^{\otimes l} \Phi_{AB} \right] \\ =& \frac{(-1)^i}{(d^2-1)^n} \sum_{l=0}^{n} \sum_{k=0}^{l} \frac{(-1)^l \binom{i}{k} \binom{n-i}{l-k}}{\min\{d^{i+2l-2k}, d^{n+i-2k}\}}, \end{aligned} \tag{4.45}$$

where we have used the relation

$$\mathrm{tr}\left[V^{\otimes (n-l)} \otimes \mathbb{1}^{\otimes l} \Phi_{AB} \right] = \frac{1}{\min\{d^l, d^{n-l}\}}, \tag{4.46}$$

whose proof is similar to Eq. (4.41).

Finally, we show that the symmetrized Φ_{AB} with coefficients determined by Eq. (4.45) is indeed compatible with the constraints in Eqs. (4.32, 4.34, 4.35). To this end, we show that Eq. (4.41) implies that $V_{AB}\Phi_{AB} = \Phi_{AB}$ and Eq. (4.38). As $\mathrm{tr}(X_i \Phi_{AB}) = \mathrm{tr}(X_{n-i} \Phi_{AB})$ by Eq. (4.41) and $X_i V_{AB} = X_i V^{\otimes n} = X_{n-i}$, it holds that

$$\mathrm{tr}(X_i \Phi_{AB}) = \mathrm{tr}(X_i V_{AB} \Phi_{AB}) \; \forall \; i = 0, 1, \ldots, n. \tag{4.47}$$

From the uniqueness statement in Lemma 4.4, it follows that $V_{AB}\Phi_{AB} = \Phi_{AB}$. To prove Eq. (4.38), we define $\mathcal{R}$ to be the space generated by the linearly independent operators

$$R_i = \mathcal{P}\{ V^{\otimes i} \otimes \mathbb{1}^{\otimes (r-i)} \} \; \forall \; i = 0, 1, \ldots, r. \tag{4.48}$$

Eq. (4.41) and the permutation symmetry of $\Phi_{AB} \in \mathcal{S}$ imply that

$$\mathrm{tr}\left[V^{\otimes i} \otimes \mathbb{1}^{\otimes (n-i)} \Phi_{AB} \right] = \frac{1}{d^i} \; \forall \; i = 0, 1, \ldots, r. \tag{4.49}$$

Thus,

$$\mathrm{tr}[R_i \, \mathrm{tr}_{A_{I^c} B_{I^c}}(\Phi_{AB})] = \binom{r}{i} \mathrm{tr}\left[V^{\otimes i} \otimes \mathbb{1}^{\otimes (n-i)} \Phi_{AB} \right] = \frac{\binom{r}{i}}{d^i}, \; \forall \; i = 0, 1, \ldots, r \; \forall \; I \in \mathcal{I}_r. \tag{4.50}$$

Furthermore, one can easily check that

$$\mathrm{tr}\left[R_i \frac{\mathbb{1}_{d^r}}{d^r} \otimes \frac{\mathbb{1}_{d^r}}{d^r} \right] = \frac{\binom{r}{i}}{d^i} \; \forall \; i = 0, 1, \ldots, r. \tag{4.51}$$

Then, applying the uniqueness statement in Lemma 4.4 to $\mathcal{R}$ implies Eq. (4.38). Hence, we proved the compatibility of Φ_{AB} with Eqs. (4.32, 4.34, 4.35).

This means that the two-party extension under the symmetries is independent of the specific AME state, which is an interesting structural result considering that there exist even infinite families of AME(n,d) states that are not SLOCC equivalent [189]. Together with Theorem 4.1, this result implies that an AME state exists if, and only if, Φ_{AB} is a separable quantum state.

Theorem 4.5. *An* AME(n,d) *state exists if, and only if, the state* Φ_{AB} *defined by Eqs. (4.31) and (4.45) is a separable state w.r.t. the bipartition* $(A|B) = (A_1A_2\dots A_n|B_1B_2\dots B_n)$.

To check the separability of Φ_{AB}, we first consider the positivity condition and the PPT condition. It is easy to see that Φ_{AB} can be written as

$$\Phi_{AB} = \sum_{i=0}^{n} p_i \mathcal{P}\left\{P_+^{\otimes(n-i)} \otimes P_-^{\otimes i}\right\}, \tag{4.52}$$

and $\Phi_{AB}^{T_B}$ can be written as

$$\Phi_{AB}^{T_B} = \sum_{i=0}^{n} q_i \mathcal{P}\left\{P_\phi^{\otimes(n-i)} \otimes P_\perp^{\otimes i}\right\}, \tag{4.53}$$

where

$$P_\pm = \frac{1}{2}(\mathbb{1} \pm V),\ P_\phi = |\phi^+\rangle\langle\phi^+|,\ P_\perp = \mathbb{1} - P_\phi, \tag{4.54}$$

with $|\phi^+\rangle = \frac{1}{\sqrt{d}}\sum_{k=1}^{d}|k\rangle|k\rangle$. Here p_i and q_i are the eigenvalues of Φ_{AB} and $\Phi_{AB}^{T_B}$, respectively. To get a closed form of the positivity and PPT conditions for AME states, we will use the following relations

$$\begin{aligned} \operatorname{tr}\left(V^{\otimes l}\otimes\mathbb{1}^{\otimes(n-l)}\Phi_{AB}\right) &= \frac{1}{\min\{d^l, d^{n-l}\}}, \\ \operatorname{tr}\left(|\phi^+\rangle\langle\phi^+|^{\otimes l}\otimes\mathbb{1}^{\otimes(n-l)}\Phi_{AB}^{T_B}\right) &= \frac{1}{\min\{d^{2l}, d^n\}}, \end{aligned} \tag{4.55}$$

where the proof of the first relation is similar to Eqs. (4.41, 4.49) and the second relation follows from the observation that $\operatorname{tr}(W\Phi_{AB}^{T_B}) = \operatorname{tr}(W^{T_B}\Phi_{AB})$. From Eq. (4.52) it follows that the positivity condition is equivalent to $\operatorname{tr}(P_+^{\otimes(n-i)}\otimes P_-^{\otimes i}\Phi_{AB}) \geq 0$. This gives

$$\begin{aligned} &\operatorname{tr}\left[(\mathbb{1}+V)^{\otimes(n-i)}\otimes(\mathbb{1}-V)^{\otimes i}\Phi_{AB}\right] \\ &= \operatorname{tr}\left[\sum_{l=0}^{n}\sum_{k=0}^{l}(-1)^k\binom{i}{k}\binom{n-i}{l-k}V^{\otimes l}\otimes\mathbb{1}^{\otimes(n-l)}\Phi_{AB}\right] \\ &= \sum_{l=0}^{n}\sum_{k=0}^{l}\frac{(-1)^k\binom{i}{k}\binom{n-i}{l-k}}{\min\{d^l, d^{n-l}\}} \geq 0\ \forall\ i = 0,1,\dots,n. \end{aligned} \tag{4.56}$$

Similarly due to Eq. (4.53), the PPT condition is equivalent to

$$\begin{aligned}
&\mathrm{tr}\left[|\phi^+\rangle\langle\phi^+|^{\otimes(n-i)}\otimes(\mathbb{1}-|\phi^+\rangle\langle\phi^+|)^{\otimes i}\Phi_{AB}^{T_B}\right]\\
=&\,\mathrm{tr}\left[\sum_{k=0}^{i}(-1)^k\binom{i}{k}|\phi^+\rangle\langle\phi^+|^{\otimes(n+k-i)}\otimes\mathbb{1}^{\otimes(i-k)}\Phi_{AB}^{T_B}\right]\\
=&\sum_{k=0}^{i}\frac{(-1)^k\binom{i}{k}}{\min\{d^{2(n+k-i)},d^n\}}\geq 0\ \forall\ i=0,1,\ldots,n.
\end{aligned}\tag{4.57}$$

By noticing that

$$\begin{aligned}
&\mathrm{tr}[(\mathbb{1}+V)^{\otimes(n-i)}\otimes(\mathbb{1}-V)^{\otimes i}]=d^n(d+1)^{n-i}(d-1)^i\\
&\mathrm{tr}[|\phi^+\rangle\langle\phi^+|^{\otimes(n-i)}\otimes(\mathbb{1}-|\phi^+\rangle\langle\phi^+|)^{\otimes i}]=(d^2-1)^i,
\end{aligned}\tag{4.58}$$

we also obtain an explicit expressions for p_i and q_i

$$\begin{aligned}
p_i&=\frac{1}{d^n(d+1)^{n-i}(d-1)^i}\sum_{l=0}^{n}\sum_{k=0}^{l}\frac{(-1)^k\binom{i}{k}\binom{n-i}{l-k}}{\min\{d^l,d^{n-l}\}},\\
q_i&=\frac{1}{(d^2-1)^i}\sum_{k=0}^{i}\frac{(-1)^k\binom{i}{k}}{\min\{d^{2(n+k-i)},d^n\}}.
\end{aligned}\tag{4.59}$$

For example, for the existence of the 4-qubit AME state, the eigenvalues of the matrix Φ_{AB} are

$$(p_0,p_1,p_2,p_3,p_4)=\left(\frac{5}{864},0,\frac{1}{96},0,-\frac{1}{32}\right).\tag{4.60}$$

The last negative eigenvalue implies that no AME$(4,2)$ state exists.

The positivity and PPT conditions can already rule out the existence of many AME states. Actually, they can reproduce all the known nonexistence results [17] except AME$(7,2)$ [172] and AME$(4,6)$ [181]. To get a higher-order approximation, we provide a general framework for performing the symmetric extension.

4.4 Multi-party extension: primal problem

We are going to analyze and simplify the hierarchy of SDPs stated in Theorem 4.3 for the case of the existence of AME states, i.e.,

$$\begin{aligned}
\text{find}\quad&\Phi_{AB\cdots Z}\\
\text{s.t.}\quad&P_N^+\Phi_{AB\cdots Z}P_N^+=\Phi_{AB\cdots Z},\\
&\Phi_{AB\cdots Z}\geq 0,\ \mathrm{tr}(\Phi_{AB\cdots Z})=1,\\
&\mathrm{tr}_{A_{I^c}}(\Phi_{AB\cdots Z})=\frac{\mathbb{1}_{d^r}}{d^r}\otimes\mathrm{tr}(\Phi_{B\cdots Z})\ \forall\ I\in\mathcal{I}_r.
\end{aligned}\tag{4.61}$$

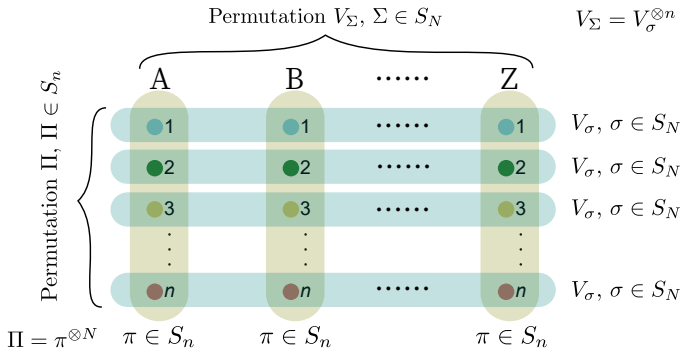


FIGURE 4.3: [134] Extended illustration of the multipartite extension. If the marginal problem has a solution $|\varphi\rangle$, then there are multi-party extensions $\Phi_{AB\cdots Z}$ for any number of copies, obeying some semidefinite constraints.

Similar to the two-party case, we can view the N-party state $\Phi_{AB\cdots Z}$ as $\Phi_{12\ldots n}$, where i labels the subsystems $A_iB_i\cdots Z_i$. The permutations on $A_iB_i\cdots Z_i$ are denoted with subscripts $ab\cdots z$. For example, V_{AB} and V_{ABC} can be written as $V_{ab}^{\otimes n}$ and $V_{abc}^{\otimes n}$, respectively, where V_{ab} are the permutations $A_i \leftrightarrow B_i$ and V_{abc} are the permutations $A_i \to B_i \to C_i \to A_i$. Generally, we use σ and Σ to denote the permutations on $ab\cdots z$ and $AB\cdots Z$, respectively, and in addition $V_\Sigma = V_\sigma^{\otimes n}$.

Again, as the set of AME(n,d) is invariant under local unitaries and permutations on the n particles, we can assume that $\Phi_{AB\cdots Z}$ is symmetric under the following operations,

$$\begin{aligned} &[U_1 \otimes \cdots \otimes U_n]^{\otimes N} \;\forall\; U_i \in SU(d), \\ &\pi^{\otimes N} \;\forall\; \pi \in S_n. \end{aligned} \tag{4.62}$$

Note that $\pi \in S_n$ denotes a permutation on $12\cdots n$ (vertical permutation in Fig. 4.3), while $\sigma \in S_N$ in the previous paragraph denotes a permutation on $ab\cdots z$ (horizontal permutation in Fig. 4.3). According to Schur-Weyl duality [190], any operator Φ such that $[\Phi, U^{\otimes N}] = 0$ must have the form

$$\Phi = \sum_\sigma x_\sigma V_\sigma. \tag{4.63}$$

Thus, the $[U_1 \otimes \cdots \otimes U_n]^{\otimes N}$ symmetry implies that

$$\Phi_{AB\cdots Z} = \sum_{\sigma_1\sigma_2\cdots\sigma_n} x_{\sigma_1\sigma_2\cdots\sigma_n} V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}. \tag{4.64}$$

The number of parameters can be further reduced by taking advantage of the vertical permutation symmetry $\{\Pi = \pi^{\otimes N} \mid \pi \in S_n\}$, i.e.,

$$x_{\sigma_1\sigma_2\cdots\sigma_n} = x_{\sigma_1'\sigma_2'\cdots\sigma_n'} \tag{4.65}$$

when $\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$ and $\{\sigma'_1, \sigma'_2, \cdots, \sigma'_n\}$ are the same multiset (set that allows repeated elements).

We are now ready to express the constraints in Eq. (4.61) in terms of the variables $x_{\sigma_1\sigma_2\cdots\sigma_n}$ in Eq. (4.64). Naively plugging Eq. (4.64) into Eq. (4.61) results in relations between large matrices; however the symmetry of the problem allows one to also simplify these constraints.

Notice that the partial trace operation can also be expressed under the basis $\{V_\sigma \mid \sigma \in S_N\}$. For example,

$$\begin{aligned} \mathrm{tr}_c(\mathbb{1}) \otimes \mathbb{1}_c = d\mathbb{1}, \quad \mathrm{tr}_c(V_{ab}) \otimes \mathbb{1}_c = dV_{ab}, \quad \mathrm{tr}_c(V_{ac}) \otimes \mathbb{1}_c = \mathbb{1}, \\ \mathrm{tr}_c(V_{bc}) \otimes \mathbb{1}_c = \mathbb{1}, \quad \mathrm{tr}_c(V_{abc}) \otimes \mathbb{1}_c = V_{ab}, \quad \mathrm{tr}_c(V_{cba}) \otimes \mathbb{1}_c = V_{ab}, \end{aligned} \tag{4.66}$$

where all V_σ are operators on abc and we perform $\otimes\mathbb{1}_c$ to ensure that the operator stays within the original space. Similarly, we can implement the trace operation. In this way, the equality constraints regarding the marginals in Eq. (4.61) can be written in terms of the basis operators $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$ without referring to explicit matrix elements. Also, the symmetric projection P_N^+ takes the form

$$P_N^+ = \frac{1}{N!} \sum_{\sigma \in S_N} V_\sigma^{\otimes n}. \tag{4.67}$$

Therefore the equality $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$ can also be expressed in terms of basis operators $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$.

Let us now consider the positivity constraint $\Phi_{AB\cdots Z} \geq 0$. Here, the crucial observation is that $\Phi_{AB\cdots Z}$ is simply a linear combination of the basis matrices $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$. The matrices V_{σ_i} in fact form a so-called (unitary linear) representation of the group S_N [190]. By the general theory of linear representations of groups, there is an orthogonal basis such that all of these matrices are block-diagonalized. Moreover, the possible blocks that appear in the block-diagonal form of these matrices are also completely specified by the group, known as the unitary irreducible representations of the group. In this way, the positivity constraint on $\Phi_{AB\cdots Z} \geq 0$ is reduced to the positivity of each of the different irreducible blocks.

For the symmetric group S_N, the irreducible representations are conveniently labeled by the partitions of N. A partition λ of length $k = |\lambda|$ is a tuple of positive integer numbers $\lambda = (N_1, N_2, \ldots, N_k)$ such that $N_1 \geq N_2 \geq \cdots \geq N_k$ and $N_1 + N_2 + \cdots + N_k = N$. We denote the set of all partitions by Λ_N. For each partition λ, there is an associated unitary irreducible representation M_λ, that is, the set of unitary matrices $M_\lambda(\sigma)$ for $\sigma \in S_N$. Concretely, by choosing a suitable orthonormal basis (independent of σ), all

V_σ can be written as

$$V_\sigma = \bigoplus_\lambda M_\lambda(\sigma) \otimes \mathbb{1}_{d_\lambda} \tag{4.68}$$

where the $M_\lambda(\sigma)$ correspond to the unitary irreducible representations and d_λ are the corresponding multiplicities. The matrix elements of $M_\lambda(\sigma)$ can also be constructed explicitly by taking advantage of the Young tableaux [191]. For practical purposes, these matrices can be called from an appropriate computer algebra system such as GAP [192]. For the representation V_σ, it is also known that $M_\lambda(\sigma)$ is present ($d_\lambda \neq 0$) in the block-diagonal form of V_σ if, and only if, the length of λ is smaller than the local dimension $|\lambda| \leq d$ [190]. We thus have the following observation.

Observation 4.6. For $\Phi_{AB\cdots Z}$ in Eq. (4.64), $\Phi_{AB\cdots Z} \geq 0$ if, and only if,

$$\sum_{\sigma_1\sigma_2\cdots\sigma_n} x_{\sigma_1\sigma_2\cdots\sigma_n} M_{\lambda_1}(\sigma_1) \otimes M_{\lambda_2}(\sigma_2) \otimes \cdots \otimes M_{\lambda_n}(\sigma_n) \geq 0, \tag{4.69}$$

for all $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \Lambda_N^n$ such that $|\lambda_i| \leq d$. In addition, as the state $\Phi_{AB\cdots Z}$ is also permutation-invariant under $\Pi \in S_n$, we can restrict to the cases where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ with any predefined order for the partitions.

There is yet another way to parameterize the optimization problem, which additionally incorporates the constraint $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$ more directly.

Recall from above that $\Phi_{AB\cdots Z}$ as well as P_N^+ are linear combinations of operators of the form $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$. Thus, by choosing a suitable basis such that V_{σ_i} are all block-diagonal, both $\Phi_{AB\cdots Z}$ and P_N^+ are also block-diagonal. The possible blocks of V_σ are labeled by partitions of the form $\lambda = (N_1, N_2, \ldots, N_k)$ with $k = |\lambda| \leq d$. Correspondingly, the possible blocks of $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$ are labeled by a tuple of partitions $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $|\lambda_i| \leq d$. Each of such blocks may appear multiple times, but because of Eq. (4.68), this simply results in exactly the same blocks in $\Phi_{AB\cdots Z}$ as well as P_N^+. Therefore, considering just one time of appearance of each block is sufficient. Moreover, because of the symmetry of coefficients in the linear combination under vertical permutations as in Eq. (4.65), only a single representative of the tuples of partitions that are different by a vertical permutation needs to be considered. Hence, we are left with analyzing the constraint $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$ within the blocks corresponding to $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

More specifically, let $\mathcal{H}_{\lambda_i}$ denote the subspace corresponding to the blocks λ_i of the operators V_{σ_i}. Then the subspace corresponding to the block $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ of $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$ is given by

$$\mathcal{H}_{\boldsymbol{\lambda}} = \mathcal{H}_{\lambda_1} \otimes \mathcal{H}_{\lambda_2} \otimes \cdots \otimes \mathcal{H}_{\lambda_n}. \tag{4.70}$$

In this subspace, the symmetric projection P_N^+ reads

$$(P_N^+)^{\boldsymbol{\lambda}} = \frac{1}{N!} \sum_{\sigma \in S_N} M_{\lambda_1}(\sigma) \otimes M_{\lambda_2}(\sigma) \cdots \otimes M_{\lambda_n}(\sigma). \tag{4.71}$$

The constraint $P_N^+ \Phi_{AB\cdots Z} P_N^+ = \Phi_{AB\cdots Z}$ restricted to the subspace $\mathcal{H}_{\boldsymbol{\lambda}}$ means that the corresponding block of $\Phi_{AB\cdots Z}$, denoted as $\Phi_{AB\cdots Z}^{\boldsymbol{\lambda}}$, is supported only on the symmetric subspace defined by the projection $(P_N^+)^{\boldsymbol{\lambda}}$,

$$\mathcal{K}_{\boldsymbol{\lambda}} = \text{Image}\left[(P_N^+)^{\boldsymbol{\lambda}}\right]. \tag{4.72}$$

Thus, if one chooses a basis $\{|\Psi_i^{\boldsymbol{\lambda}}\rangle\}_{i=1}^{k_{\boldsymbol{\lambda}}}$, where $k_{\boldsymbol{\lambda}} = \dim(\mathcal{K}_{\boldsymbol{\lambda}})$, for this subspace $\mathcal{K}_{\boldsymbol{\lambda}}$, then the corresponding block of $\Phi_{AB\cdots Z}$ is of the form

$$\Phi_{AB\cdots Z}^{\boldsymbol{\lambda}} = \sum_{i,j=1}^{k_{\boldsymbol{\lambda}}} X_{ij}^{\boldsymbol{\lambda}} \, |\Psi_i^{\boldsymbol{\lambda}}\rangle \langle \Psi_j^{\boldsymbol{\lambda}}| \, . \tag{4.73}$$

In this way, $\Phi_{AB\cdots Z}^{\boldsymbol{\lambda}}$ is parameterized by the matrix $X^{\boldsymbol{\lambda}}$, and its positivity reduces to the positivity of $X^{\boldsymbol{\lambda}}$.

In short, let us summarize the procedure to implement the optimization problem. First, enumerate all irreducible representations of S_N, i.e., all possible partitions λ. Then, select those partitions that have length $|\lambda|$ no longer than d. Based on that, enumerate all tuples of partitions $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $|\lambda_i| \leq d$. For each of those tuples $\boldsymbol{\lambda}$, compute the symmetric projection $(P_N^+)^{\boldsymbol{\lambda}}$ by Eq. (4.71) and select a basis for $\mathcal{K}_{\boldsymbol{\lambda}} = \text{Image}(P_N^+)^{\boldsymbol{\lambda}}$. Finally, for each partition tuple $\boldsymbol{\lambda}$, consider the associated positive semidefinite Hermitian matrix variable $X^{\boldsymbol{\lambda}}$ and write down the constraints corresponding to the condition on the marginals in Eq. (4.61) to complete the SDP.

In addition, we provide some more details for the construction of the basis of $\mathcal{K}_{\boldsymbol{\lambda}}$. For readers who are familiar with the representation theory of groups, there is a simple characterization of $\mathcal{K}_{\boldsymbol{\lambda}}$ that helps carrying out the practical implementation. In the language of representation theory, $\mathcal{H}_{\lambda_i}$ is an irreducible representation of S_N, while $\mathcal{H}_{\boldsymbol{\lambda}}$ is an irreducible representation of $(S_N)^n$. This space is also a representation of S_N via the diagonal embedding into $(S_N)^n$, which maps $\sigma \in S_N$ to $(\sigma, \sigma, \ldots, \sigma) \in (S_N)^n$. As a representation of S_N, $\mathcal{H}_{\boldsymbol{\lambda}}$ contains a subrepresentation $\mathcal{K}_{\boldsymbol{\lambda}}$ on which S_N acts trivially (this is technically known as the isotropic component of the trivial representation). Methods of representation theory then allow for detailed characterization of $\mathcal{K}_{\boldsymbol{\lambda}}$. In particular, one obtains the dimension of $\mathcal{K}_{\boldsymbol{\lambda}}$ as [190]

$$k_{\boldsymbol{\lambda}} = \frac{1}{N!} \sum_{\sigma \in S_N} \prod_{i=1}^{n} \text{tr}(M_{\lambda_i}(\sigma)). \tag{4.74}$$

The symmetric projection $(P_N^+)^\lambda$ in Eq. (4.71) is in fact also known as the twirling operator: it maps a vector of $\mathcal{H}_\lambda$ to its average under the action of the group S_N. A basis of this space can be found by applying the twirling operation $(P_N^+)^\lambda$ to a set of k^λ random vectors in $\mathcal{H}_\lambda$; if the resulted vectors are linearly independent, they form a basis of $\mathcal{K}_\lambda$, else one can start over with another random set of vectors. As an alternative method, Eqs. (4.71, 4.72) imply that $\mathcal{K}_\lambda$ is the common unit eigenspace of $M_{\lambda_1}(\sigma) \otimes M_{\lambda_2}(\sigma) \otimes \cdots \otimes M_{\lambda_n}(\sigma)$ for all $\sigma \in S_N$. As all eigenvalues of $M_{\lambda_i}(\sigma)$ are always in the unit circle, a basis of $\mathcal{K}_\lambda$ can also be constructed from calculating the kernel of

$$M_{\lambda_1}(\sigma_s) \otimes M_{\lambda_2}(\sigma_s) \otimes \cdots \otimes M_{\lambda_n}(\sigma_s) + M_{\lambda_1}(\sigma_c) \otimes M_{\lambda_2}(\sigma_c) \otimes \cdots \otimes M_{\lambda_n}(\sigma_c) - 2\mathbb{1}, \quad (4.75)$$

where $\sigma_s = (ab)$ and $\sigma_c = (ab \cdots z)$ form a set of generators of S_N.

As another technical remark, working with unitary representations requires computation with cyclotomic numbers, which is often slow. Therefore, one may adjust the procedure by implementing intermediate computations in non-unitary representations (or equivalently, working in nonorthogonal bases) where matrix elements (of the representations of symmetric groups) are all rationals.

4.5 Multi-party extension: dual problem and entanglement witness

Specifically for the existence problem of AME states, as Φ_{AB} is uniquely determined, one can easily verify that the following equation is a relaxed but still complete hierarchy of Theorem 4.3,

$$\begin{aligned} \text{find} \quad & \Phi_{ABC\cdots Z} \\ \text{s.t.} \quad & \mathrm{tr}_{C\cdots Z}(P_N^+ \Phi_{ABC\cdots Z} P_N^+) = \Phi_{AB}, \\ & P_N^+ \Phi_{ABC\cdots Z} P_N^+ \geq 0, \end{aligned} \quad (4.76)$$

where Φ_{AB} is the unique quantum state given by Theorem 4.5. Alternatively, we can write the objective function in Eq. (4.76) as $\max_{\Phi_{ABC\cdots Z}}\{0\}$, such that the dual problem reads

$$\begin{aligned} \min_{W_{AB}} \quad & \mathrm{tr}(W_{AB}\Phi_{AB}) \\ \text{s.t.} \quad & P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+ \geq 0, \end{aligned} \quad (4.77)$$

where W_{AB} is Hermitian. One can easily verify that strong duality holds from Slater's condition [89] with positivity considered on the symmetric subspace, which means

the problem in Eq. (4.76) is feasible if, and only if, the solution of the dual problem in Eq. (4.77) equals zero. Thus, if $\mathrm{tr}(W_{AB}\Phi_{AB}) < 0$, we know that Φ_{AB} is entangled and the corresponding AME state does not exist from Theorem 4.5. Notice that numerically determining the negativity of the dual problem in Eq. (4.77) is less sensitive to small numerical errors, and hence, more stable than solving the primal feasibility problem in Eq. (4.76). Moreover, the physical meaning of W_{AB} is also clear: a feasible point W_{AB} of Eq. (4.77) with a negative objective value provides an entanglement witness for Φ_{AB} in the symmetric subspace $P_2^+ = \frac{1}{2}(\mathbb{1}_{AB} + V_{AB})$. Indeed, because the set of separable states in P_2^+ is given by $\mathrm{conv}\{|\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|\}$, the constraint in Eq. (4.77) implies that

$$\langle\psi|\langle\psi| W_{AB} |\psi\rangle|\psi\rangle = \langle\psi|^{\otimes N} P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+ |\psi\rangle^{\otimes N} \geq 0. \tag{4.78}$$

The analysis of the symmetry and parametrization of the dual problem Eq. (4.77) is similar to that for the primal problem as discussed in Section 4.4; in fact, it is more straightforward for the dual problem. For $g \in G$ defined in Eqs. (4.100,4.101), we have

$$g\Phi_{AB}g^\dagger = \Phi_{AB}, \quad gP_N^+g^\dagger = P_N^+. \tag{4.79}$$

In addition, we know that Φ_{AB} and P_N^+ are also in the symmetric subspace P_2^+, i.e.,

$$P_2^+\Phi_{AB}P_2^+ = \Phi_{AB}, \quad \left(P_2^+ \otimes \mathbb{1}_{C\cdots Z}\right) P_N^+ \left(P_2^+ \otimes \mathbb{1}_{C\cdots Z}\right) = P_N^+. \tag{4.80}$$

Thus, we can assume that W_{AB} is invariant under G and constrained to P_2^+, i.e.,

$$gW_{AB}g^\dagger = W_{AB} \;\forall\; g \in G, \quad P_2^+W_{AB}P_2^+ = W_{AB}. \tag{4.81}$$

Similar to the analysis of Eq. (4.31), one can easily see that $gW_{AB}g^\dagger = W_{AB}$ for all $g \in G$ implying that

$$W_{AB} = \sum_{l=0}^{n} w_l \mathcal{P}\{V^{\otimes l} \otimes \mathbb{1}^{\otimes(n-l)}\}, \tag{4.82}$$

where again $\mathcal{P}$ denotes the sum over all permutations of the tensor product under its argument. Furthermore, $P_+W_{AB}P_+ = W_{AB}$ implies that

$$w_l = w_{n-l} \;\forall\; l = 0, 1, \ldots, n-r-1. \tag{4.83}$$

Hence, the objective function $\mathrm{tr}(W_{AB}\Phi_{AB})$ can be expressed as

$$\mathrm{tr}(W_{AB}\Phi_{AB}) = \sum_{l=0}^{n} a_l w_l, \tag{4.84}$$

where

$$a_l = \mathrm{tr}(\mathcal{P}\{V^{\otimes l} \otimes \mathbb{1}^{\otimes(n-l)}\}\Phi_{AB}) = \frac{\binom{n}{l}}{\min\{d^l, d^{n-l}\}}, \tag{4.85}$$

from Eq. (4.41).

The constraint $P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+ \geq 0$ can be expressed in terms of the variables w_l similar to Section 4.4. Let us summarize the arguments once more for completeness. The fact that $W_{AB} \otimes \mathbb{1}_{C\cdots Z}$ and P_N^+ are both linear combinations of $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$ implies that they are block-diagonal when one chooses a basis such that the V_{σ_i} are block-diagonal. Let $\mathcal{H}_{\lambda_i}$ denote the subspace corresponding to the block of V_{σ_i} labeled by partition λ_i with $|\lambda_i| \leq d$. Then $\mathcal{H}_{\boldsymbol{\lambda}} = \mathcal{H}_{\lambda_1} \otimes \mathcal{H}_{\lambda_2} \otimes \mathcal{H}_{\lambda_n}$ denotes the subspace corresponding to a block of $V_{\sigma_1} \otimes V_{\sigma_2} \otimes \cdots \otimes V_{\sigma_n}$ labeled by a tuple of partitions $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Moreover, within this subspace, $(P_N^+)^{\boldsymbol{\lambda}}$ is a projection onto the symmetric subspace, which is typically low-rank. Let $\mathcal{K}_{\boldsymbol{\lambda}}$ denote the image of $(P_N^+)^{\boldsymbol{\lambda}}$ and $\{|\Psi_i^{\boldsymbol{\lambda}}\rangle\}_{i=1}^{k_{\boldsymbol{\lambda}}}$ denote a basis of $\mathcal{K}_{\boldsymbol{\lambda}}$. One defines the matrix $Y^{\boldsymbol{\lambda}}$ as

$$Y_{ij}^{\boldsymbol{\lambda}} = \langle \Psi_i^{\boldsymbol{\lambda}} | \, P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+ \, | \Psi_j^{\boldsymbol{\lambda}} \rangle \, . \tag{4.86}$$

Notice that in computing these matrix elements, we only need the blocks of P_N^+ and $W_{AB} \otimes \mathbb{1}_{C\cdots Z}$ corresponding to partitions $\boldsymbol{\lambda}$. Then, $P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+ \geq 0$ is equivalent to $Y^{\boldsymbol{\lambda}} \geq 0$ for all tuples of partitions $\boldsymbol{\lambda}$ with $|\lambda_i| \leq d$. Moreover, since the problem is symmetric under vertical permutations, tuples of partitions $\boldsymbol{\lambda}$ that are different by a vertical permutation are considered just once.

As a final remark, we can consider relaxations of the constraints in Eq. (4.77). If the optimal value of a relaxed problem is nonnegative, this is also the case for the optimal value of Eq. (4.77). In particular, ignoring some tuples of partitions $\boldsymbol{\lambda}$ in the constraints $Y^{\boldsymbol{\lambda}} \geq 0$ corresponds to a relaxation of Eq. (4.77). For example, one can consider only $\boldsymbol{\lambda}$ such that $(P_N^+)^{\boldsymbol{\lambda}}$ is rank-1 and hence, obtain a linear program relaxation of Eq. (4.77).

4.6 Multi-party extension: PPT criterion with respect to any bipartition

In Sections 4.4 and 4.5, we used the representation theory of S_N to simplify the positivity constraints on the multipartite extension by block-diagonalizing the permutation matrices that appear in the decomposition of $\Phi_{ABC\cdots Z}$ and the symmetrized witness $P_N^+ W_{AB} \otimes \mathbb{1}_{C\cdots Z} P_N^+$. However, from Lemma 3.3 we can only expect a linear scaling in the entanglement detection performance. To improve this scaling, we need to add extra entanglement criteria to ensure full separability of the extension. Naturally,

$\Phi_{ABC\cdots Z}$ has a positive partial transpose with respect to any bipartition, a criterion which is known to lead to more effective detection performance [136].

From Eq. (4.64), it is clear that the partially transposed $\Phi_{ABC\cdots Z}$ can be written as a linear combination of partially transposed permutation matrices. Instead of a $[U_1 \otimes \cdots \otimes U_n]^{\otimes N}$ symmetry, $\Phi^{T_k}_{ABC\cdots Z}$, where the first k subsystems are transposed, satisfies a $[U_1^* \otimes \cdots \otimes U_k^* \otimes U_{k+1} \otimes \cdots \otimes U_n]^{\otimes N}$ symmetry. Fortunately, this implies that the partially transposed permutation matrices also form an algebra and hence, it is in principle possible to find the invariant subspaces and block-diagonalize the $V_\sigma^{T_k}$ in analogy to Eq. (4.68).

While the representation theory of the symmetric group S_N is well understood [190], there is much less known about the algebra generated by partially transposed permutation matrices beyond the partial transposition of a single subsystem [193–196]. For $N = 3$, these considerations are sufficient as there is only a single relevant partial transpose because Φ_{ABC} is symmetric and $(\Phi^{T_A}_{ABC})^T = \Phi^{T_{BC}}_{ABC}$, where the transpose preserves the spectrum and hence, also the positive semidefiniteness. In Ref. [197], the block-diagonal basis for $N = 3$ is explicitly given.

In general, only the projectors onto the invariant subspaces are needed to ensure positivity since a corresponding basis can be computed by projecting random vectors onto the different subspaces until a complete (nonorthogonal) basis is found. Such a projector S onto an invariant subspace is apparently characterized by the equations $V_\sigma^{T_k} S = S V_\sigma^{T_k} S$ or equivalently

$$(\mathbb{1} - S) V_\sigma^{T_k} S = 0, \tag{4.87}$$

for all permutations σ. We express S as a vector of $N!$ coefficients s_σ such that $S = \sum_\sigma s_\sigma V_\sigma$. Furthermore, we express $V_\sigma^{T_k}$ and $(\mathbb{1} - S)$ as matrices via their action on this vector space. This allows us to describe the Eqs. (4.87) with matrices and vectors whose size is independent of the considered local dimension d. However, the coefficients of the matrices and vectors depend, of course, on d.

We use a computer algebra system to solve the resulting system of quadratic equations for $N = 4$ and partial transposes w.r.t. the first and the first two subsystems and fixed dimension d. Then, the partial result for small dimensions d enables us to guess a general solution whose correctness can easily be checked through Eqs. (4.87). The projectors onto the invariant subspaces are given explicitly in Appendix A. To facilitate an improved multipartite extension technique by adding PPT constraints for every bipartition, we hope to extend the computations also to $N = 5$ in the near future using, e.g., Gröbner bases [198].

Finally, note that $\Phi^{T_k}_{ABC\cdots Z}$ is not in the symmetric subspace anymore, however, it still holds that

$$P_k^+ \otimes P_{N-k}^+ \Phi^{T_k}_{ABC\cdots Z} = \left(P_k^+ \otimes P_{N-k}^+\right)^{T_k} \Phi^{T_k}_{ABC\cdots Z} = \Phi^{T_k}_{ABC\cdots Z}, \tag{4.88}$$

which allows to simplify the analysis of the PPT criterion further, similar to the decomposition of $\Phi_{ABC\cdots Z}$ in Eq. (4.73).

4.7 Quantum codes

As another application, we show that our method can also be used to analyze the existence of quantum error correcting codes introduced in Section 2.7. First, we only consider pure quantum codes [87]. Our starting point is the fact that pure quantum codes are closely related to m-uniform states [86]. More precisely, an $((n, K, m+1))_d$ pure code exists if, and only if, there exists a K-dimensional subspace $\mathcal{Q}$ of $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i = (\mathbb{C}^d)^{\otimes n}$ such that all states in $\mathcal{Q}$ are m-uniform, i.e., for all $|\varphi\rangle \in \mathcal{Q}$

$$\mathrm{tr}_{I^c}(|\varphi\rangle\langle\varphi|) = \frac{\mathbb{1}_{d^m}}{d^m} \;\forall\; I \in \mathcal{I}_m, \tag{4.89}$$

where $\mathcal{I}_m = \{I \in [n] \mid |I| = m\}$ and $I^c = [n] \setminus I$. The existence of $((n, 1, m+1))_d$ pure codes reduces to the existence of m-uniform states, for which the methods from the last section are directly applicable. Here, we show that the existence of $((n, K, m+1))_d$ pure codes can still be written as a marginal problem if $K > 1$. To do so, we define an auxiliary system $\mathcal{H}_0 = \mathbb{C}^K$ and let $\widetilde{\mathcal{H}} = \mathcal{H}_0 \otimes \mathcal{H} = \bigotimes_{i=0}^n \mathcal{H}_i = \mathbb{C}^K \otimes (\mathbb{C}^d)^{\otimes n}$. Now, we can write the existence of $((n, K, m+1))_d$ pure codes as a marginal problem on $\widetilde{\mathcal{H}}$.

Lemma 4.7. *A quantum $((n, K, m+1))_d$ pure code exists if, and only if, there exists a quantum state $|Q\rangle$ in $\widetilde{\mathcal{H}}$ such that*

$$\mathrm{tr}_{I^c}(|Q\rangle\langle Q|) = \frac{\mathbb{1}_{Kd^m}}{Kd^m} \;\forall\; I \in \mathcal{I}_m, \tag{4.90}$$

where I^c is still defined as $\{1, 2, \ldots, n\} \setminus I$.

Proof. We first show the necessity part. Suppose that a $((n, K, m+1))_d$ code with corresponding subspace $\mathcal{Q}$ exists. We define an entangled state $|Q\rangle$ in $\mathcal{H}_0 \otimes \mathcal{Q} \subset \widetilde{\mathcal{H}}$ as

$$|Q\rangle = \frac{1}{\sqrt{K}} \sum_{k=1}^{K} |k\rangle\,|k_L\rangle\,, \tag{4.91}$$

where $\{|k\rangle\}_{k=1}^{K}$ and $\{|k_L\rangle\}_{k=1}^{K}$ are orthonormal bases for $\mathcal{H}_0$ and $\mathcal{Q}$, respectively. Then for any pure state $|a\rangle$ in $\mathcal{H}_0$, $\sqrt{K}\langle a|Q\rangle \in \mathcal{Q}$. Hence, Eq. (4.89) implies that

$$\mathrm{tr}_0[\mathrm{tr}_{I^c}(|a\rangle\langle a| \otimes \mathbb{1}_{d^n} |Q\rangle\langle Q|)] = \frac{\mathbb{1}_{d^m}}{Kd^m} \;\forall\; I \in \mathcal{I}_m, \tag{4.92}$$

for all $|a\rangle$ in $\mathcal{H}_0$, which in turn implies Eq. (4.90).

To prove the sufficiency part, let $\mathcal{Q}$ be the space generated by the pure states $|\varphi_a\rangle = \sqrt{K}\langle a|Q\rangle$ for all $|a\rangle$ in $\mathcal{H}_0$. Then, Eq. (4.90) implies that all $|\varphi_a\rangle$ are m-uniform states. Furthermore, from $\mathrm{rank}(\mathrm{tr}_0(|Q\rangle\langle Q|)) = \mathrm{rank}(\mathrm{tr}_{12\cdots n}(|Q\rangle\langle Q|)) = \mathrm{rank}(\mathbb{1}_K/K) = K$ it follows that $\mathcal{Q}$ is a K-dimensional subspace. □

Thus, Theorem 4.1 gives a necessary and sufficient condition for the existence of $((n, K, m+1))_d$ pure codes.

Proposition 4.8. *A quantum $((n, K, m+1))_d$ pure code exists if, and only if, there exists Φ_{AB} in $\widetilde{\mathcal{H}}_A \otimes \widetilde{\mathcal{H}}_B = [\mathbb{C}^K \otimes (\mathbb{C}^d)^{\otimes n}]^{\otimes 2}$ such that*

$$\Phi_{AB} \in \mathrm{SEP},\; V_{AB}\Phi_{AB} = \Phi_{AB},\; \mathrm{tr}(\Phi_{AB}) = 1, \tag{4.93}$$

$$\mathrm{tr}_{A_{I^c}}(\Phi_{AB}) = \frac{\mathbb{1}_{Kd^m}}{Kd^m} \otimes \mathrm{tr}_A(\Phi_{AB}) \;\forall\; I \in \mathcal{I}_m, \tag{4.94}$$

where SEP *denotes the set of separable states w.r.t. the bipartition $(A|B)$, V_{AB} is the swap operator between $\widetilde{\mathcal{H}}_A$ and $\widetilde{\mathcal{H}}_B$, and A_{I^c} denotes all subsystems A_i for $i \in I^c$.*

Furthermore, the multi-party extension and symmetrization techniques that we developed for AME states can be easily adapted to the quantum error correcting codes. For instance, the PPT relaxation can be written as a linear program and the symmetric extensions can be written as SDPs. An important difference is that the symmetrized Φ_{AB} for quantum error correcting codes is no longer uniquely determined by the marginals in general. Finally, we would like to mention that Lemma 4.7 is of independent interest on its own. For example, Eq. (4.90) implies that $Kd^m \leq \sqrt{Kd^n}$, as $\mathrm{rank}(\mathrm{tr}_{I^c}(|Q\rangle\langle Q|)) \leq \sqrt{\dim(\widetilde{\mathcal{H}})}$. This provides a simple proof for the quantum Singleton bound [85, 87] $K \leq d^{n-2m}$ for pure codes.

In general, a quantum $((n, K, m+1))_d$ code exists if, and only if, there exists a K-dimensional subspace $\mathcal{Q}$ of $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i = (\mathbb{C}^d)^{\otimes n}$ such that for all $|\varphi\rangle \in \mathcal{Q}$

$$\mathrm{tr}_{I^c}(|\varphi\rangle\langle\varphi|) = \rho_I \;\forall\; I \in \mathcal{I}_m, \tag{4.95}$$

where ρ_I are marginals that are arbitrary but independent of $|\varphi\rangle$, $\mathcal{I}_m = \{I \in [n] \mid |I| = m\}$, and $I^c = [n] \setminus I = \{1, 2, \ldots, n\} \setminus I$. Similar to the case of pure codes, we can prove the following lemma.

Lemma 4.9. *A quantum $((n, K, m+1))_d$ code exists if, and only if, there exists a quantum state $|Q\rangle$ in $\widetilde{\mathcal{H}}$ and marginal states ρ_I such that*

$$\mathrm{tr}_{I^c}(|Q\rangle\langle Q|) = \frac{\mathbb{1}_K}{K} \otimes \rho_I \;\forall\; I \in \mathcal{I}_m, \tag{4.96}$$

where $\widetilde{\mathcal{H}} = \mathcal{H}_0 \otimes \mathcal{H} = \bigotimes_{i=0}^{n} \mathcal{H}_i = \mathbb{C}^K \otimes (\mathbb{C}^d)^{\otimes n}$ and $I^c = [n] \setminus I = \{1, 2, \ldots, n\} \setminus I$.

If the marginals ρ_I are given like in the case of pure codes, the problem reduces to a marginal problem. However, to ensure the existence of $((n, K, m+1))_d$ codes, an arbitrary set of marginals is sufficient. This makes the problem no longer a marginal problem, however, we can circumvent this issue by observing that Eq. (4.96) is equivalent to

$$\mathrm{tr}_0[(M_0 \otimes \mathbb{1}_I)\,\mathrm{tr}_{I^c}(|Q\rangle\langle Q|)] = 0 \;\forall\; I \in \mathcal{I}_m, \tag{4.97}$$

for all M_0 such that $\mathrm{tr}(M_0) = 0$. Moreover, we can choose an arbitrary basis $\mathcal{B}$ for $\{M_0 \mid \mathrm{tr}(M_0) = 0,\ M_0^\dagger = M_0\}$. Then, with the general result on rank-constrained optimization in Theorem 3.2, we obtain the following theorem, and similar to the AME existence problem, a complete hierarchy can be constructed using the symmetric extension technique.

Proposition 4.10. *A quantum $((n, K, m+1))_d$ code exists if, and only if, there exists Φ_{AB} in $\widetilde{\mathcal{H}}_A \otimes \widetilde{\mathcal{H}}_B = [\mathbb{C}^K \otimes (\mathbb{C}^d)^{\otimes n}]^{\otimes 2}$ such that*

$$\Phi_{AB} \in \mathrm{SEP},\ V_{AB}\Phi_{AB} = \Phi_{AB},\ \mathrm{tr}(\Phi_{AB}) = 1, \tag{4.98}$$

$$\mathrm{tr}_{A_0}\,\mathrm{tr}_{A_{I^c}}[(M_{A_0} \otimes \mathbb{1}_{A_0^c})\Phi_{AB}] = 0, \tag{4.99}$$

for all $I \in \mathcal{I}_m$ and $M_{A_0} \in \mathcal{B}$, where the SEP *means the separability with respect to the bipartition $(A|B) = (A_0A_1\cdots A_n|B_0B_1\cdots B_n)$, V_{AB} is the swap operator between $\widetilde{\mathcal{H}}_A$ and $\widetilde{\mathcal{H}}_B$, A_{I^c} denotes all subsystems A_i for $i \in I^c$, and $\mathbb{1}_{A_0^c}$ denotes the identity operator on $AB \setminus A_0 = A_1A_2\cdots A_nB_0B_1B_2\cdots B_n$.*

By noticing that the set of $((n, K, m+1))_d$ (pure or general) codes, or rather, the set of states $|Q\rangle$, is invariant under local unitaries and permutations on the bodies $123\cdots n$, we can assume that Φ_{AB} is invariant under the following two classes of unitaries

$$U_0 \otimes U_1 \otimes \cdots \otimes U_n \otimes U_0 \otimes U_1 \otimes \cdots \otimes U_n, \tag{4.100}$$

$$\mathrm{Id}_0 \otimes \pi \otimes \mathrm{Id}_0 \otimes \pi. \tag{4.101}$$

for all $U_0 \in SU(K)$, $U_i \in SU(d)$, and $\pi \in S_n$. Thus, the symmetrized Φ_{AB} is of the form

$$\begin{aligned}\Phi_{AB} =& \mathbb{1}_{K^2} \otimes \sum_{i=0}^{n} x_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\} \\ &+ V_{A_0B_0} \otimes \sum_{i=0}^{n} y_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\},\end{aligned} \tag{4.102}$$

for $x_i, y_i \in \mathbb{R}$. The constraints in Eq. (4.99) for all $M_{A_0} \in \mathcal{B}$ are indeed equivalent for the symmetrized Φ_{AB} reducing to

$$\mathrm{tr}_{A_{I^c}} \sum_{i=0}^{n} y_i \mathcal{P}\{V^{\otimes i} \otimes \mathbb{1}^{\otimes(n-i)}\} = 0, \tag{4.103}$$

simplifying the application of Proposition 4.10. Hence, all the techniques we developed for AME states can be easily adapted to the quantum error correcting codes.

4.8 Conclusion

We have shown that the marginal problem for multiparticle quantum systems is closely related to the problem of entanglement and separability for two-party systems. More precisely, we have shown that the existence of a pure multiparticle state with given marginals can be reformulated as the existence of a two-party separable state with additional semidefinite constraints. This allows for further refinements: First, one may use the multi-party extension technique to develop a complete hierarchy for the quantum marginal problem. Second, one can use symmetries of the original marginal problem to restrict the search of the two-party separable state further. For the AME problem, this allows us to determine a unique candidate for the state, and it remains to check its separability properties. Furthermore, we provide the projectors onto the invariant subspaces of the algebra generated by the partially transposed permutation matrices of four elements explicitly. Our approach might be used to also compute these projectors for permutations of five or more elements. Finally, we extend the approach to characterize the existence of quantum codes.

Our work provides new insights in several subfields of quantum information theory. First, it may provide a significant step towards solving the general AME existence problem beyond the case of AME$(4,6)$ or quantum orthogonal Latin squares, a problem which has been highlighted as an outstanding problem in quantum information theory [180, 181]. Second, there are already a variety of results on the separability problem, and in the future, these can be used to study marginal problems in various situations. Finally, it would be interesting to extend our work to other versions of the

marginal problem, e.g., in fermionic systems or with a relaxed version of the purity constraint. We believe that our approach can also lead to progress in these cases.

5 Entanglement detection with scrambled data

Prerequisites

2.2 Quantum mechanics
2.5 Entanglement
2.8 Semidefinite programming
2.9 Classical entropy and majorization

5.1 Introduction

The main parts of this chapter have been published as Publication (A) [199]. Since entanglement is such an intriguing quantum phenomenon and a vital resource for quantum information protocols, its characterization is a central problem in many experiments. From a theoretical point of view, methods like quantum state tomography or entanglement witnesses are available. In practice, however, the situation is not so simple, as experimental procedures are always imperfect, and the imperfections are difficult to characterize. To give an example, the usual schemes for quantum tomography require the performance of measurements in a well-characterized basis such as the Pauli basis, but in practice the measurements may be misaligned in an uncontrolled manner. Thus, the question arises how to characterize states with relaxed assumptions on the measurements or on the obtained data.

For the case that the measurements are not completely characterized, several methods exist to learn properties of quantum states in a calibration-robust or even device-independent manner [200–204]. But even if the measurements are well characterized and trustworthy, there may be problems with the interpretation of the observed probabilities. For instance, in some ion trap experiments [205] the individual ions cannot be resolved, so that some of the observed probabilities cannot be uniquely assigned to the measurement operators in a quantum mechanical description. More generally, we consider the situation where the connection between the outcomes of a measurement and the observed probabilties is lost, in the sense that the probabilities are permuted

in an uncontrolled way. We call this situation the "scrambled data" scenario. Because we still assume that the measurements have a well-characterized quantum mechanical description, the examined situation is complementary to the calibration-robust or device-independent scenario.

One simple example of scrambling is the permutation of particles. Permutation-invariant states have been considered extensively in the literature [133, 206–209]. Scrambling, however, allows for more manipulations than just permuting particles. Thus, entanglement detection methods relying just on the scrambled data require only few assumptions.

In this chapter, we present a detailed study of different methods of entanglement detection using scrambled data. After explaining the setup and the main definitions, our focus lies on the two-qubit case and Pauli measurements. We first study the use of entropies for entanglement detection. Entropies are natural candidates for this task, as they are invariant under permutations of the probabilities. We demonstrate that Tsallis- and Rényi entropies can detect entanglement in our scenario, while the Shannon entropy is sometimes useless. For deriving our criteria, we prove some entropic uncertainty relations, which may be of independent interest. Second, we introduce scrambling invariant entanglement witnesses. The key observation is here that for certain witnesses the permutation of the data corresponds to the evaluation of another witness, so that the scrambling of the data does not matter. Third, we characterize the states for which the scrambled data may origin from a separable state, meaning that their entanglement cannot be detected in the scrambled data scenario. We show that this set of states is generally not convex, which gives an intuition why entanglement detection with scrambled data is a hard problem in general.

5.2 Setup and Definitions

Consider an experiment with two qubits and local projective dichotomic measurements $A \otimes B$, i.e., local measurements with two outcomes each described by observables with eigenvalues ± 1. Then, the data consists of four outcome probabilities $p(A = \pm 1, B = \pm 1)$. We define the *scrambled data* as a random permutation of these probabilities within but not in-between measurements, such that the assignment of probabilities to outcomes is erased. The restriction that permutations in-between measurements are excluded is natural since they are generically inconsistent because the probabilities are not normalized anymore.

	$\sigma_x \otimes \sigma_x$				$\sigma_z \otimes \sigma_z$			
	p_{++}	p_{+-}	p_{-+}	p_{--}	p_{00}	p_{01}	p_{10}	p_{11}
$\|\psi^-\rangle$	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0
$\|+\rangle \otimes \|0\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0

TABLE 5.1: This table shows the measurement data for the singlet state $|\psi^-\rangle = (|+-\rangle - |-+\rangle)/\sqrt{2}$ and the product state $|+\rangle \otimes |0\rangle$ and local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$. The scrambled data is the same for the two states. Thus, detecting the entanglement of the singlet state with these measurements is impossible using only the scrambled data.

As an example for scrambled data, we consider the singlet state $|\psi^-\rangle = (|+-\rangle - |-+\rangle)/\sqrt{2}$ and the product state $|0\rangle \otimes |+\rangle$. In order to detect the entanglement of $|\psi^-\rangle$, it is usually sensible to perform the local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, as there exists an entanglement witness $W = \mathbb{1} + \sigma_x \otimes \sigma_x + \sigma_z \otimes \sigma_z$ detecting this state [52]. These measurements yield the outcome probabilities p_{++}, p_{+-}, p_{-+}, and p_{--} and p_{00}, p_{01}, p_{10}, and p_{11}, respectively. From Table 5.1, we clearly see that the measurement data is different for the two states. However, it is easy to see that there is no way of distinguishing the two states using these measurements if one has access only to the *scrambled* data since the probability distributions are mere permutations of each other. Thus, it is impossible to detect the entanglement of the singlet state because there is a separable state realizing the same scrambled data. We call states whose scrambled data can be realized by a separable state *possibly separable* as the entanglement cannot be detected in this scenario.

The above observation motivates to focus specifically on the local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ in the following analysis. However, all results hold more generally for local measurements $A_1 \otimes B_1$ and $A_2 \otimes B_2$ if the eigenstates of both A_1, A_2 and B_1, B_2 form mutually unbiased bases, i.e., $|\langle a_1^{(i)}|a_2^{(j)}\rangle| = |\langle b_1^{(i)}|b_2^{(j)}\rangle| = 1/\sqrt{d}$. This is clear from the Bloch sphere representation because any orthogonal basis can be rotated to match the analysis in this work. Indeed, in dimension two and three, all pairs of mutually unbiased bases are equivalent under local unitaries [210], including the locally two-dimensional case considered here.

In total, there are $(4!)^2 = 576$ different scrambling permutations of the eight probabilities. 96 of the permutations can be realized by quantum channels, i.e., completely positive trace-preserving maps. Namely, those transformations can be implemented using the local channels $\mathcal{M}(\rho) = (\sigma_x \otimes \sigma_x)\rho(\sigma_x \otimes \sigma_x)$ and $\mathcal{N}(\rho) = (\sigma_z \otimes \sigma_z)\rho(\sigma_z \otimes \sigma_z)$ plus the two-qubit SWAP-gate and the controlled-NOT gate, as well as compositions

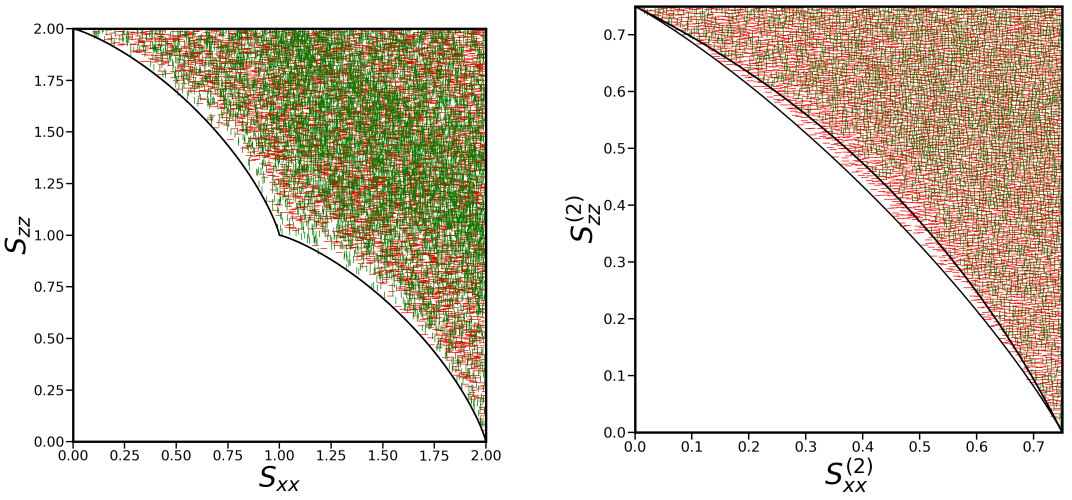


FIGURE 5.1: [199] These plots show entropy samples of local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ for Shannon entropy (left) and Tsallis-2 entropy (right) where separable and entangled states are represented by green vertical and red horizontal lines, respectively. The plot indicates that Shannon entropy is useless for entanglement detection, while Tsallis-2 entropy is suitable.

of these channels. In particular, this means that a violation of our entanglement criteria certifies that the experimental setup allows for entanglement generation beyond the controlled-NOT gate.

5.3 Entropic uncertainty relations

Entropies provide a natural framework to examine scrambled data because they are invariant under permutation of probabilities and hence, robust against scrambling. In this section, we show that measuring Tsallis-q or Rényi-α entropies for the two local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ in many cases allows for the detection of entanglement and show a new family of non-linear entropic uncertainty relations.

For local measurements $\sigma_i \otimes \sigma_i$, S_{ii} and $S_{ii}^{(q)}$ where $i \in \{x, y, z\}$ shall denote the Shannon and Tsallis-q entropy of the corresponding four probabilities, respectively. In order to detect entanglement, we investigate the possible pairs of $S_{xx}^{(\bar{q})}$ and $S_{zz}^{(q)}$ that can be realized by physical states. For gaining some intuition, we have plotted in Fig. 5.1 random samples of separable and entangled two-qubit states, where separability can be checked using the PPT criterion [57, 58]. As the figures indicate, the accessible region for both kinds of states does not differ in the case of Shannon entropy and hence, entanglement detection seems impossible in this case. This is supported by findings in earlier works: It has been shown in Ref. [211] that in the case of Shannon entropy and two local measurements, linear entropic uncertainty relations of the type

$\alpha S_{xx} + \beta S_{zz} \geq c_{\text{sep}} \geq c$ with bounds c_{sep} for separable and c for all states, are infeasible to detect entanglement, i.e., $c_{\text{sep}} = c$. Furthermore, Conjecture V.6 in Ref. [212] states that in the example of local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, even non-linear entropic uncertainty relations cannot be used to detect entanglement. However, non-linear relations are unknown in most cases [212].

In contrast to the case of Shannon entropy, using Tsallis-2 entropy, we identify a distinct region occupied by entangled states only. In the following, we will show that also $(S_{xx}^{(\tilde{q})}, S_{zz}^{(q)})$-plots with $q, \tilde{q} \geqslant 2$ exhibit this feature by determining the lower bounds of the set of all and the set of separable states explicitly. First, note that a vanishing entropy of $S_{ii}^{q} = 0$ implies that the system is in an eigenstate of the measurement operator $\sigma_i \otimes \sigma_i$. Since the measurements define mutually unbiased bases, it is clear that in this case the other entropy is maximal. Hence, the states $|00\rangle$ and $|{+}{+}\rangle$ lie on the boundary of the realizable region. The mixture of these states with white noise $\mathbb{1}/4$ leaves the maximal entropy of one measurement unchanged while increasing the entropy of the other measurement continuously. Therefore, the upper and the right boundary of the region, corresponding to maximal S_{zz} and S_{xx}, respectively, is reached by separable states (see Fig. 5.1). We will see later that the lower boundaries for all and for separable states are both realized by continuous one-parameter families of states. Thus, the mixture of these states with white noise forms a continuous family of curves connecting the lower boundary with the point where both entropies are maximal. Hence, these states realize any accessible point in the entropy plot and it is sufficient to only determine the lower boundary.

5.3.1 Entropic bound for general states

We begin by determining the bounds in the $(S_{xx}^{(\tilde{q})}, S_{zz}^{(q)})$-plot for all states. In Ref. [212], Theorem V.2 states that for two concave functionals f_1, f_2 on the state space, for any state ρ, there is a pure state $|\psi\rangle$ such that $f_1(|\psi\rangle\langle\psi|) \leqslant f_1(\rho)$ and $f_2(|\psi\rangle\langle\psi|) \leqslant f_2(\rho)$. Furthermore, it is shown in Theorem V.3 that the state $|\psi\rangle$ can additionally be chosen real if the inputs of the functionals are linked by a real unitary matrix. Thus, in the case of general two-qubit states and local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, the analysis of the boundary of the entropy plots can be reduced to pure real states. First, we will solve the special case where $q = 2$, also known as linear entropy. This result can then be used as an anchor to prove the bound for all $q \geqslant 2$.

Lemma 5.1. *For two-qubit states ρ and fixed $S_{zz}^{(2)}(\rho)$, minimal $S_{xx}^{(2)}(\rho)$ is reached by the unique state $\rho_t = |\psi_t\rangle\langle\psi_t|$ where $|\psi_t\rangle = \frac{1}{\sqrt{3+t^2}}(t\,|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ and some $t \geqslant 1$ determined by the given entropy $S_{zz}^{(2)}(\rho)$.*

Proof. According to Theorem V.3 in Ref. [212], if two entropies S_1 and S_2 are considered where the measurement bases are related by a real unitary transformation, then for any state ρ, there is always a pure and real state $|\psi\rangle$ with $S_1(|\psi\rangle\langle\psi|) \leqslant S_1(\rho)$ and $S_2(|\psi\rangle\langle\psi|) \leqslant S_2(\rho)$. As in our case $\sigma_x = H\sigma_z H^\dagger$ where H is the Hadamard matrix, it is sufficient to consider pure real states to obtain minimal $S_{zz}^{(2)}$ for given $S_{xx}^{(2)}$. For a general pure real state $|\psi\rangle = (x_1, x_2, x_3, x_4)^T$, the problem boils down to the following maximization problem under constraints

$$\begin{aligned} &\max_{x_i} f(x_1, x_2, x_3, x_4), \\ &\text{s.t. } x_1^4 + x_2^4 + x_3^4 + x_4^4 = k = \text{const.}, \\ &\quad x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 \end{aligned} \tag{5.1}$$

where $f(\{x_i\}) = (x_1 + x_2 + x_3 + x_4)^4 + (x_1 + x_2 - x_3 - x_4)^4 + (x_1 - x_2 + x_3 - x_4)^4 + (x_1 - x_2 - x_3 + x_4)^4 = 1 - S_{xx}^{(2)}$ and $k = 1 - S_{zz}^{(2)}$. Note that $\frac{1}{4} \leqslant k \leqslant 1$. It is straightforward to see that $\frac{1}{96}[f(\{x_i\}) - 12 \times 1^2 + 8k] = \frac{1}{96}[f(\{x_i\}) - 12(x_1^2 + x_2^2 + x_3^2 + x_4^2)^2 + 8(x_1^4 + x_2^4 + x_3^4 + x_4^4)] = x_1x_2x_3x_4$, using the constraints. So, we can replace f by $x_1x_2x_3x_4$. Clearly, the x_j can be chosen greater than 0 in case of a maximum. Consequently, x_i can be substituted by $\sqrt{y_i}$ because the square root is a monotone function. Thus, the problem reduces to

$$\begin{aligned} &\max_{y_i} y_1y_2y_3y_4, \\ &\text{s.t. } y_1^2 + y_2^2 + y_3^2 + y_4^2 = k = \text{const.}, \\ &\quad y_1 + y_2 + y_3 + y_4 = 1 \end{aligned} \tag{5.2}$$

where all y_i are positive. Using Lagrange multipliers, it is straightforward to obtain the optimal solution. For given $S_{zz}^{(2)}$, the minimal $S_{xx}^{(2)}$ is reached by the state

$$|\psi_t\rangle = \frac{1}{\sqrt{3 + t^2}}(t\,|00\rangle + |01\rangle + |10\rangle + |11\rangle), \tag{5.3}$$

for some $t \geqslant 1$. Since the minimal $S_{zz}^{(2)}$-entropy state $|\psi_\infty\rangle = |00\rangle$ and maximal $S_{zz}^{(2)}$-entropy state $|\psi_0\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ are part of the family $|\psi_t\rangle$ and $\frac{dS_{zz}^{(2)}(|\psi_t\rangle)}{dt} < 0$, fixing $S_{zz}^{(2)}$ uniquely determines t and hence, also $\rho_t = |\psi_t\rangle\langle\psi_t|$. □

This result holds for the Tsallis-2 entropy. However, it can be generalized to any pair of Tsallis-q and Tsallis-$\tilde{q}$ entropies with $q, \tilde{q} \geqslant 2$. To that end, we use a result from Ref. [213]. There, the authors consider entropy measures $H_f = \sum_i f(p_i)$ and $H_g = \sum_i g(p_i)$ where $f(0) = g(0) = 0$ and the functions f, g are strictly convex (implying that $g'(p)$ is invertible) with their first derivatives being continuous in the

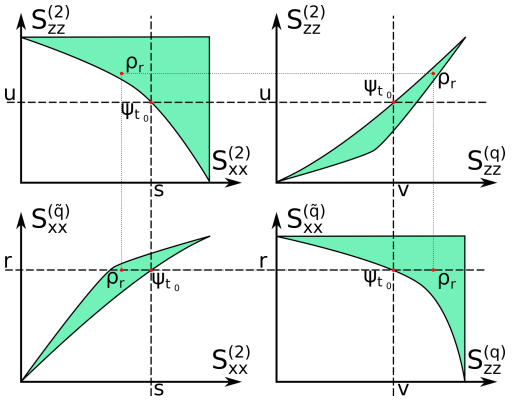


FIGURE 5.2: [199] These sketched plots depict the proof of Theorem 5.2. Starting with the lower right picture, for fixed ρ_r with $S_{xx}^{(\tilde{q})}(\rho_r) = r$, we consider the state $|\psi_{t_0}\rangle$ defined in Lemma 5.1, with t_0 such that also $S_{xx}^{(\tilde{q})}(\psi_{t_0}) = r$. The state $|\psi_{t_0}\rangle$ has the largest $S_{xx}^{(2)}$-entropy among all states ρ with $S_{xx}^{(\tilde{q})}(\rho) = r$ [213], particularly including ρ_r (see lower left). From Lemma 5.1, it follows that $S_{zz}^{(2)}(\psi_{t_0}) \leqslant S_{zz}^{(2)}(\rho_r)$ which is shown in the upper left. This, in turn, implies that $S_{zz}^{(q)}(\psi_{t_0}) \leqslant S_{zz}^{(q)}(\rho_r)$ [213] (see plot on the upper right). In summary, we have that for any state ρ_r, there exists a state $|\psi_{t_0}\rangle$ with $S_{xx}^{(\tilde{q})}(\psi_{t_0}) = S_{xx}^{(\tilde{q})}(\rho)$ and $S_{zz}^{(q)}(\psi_{t_0}) \leqslant S_{zz}^{(q)}(\rho)$. This proves that the boundary is realized by the states $|\psi_t\rangle$, which is illustrated again in the lower right.

interval $(0,1)$. They show that then the maximum (minimum) of H_f for fixed H_g is obtained by the probability distribution $p_1 \geqslant p_2 = \cdots = p_n$ if $f'[p(g')]$ as a function of g' is strictly convex (concave). Furthermore, for each value of H_g, there is a unique probability distribution of this form.

In the specific case of Tsallis entropies with parameters q and $\tilde{q}$, it is shown that if

$$\frac{q(q-1)}{\tilde{q}(\tilde{q}-1)} p^{q-\tilde{q}} \tag{5.4}$$

is monotonically increasing (decreasing), the minimum (maximum) S^q for fixed $S^{\tilde{q}}$ is reached by the probability distribution described above, when considering the same measurement for different $q, \tilde{q}$. That is exactly the probability distribution obtained by measuring $\sigma_x \otimes \sigma_x$ or $\sigma_z \otimes \sigma_z$ locally in the state $|\psi_t\rangle$ since

$$\begin{aligned} |\psi_t\rangle &\propto t\,|00\rangle + |01\rangle + |10\rangle + |11\rangle \\ &\propto (t+3)\,|{+}{+}\rangle + (t-1)(|{+}{-}\rangle + |{-}{+}\rangle + |{-}{-}\rangle), \end{aligned} \tag{5.5}$$

where $t^2 \geqslant 1$ and $(t+3)^2 \geqslant (t-1)^2$. This observation assists in proving the following theorem.

Theorem 5.2. *For all $q, \tilde{q} \geqslant 2$, the lower boundary in the $(S_{xx}^{(\tilde{q})}, S_{zz}^{(q)})$-plot is realized by the family of states $|\psi_t\rangle = \frac{1}{\sqrt{3+t^2}}(t\,|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ where $t \geqslant 1$.*

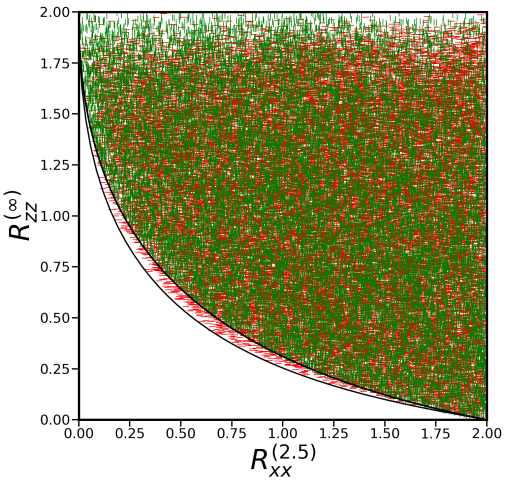


FIGURE 5.3: [199] This plots shows entropy samples of local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ for Rényi-2.5- and Rényi-∞-entropies, respectively. Separable states are represented by green vertical lines, while red horizontal lines indicate entangled states. The lower boundary is given by the states $|\psi_t\rangle$ defined in Lemma 5.1.

Proof. For fixed r and a state ρ_r with $S_{xx}^{(\tilde{q})}(\rho_r) = r$, there exists a unique state $|\psi_{t_0}\rangle$ with $S_{xx}^{(\tilde{q})}(\psi_{t_0}) = r$. From Theorem 1 in Ref. [213], it follows that

$$S_{xx}^{(2)}(\rho_r) \leqslant S_{xx}^{(2)}(\psi_{t_0}) \equiv s \tag{5.6}$$

(see bottom left graph in Fig. 5.2). Since $\frac{dS_{zz}^{(2)}(|\psi_t\rangle)}{dS_{xx}^{(2)}(|\psi_t\rangle)} = \frac{dS_{zz}^{(2)}(|\psi_t\rangle)}{dt}\left(\frac{dS_{xx}^{(2)}(|\psi_t\rangle)}{dt}\right)^{-1} < 0$, it follows from Lemma 5.1 that $S_{xx}^{(2)}(\rho_r) \leqslant s$ implies

$$S_{zz}^{(2)}(\rho_r) \geqslant S_{zz}^{(2)}(\psi_{t_0}) \equiv u \tag{5.7}$$

(see top left graph in Fig. 5.2). Now, given $S_{zz}^{(2)}(\rho_r) \geqslant u$, using the fact that $\frac{dS_{zz}^{(q)}(|\psi_t\rangle)}{dS_{zz}^{(2)}(|\psi_t\rangle)} > 0$, it follows from Ref. [213] that

$$S_{zz}^{(q)}(\rho_r) \geqslant S_{zz}^{(q)}(\psi_{t_0}) \equiv v \tag{5.8}$$

(see top right graph in Fig. 5.2).

In summary, by considering all values of r, we find that for all two-qubit states ρ_r,

$$S_{xx}^{(\tilde{q})}(\rho_r) = r \Rightarrow S_{xx}^{(2)}(\rho_r) \leqslant S_{xx}^{(2)}(\psi_{t_0}) \tag{5.9}$$

$$\Rightarrow S_{zz}^{(2)}(\rho_r) \geqslant S_{zz}^{(2)}(\psi_{t_0}) \tag{5.10}$$

$$\Rightarrow S_{zz}^{(q)}(\rho_r) \leqslant S_{zz}^{(q)}(\psi_{t_0}) \tag{5.11}$$

(see also the lower right graph in Fig. 5.2), where $|\psi_{t_0}\rangle$ is uniquely determined by $S_{xx}^{(\tilde{q})}(\psi_{t_0}) = r$. All bounds, as well as the overall implication $S_{xx}^{(\tilde{q})}(\rho_r) = r \Rightarrow S_{zz}^{(q)}(\rho_r) \leqslant$

$S_{zz}^{(q)}(\psi_{t_0})$ are tight since they are saturated by the same state $|\psi_{t_0}\rangle$. Thus, the lower boundary in the $(S_{xx}^{(\tilde{q})}, S_{zz}^{(q)})$-plot is realized by the family of states $|\psi_t\rangle = \frac{1}{\sqrt{3+t^2}}(t\,|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ where $t \geqslant 1$. □

In the above proof, we used the $q = \tilde{q} = 2$-case as an anchor to derive the result for all $q, \tilde{q} \geqslant 2$. The same argument also holds if we would use any other anchor case where the $|\psi_t\rangle$ are the optimal states. Numerical evidence suggests that the conclusion is indeed valid for any $q, \tilde{q} \gtrsim 1.37$. Furthermore, the result can also be interpreted as a family of entropic uncertainty relations.

Corollary 5.3. *For all two-qubit states ρ and $q, \tilde{q} \geqslant 2$,*

$$\begin{aligned} F[S_{xx}^{(\tilde{q})}(\rho), S_{zz}^{(q)}(\rho)] &\equiv S_{zz}^{(q)}(\rho) - S_{zz}^{(q)}(|\psi_t\rangle\,[S_{xx}^{(\tilde{q})}(\rho)]) \\ &= S_{zz}^{(q)}(\rho) - \frac{1}{q-1}\left(1 - \frac{3 + t^{2q}[S_{xx}^{(\tilde{q})}(\rho)]}{\{3 + t^2[S_{xx}^{(\tilde{q})}(\rho)]\}^q}\right) \geqslant 0. \end{aligned} \tag{5.12}$$

Here, $|\psi_t\rangle\,[S_{xx}^{(\tilde{q})}(\rho)]$ and $t[S_{xx}^{(\tilde{q})}(\rho)]$ are the unique state $|\psi_t\rangle$ and parameter t in dependence on $S_{xx}^{(\tilde{q})}(\rho)$, such that $S_{xx}^{(\tilde{q})}(|\psi_t\rangle) = S_{xx}^{(\tilde{q})}(\rho)$.

In the case of $q = \tilde{q} = 2$, we have

$$F[S_{xx}^{(2)}(\rho), S_{zz}^{(2)}(\rho)] = S_{zz}^{(2)} - \frac{3QT^2 - T^4}{3Q^2} \geqslant 0, \tag{5.13}$$

where

$$T = \sqrt{9 - 12S_{xx}^{(2)}}, \tag{5.14}$$

$$Q = 3 + T + \sqrt{3}\sqrt{(1+T)(3-T)}. \tag{5.15}$$

This bound is displayed in Fig. 5.1.

Note that this result is also valid for Rényi-α entropies [98] with $\alpha, \tilde{\alpha} \geqslant 2$ as Rényi-α and Tsallis-q entropies are monotone functions of each other for $\alpha = q$. Thus, the change from Tsallis- to Rényi entropies merely induces a rescaling of the axes in the (S_{xx}, S_{zz})-plot. An example is given in Fig. 5.3 where $\alpha = 2.5$ and $\tilde{\alpha} = \infty$.

In contrast to any linear bounds which are usually considered [211] the uncertainty relations found here are optimal. That means, for any entropic uncertainty relation defined in Corollary 5.3 and any $S_{zz}^{(q)}$, there exists a state, namely the $|\psi_t\rangle$ with the given entropy, saturating the corresponding bound.

5.3.2 Entropic bound for mutually unbiased bases

Before proceeding with the entropic bound for separable states measuring $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ locally, we want to discuss the general problem of an entropic bound for measurements in the computational and Fourier basis, given by $|j\rangle$ and

$$|\tilde{j}\rangle = \frac{1}{\sqrt{d}} \sum_k e^{i\omega jk} |k\rangle \,, \tag{5.16}$$

respectively, where $\omega = 2\pi/d$. In particular, we aim at generalizing Lemma 5.1, i.e., minimizing the Tsallis-2 entropy for one measurement for fixed Tsallis-2 entropy of the other measurement.

Again, it is sufficient to consider pure states [212]. The coefficients of a general pure state $|\psi\rangle = \sum_j z_j |j\rangle$, where $z_j = r_j e^{-i\phi_j}$ with $r_j \geq 0$ and $0 \leq \phi_j \leq 2\pi$, in the Fourier basis are given by $\tilde{z}_j = \sum_k z_k e^{-i\omega jk}/\sqrt{d}$. Then, we have for the entropy measuring in the Fourier basis,

$$\begin{aligned}
1 - S_F^{(2)} &= \sum_j |\tilde{z}_j|^4 = \frac{1}{d^2} \sum_j \left| \sum_k z_k e^{-i\omega jk} \right|^4 \\
&= \frac{1}{d^2} \sum_j \left\{ \left[\sum_k r_k \cos(\omega jk + \phi_k) \right]^2 + \left[\sum_k r_k \sin(\omega jk + \phi_k) \right]^2 \right\}^2 \\
&= \frac{1}{d^2} \sum_j \left[\sum_{k,l} r_k r_l \cos(\omega jk + \phi_k) \cos(\omega jl + \phi_l) + \sin(\omega jk + \phi_k) \sin(\omega jl + \phi_l) \right]^2 \\
&= \frac{1}{d^2} \sum_{j,k,l,m,n} r_k r_l r_m r_n \cos(\omega j(k-l) + \phi_k - \phi_l) \cos(\omega j(m-n) + \phi_m - \phi_n) \\
&= \frac{1}{2d^2} \sum_{k,l,m,n} r_k r_l r_m r_n \left[\sum_j \cos(\omega j(k-l+m-n) + \phi_k - \phi_l + \phi_m - \phi_n) \right. \\
&\qquad \left. + \sum_j \cos(\omega j(k-l-(m-n)) + \phi_k - \phi_l - (\phi_m - \phi_n)) \right] \\
&= \frac{1}{2d} \sum_{k,l,m,n} r_k r_l r_m r_n \left[\delta_{k-l+m-n \mod d,0} \cos(\phi_k - \phi_l + \phi_m - \phi_n) \right. \\
&\qquad \left. + \delta_{k-l-(m-n) \mod d,0} \cos(\phi_k - \phi_l - (\phi_m - \phi_n)) \right] ,
\end{aligned} \tag{5.17}$$

where we used the well-known fact that $\sum_j \cos(q\omega + \phi) = d\delta_{q \mod d,0}$ for whole numbers q, where $\delta_{q \mod d,0}$ is equal to 1 if q is a multiple of d and 0 otherwise.

Since all $r_j \geqslant 0$, the minimal entropy with respect to the angles ϕ_j is achieved if

$$\phi_k - \phi_l \pm (\phi_m - \phi_n) \equiv 0 \mod 2\pi, \tag{5.18}$$

for all k, l, m, n for which $k - l \pm (m - n) \equiv 0 \mod d$ holds. Choosing a global phase such that $\phi_0 = 0$, which leaves the state $|\psi\rangle$ invariant, this condition can be simplified to

$$\phi_j \equiv j\phi_1 \mod d \tag{5.19}$$

where $j = 0, \ldots, d$. Hence, ϕ_1 determines all other phases and has to satisfy $d\phi_1 \equiv 0 \mod 2\pi$.

Because of the symmetry between the computational and the dual basis, we can also fix the entropy $S_F^{(2)}$ while minimizing $S_C^{(2)}$, the entropy of a measurement in the computational basis. Again, we obtain the analogous condition $\tilde{\phi}_j \equiv j\tilde{\phi}_1 \mod d$. If a state $|\psi\rangle$ satisfies only one of the conditions, we can decrease one of the entropies while keeping the other constant by adjusting the phases ϕ_j or $\tilde{\phi}_j$. Hence, the optimal states satisfy both $\phi_j \equiv j\phi_1 \mod d$ and $\tilde{\phi}_j \equiv j\tilde{\phi}_1 \mod d$. Now, the choice of different feasible ϕ_1 corresponds merely to a cyclic permutation of the $\tilde{z}_j$, and vice versa, because

$$\tilde{z}_j = \frac{1}{\sqrt{d}} \sum_k r_k e^{-i\omega k(j+\xi)}, \tag{5.20}$$

where $\phi_1 = \xi\omega$ and $0 \leq \xi < d$ is a natural number. Thus, we can choose $\phi_1 = \tilde{\phi}_1 = 0$ making the coefficients in both bases positive real numbers. Then, from $\tilde{x}_j = \tilde{x}_j^*$, it follows that

$$0 = \sum_{k=0}^{d-1} x_k \sin(\omega jk) = \frac{1}{2} \sum_{k=1}^{\lfloor (d-1)/2 \rfloor} (x_k - x_{d-k}) \sin(\omega jk), \tag{5.21}$$

which are independent for $j = 1, \ldots, \lfloor (d-1)/2 \rfloor$. Using an inverse-Fourier-like transform for $l = 1, \ldots, \lfloor (d-1)/2 \rfloor$,

$$\begin{aligned} 0 &= \sum_{j=0}^{d-1} \sum_{k=1}^{\lfloor (d-1)/2 \rfloor} (x_k - x_{d-k}) \sin(\omega jk) \sin(\omega jl) \\ &\propto \sum_{k=1}^{\lfloor (d-1)/2 \rfloor} (x_k - x_{d-k}) \sum_{j=0}^{d-1} \left[\cos\left(\omega j(k-l)\right) - \cos\left(\omega j(k+l)\right) \right] \propto (x_l - x_{d-l}), \end{aligned} \tag{5.22}$$

because $2 \leq k + l \leq d - 1$ and hence, the sum over the second cosine term vanishes, and we again use $\sum_j \cos\left(\omega j(k-l)\right) = d\delta_{(k-l)0}$. Thus, we obtain the symmetry $x_j = x_{d-j}$.

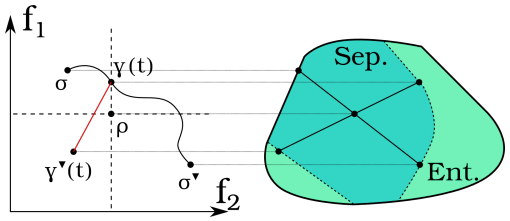


FIGURE 5.4: [199] This sketch shows the proof idea of Theorem 5.4, where the left plot is based on Fig. 6 in Ref. [212]. Any separable state ρ can be written as the mixture of two states on the topological boundary of the space of separable states. These two states can be converted into each other continuously. In this process, we find a state $\gamma^{\blacktriangledown}(t)$ on the boundary such that $f_1[\gamma^{\blacktriangledown}(t)] \leqslant f_1(\rho)$ and $f_2[\gamma^{\blacktriangledown}(t)] \leqslant f_2(\rho)$ for continuous concave functionals f_1 and f_2.

In summary, the optimization problem that needs to be solved for any $d \geqslant 3$ is given by

$$\max \sum_{k,l,m=0}^{d-1} x_k x_l x_m x_{k-l+m} \mod d \tag{5.23}$$

$$\text{s.t.} \sum_j x_j^2 = 1, \tag{5.24}$$

$$\text{s.t.} \sum_j x_j^4 = 1 - S_C^{(2)} = c, \tag{5.25}$$

with the additional symmetry constraint $x_j = x_{d-j}$ for $j = 1, \ldots, d-1$. The optimization can be easily solved for small d using Lagrange multipliers and yields indeed states of the form as the $|\psi_t\rangle$. We conjecture that these considerations might also apply to any pair of mutually unbiased bases because entropic uncertainty relations commonly only consider the absolute overlap of basis states but not their phases [96, 214]. In the cases where the $|\psi_t\rangle$-like states are optimal, the result can be generalized to Tsallis-q and Rényi-α entropies with $q, \alpha \geq 2$ as in the proof of Theorem 5.2.

5.3.3 Entropic bound for separable states

In this section, we determine the entropic bound for separable states and local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$. Theorem V.2 from Ref. [212], which shows that for any state ρ there is a pure state $|\psi\rangle$ such that $f_1(\psi) \leqslant f_1(\rho)$ and $f_2(\psi) \leqslant f_2(\rho)$, cannot be applied to separable states. This is because the boundary of the space of separable states is determined by positivity as well as separability conditions. While the former implies that states on the boundary are of lower rank, the latter gives a different constraint. However, this can still be used to simplify the optimization process.

Theorem 5.4. *Let f_1, f_2 be two continuous concave functions on the state space. Then, for every separable state ρ, there exists a separable state ρ^* of the form*

$$\rho^* = (1-p)\,|ab\rangle\langle ab| + p\,|cd\rangle\langle cd|\,, \tag{5.26}$$

where $0 \leqslant p \leqslant 1$ and $|ab\rangle\langle ab|$, $|cd\rangle\langle cd|$ are pure product states, such that $f_1(\rho^*) \leqslant f_1(\rho)$ and $f_2(\rho^*) \leqslant f_2(\rho)$.

Proof. In the range of ρ, we consider some state σ on the boundary of the space of separable states in this subspace. Then, there is some antipode $\sigma^{\blacktriangledown}$ defined as $\frac{1}{\lambda}\left[\rho - (1-\lambda)\sigma\right]$ for the smallest λ such that this expression still describes a separable state. By this definition, obviously, also $\sigma^{\blacktriangledown}$ lies on the boundary. Now, σ can be converted continuously into $\sigma^{\blacktriangledown}$ by a curve $t \mapsto \gamma(t)$ on the boundary where $\gamma(0) = \sigma$ and $\gamma(1) = \sigma^{\blacktriangledown}$, as long as the boundary is connected (see Fig.5.4). Since the functions are continuous, there must be some $t^* \in [0,1]$ such that $f_1[\gamma(t^*)] = f_1(\rho)$. At this point, either $f_2[\gamma(t^*)] \leqslant f_2(\rho)$ or it holds that $f_1[\gamma^{\blacktriangledown}(t^*)] \leqslant f_1(\rho)$ and $f_2[\gamma^{\blacktriangledown}(t^*)] \leqslant f_2(\rho)$ since otherwise concavity implies the contradiction

$$\begin{aligned} f_i(\rho) &\geqslant [1-\lambda(t)]f_i[\gamma(t)] + \lambda(t)f_i[\gamma^{\blacktriangledown}(t)] && (5.27)\\ &> [1-\lambda(t)]f_i(\rho) + \lambda(t)f_i(\rho) = f_i(\rho) && (5.28)\end{aligned}$$

for $i = 1$ or $i = 2$. Thus, we find a state γ^* with $f_{1,2}[\gamma^*] \leqslant f_{1,2}(\rho)$. Compared to ρ, this boundary state γ^* satisfies at least one additional constraint of the form

$$\gamma^*\,|\phi_0\rangle = 0, \qquad\qquad \mathrm{Tr}(\gamma^* W) = 0, \tag{5.29}$$

where $|\phi_0\rangle$ is an eigenstate of γ^* and W is an entanglement witness, because γ^* lies at the positivity or separability boundary, respectively.

Decomposing γ^* into pure product states $\gamma^* = \sum_j p_j\,|a_jb_j\rangle\langle a_jb_j|$, every $|a_jb_j\rangle\langle a_jb_j|$ satisfies the constraints individually. This is because the range of each of them has to be contained in the range of γ^*, and furthermore, for product states it holds that $\mathrm{Tr}(|a_jb_j\rangle\langle a_jb_j|\,W) \geqslant 0$ and since we have $0 = \mathrm{Tr}(\gamma^* W) = \sum_j p_j\,\mathrm{Tr}(|a_jb_j\rangle\langle a_jb_j|\,W)$, also $\mathrm{Tr}(|a_jb_j\rangle\langle a_jb_j|\,W) = 0$.

Thus, we can apply this procedure repeatedly, considering only the state space defined by the already accumulated constraints of the form given in Eq. (5.29). In the end, we either have a pure product state ρ^* or a one-dimensional state space spanned by two pure product states $|ab\rangle\langle ab|$ and $|cd\rangle\langle cd|$, whose boundary is disconnected and hence, the scheme cannot be applied anymore. This might indeed happen, as there are two-dimensional subspaces of the two-qubit space that contain exactly two product

vectors [215]. Either way, for any separable state ρ we find a state ρ^* of the form $\rho^* = (1-p)\,|ab\rangle\langle ab| + p\,|cd\rangle\langle cd|$ such that $f_1(\rho^*) \leqslant f_1(\rho)$ and $f_2(\rho^*) \leqslant f_2(\rho)$. □

Thus, in the case of local $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ measurements, we can restrict the optimization further to real states $|ab\rangle$ and $|cd\rangle$.

Observation 5.5. For any separable state ρ, there is a state $\rho^* = (1-p)\,|ab\rangle\langle ab| + p\,|cd\rangle\langle cd|$ where $0 \leqslant p \leqslant 1$ and $|ab\rangle$ and $|cd\rangle$ are pure and real product states such that $S_{xx}^{(\tilde{q})}(\rho^*) \leqslant S_{xx}^{(\tilde{q})}(\rho)$ and $S_{zz}^{(q)}(\rho^*) \leqslant S_{zz}^{(q)}(\rho)$ for any $q, \tilde{q} \in \mathbb{R}$.

Proof. Using Theorem 5.4, we immediately find a state $\sigma = (1-p)\,|ab\rangle\langle ab| + p\,|cd\rangle\langle cd|$ such that $S_{xx}^{(\tilde{q})}(\sigma) \leqslant S_{xx}^{(\tilde{q})}(\rho)$ and $S_{zz}^{(q)}(\sigma) \leqslant S_{zz}^{(q)}(\rho)$ for any $q, \tilde{q} \in \mathbb{R}$. However, the states $|ab\rangle$ and $|cd\rangle$ might not be real.

A general one-qubit state can be written as $|a\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|1\rangle$ where $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \varphi < 2\pi$. The corresponding probabilities for σ_x and σ_z measurements are then given by

$$p_0 = \cos^2\frac{\theta}{2},\ p_1 = \sin^2\frac{\theta}{2}, \tag{5.30}$$

$$p_\pm = \frac{1}{2} \pm \frac{1}{2}\sin\theta\cos\varphi. \tag{5.31}$$

Hence, by just varying φ, p_0 and p_1 remain unaffected, while $p_+ = (1-\alpha)P_{\max} + \alpha P_{\min}$ and $p_- = \alpha P_{\max} + (1-\alpha)P_{\min}$, where $P_{\max} = \frac{1}{2} + \frac{1}{2}\sin\theta$ and $P_{\max} = \frac{1}{2} - \frac{1}{2}\sin\theta$, vary continuously with $0 \leqslant \alpha \leqslant 1$. Now, consider varying the state $|a\rangle$ in such a way while leaving $|b\rangle$ and $|cd\rangle$ the same. Obviously, the probability distribution for the $\sigma_z \otimes \sigma_z$ measurement on σ stays unchanged. The $\sigma_x \otimes \sigma_x$ measurement, on the other hand, yields $(1-\alpha)\boldsymbol{p_1} + \alpha\boldsymbol{p_2}$ for some probability distributions $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$. Hence, the optimization problem over α is an optimization over a convex set of probabilities. As the entropies are concave functions of probability distributions, the optimum can be found at the boundary. Note that we only optimize the $S_{xx}^{\tilde{q}}$ while leaving S_{zz}^{q} unchanged. Thus, $|a\rangle$ can be chosen real and so can $|b\rangle$, $|c\rangle$ and $|d\rangle$. □

Reducing the optimization to real states of rank at most two, the lower number of parameters allows for robust numerical analysis. This suggests that for $q, \tilde{q} \geqslant 2$, the boundary is reached by real pure product states of the form

$$|\phi_\theta^{q,\tilde{q}}\rangle = \Big(\cos\frac{\theta}{2}\,|0\rangle + \sin\frac{\theta}{2}\,|1\rangle\Big)^{\otimes 2}, \tag{5.32}$$

which, for $q = \tilde{q} = 2$, leads to the boundary for separable states of

$$S_{zz}^{(2)}(\rho) \geqslant -\frac{9}{4} + 3\sqrt{1 - S_{xx}^{(2)}(\rho)} + S_{xx}^{(2)}(\rho), \tag{5.33}$$

shown in Fig. 5.1. In comparison, in the case of Shannon entropy, numerical analysis indicates that the boundary is realized by the states

$$|\phi_\theta^S\rangle = |0\rangle \otimes \Big(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\Big), \tag{5.34}$$

$$|\psi_\theta^S\rangle = \Big(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\Big) \otimes |+\rangle, \tag{5.35}$$

also shown in Fig. 5.1.

5.3.4 Robustness

In the previous sections, we showed that the accessible regions in the entropy plot $(S_{xx}^{(q)}, S_{zz}^{(\tilde{q})})$ are different for general two-qubit states and separable states when $q, \tilde{q} \geqslant 2$. Thus, these entropies provide a scrambling-invariant method to detect entanglement. The accessible regions for $q = \tilde{q} = 2$ are shown in Fig. 5.1.

We investigate the robustness of this detection method for different $q = \tilde{q} \geqslant 2$. The robustness is quantified by the amount of white noise that can be added to the boundary states defined in Eq. (5.3) such that they are still detectable. Numerical analysis indicates that independent of q, the most robust states are those with $S_{xx}^{(q)} = S_{zz}^{(q)}$, i.e. $t = 3$. For states $\rho_{\lambda,t} = (1-\lambda)|\psi_t\rangle\langle\psi_t| + \lambda\frac{\mathbb{1}}{4}$, it also holds that $S_{xx}^{(q)}(\rho_{\lambda,t}) = S_{zz}^{(q)}(\rho_{\lambda,t})$ independent of λ and hence, they enter the region of separable states at the point of the symmetric real pure product state $\big[\frac{1}{\sqrt{1+s^2}}(s|0\rangle + |1\rangle)\big]^{\otimes 2}$ where $s = 1 + \sqrt{2}$. The maximal noise level λ is then determined by

$$\begin{aligned} &\Big(\frac{(1-\lambda)t}{\sqrt{3+t^2}} + \frac{\lambda}{4}\Big)^{2q} + 3\Big(\frac{(1-\lambda)}{\sqrt{3+t^2}} + \frac{\lambda}{4}\Big)^{2q} \\ &= \Big(\frac{s^2}{1+s^2}\Big)^{2q} + 2\Big(\frac{s}{1+s^2}\Big)^{2q} + \Big(\frac{1}{1+s^2}\Big)^{2q} \end{aligned} \tag{5.36}$$

which can be solved analytically for large q. In the limit of $q \to \infty$, $\lambda = \frac{1}{11}(10 - \sqrt{2} - \sqrt{12} - \sqrt{24}) \approx 0.020$. Note that this is an upper bound on the robustness, since the boundary of the region of separable states was only determined numerically in the last section. However, even this upper bound is rather small and the method is not very robust. Finally, we see that the method is most robust for large q, but the limit is reached very fast.

5.3.5 Measurement scheme

Although entropies are not observable in the standard quantum mechanical measurement scheme, similar to measuring the Rényi-2 entropy of a quantum state using two copies of the state [216, 217], which has been implemented experimentally [218], also the Tsallis-2 or Rényi-2 entropy of the local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ can be directly obtained by a measurement on two copies of the quantum state.

To simplify the discussion, we restrict the following description to local measurements $\sigma_x \otimes \sigma_x$, however, local measurements $\sigma_z \otimes \sigma_z$ work analogously. We need the notion of the swap gate $V |\psi\rangle |\phi\rangle = |\phi\rangle |\psi\rangle$ and the decohering channel $\Delta_x(\rho) = \sum_j \langle j_x| \rho |j_x\rangle |j_x\rangle \langle j_x|$, where $|j_x\rangle$ is the basis $|\pm\pm\rangle$. Since $\Delta_x(\rho)$ has eigenvectors $|j_x\rangle$, the spectrum λ_j of $\Delta_x(\rho)$ are the probabilities $p_{\pm\pm}$. It is known that $\mathrm{Tr}\{[\Delta_x(\rho) \otimes \Delta_x(\rho)]V\} = \sum_j \lambda_j^2$ [216]. Moreover,

$$\mathrm{Tr}[(\rho \otimes \rho)(\Delta_x \otimes \Delta_x)(V)] = \mathrm{Tr}\{[\Delta_x(\rho) \otimes \Delta_x(\rho)]V\} = \sum_j \lambda_j^2 = \sum_{\pm\pm} p_{\pm\pm}^2. \tag{5.37}$$

Thus, measuring the expectation value of $(\Delta_x \otimes \Delta_x)(V)$ on two copies of a state ρ gives direct access to the Tsallis-2 or Rényi-2 entropy of the local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$. A similar approach can be used to determine these entropies through randomized measurements on a single system [219].

5.4 Scrambling-invariant families of entanglement witnesses

Inspired by the probability distributions of the states defined in Eq. (5.3), we define a scrambling-invariant family of entanglement witnesses. In the most general form, with local measurements $\sigma_x \otimes \sigma_x$, $\sigma_y \otimes \sigma_y$, and $\sigma_z \otimes \sigma_z$, they are given by

$$W = \mathbb{1} + \alpha |x_1 x_2\rangle \langle x_1 x_2| + \beta |y_1 y_2\rangle \langle y_1 y_2| + \gamma |z_1 z_2\rangle \langle z_1 z_2|, \tag{5.38}$$

where $|x_j\rangle \in \{|+\rangle, |-\rangle\}$, $|y_j\rangle \in \{|i_+\rangle, |i_-\rangle\}$, and $|z_j\rangle \in \{|0\rangle, |1\rangle\}$ for $j = 1, 2$. The key observation is that if for fixed α, β and γ this yields an entanglement witness, then also every other choice of x_j, y_j and z_j results in an entanglement witness. This is because using only local unitary transformations and the partial transposition, the witnesses can be transformed into each other. Consider for example $W = \mathbb{1} + \alpha |+-\rangle \langle +-| + \beta |i_+ i_+\rangle \langle i_+ i_+| + \gamma |10\rangle \langle 10|$, and the transformations $U_A = \sigma_x$, and $U_B = \mathbb{1}$. Then, $U_A^\dagger \otimes U_B^\dagger W^{T_A} U_A \otimes U_B = \mathbb{1} + \alpha |+-\rangle \langle +-| + \beta |i_+ i_+\rangle \langle i_+ i_+| + \gamma |00\rangle \langle 00|$ and one can directly check that any other witness can also be reached.

Indeed, such mappings correspond to permutations of the probabilities, as

$$\langle W \rangle = 1 + \alpha p_{x_1 x_2} + \beta p_{y_1 y_2} + \gamma p_{z_1 z_2}. \tag{5.39}$$

So, for evaluating such a witness from scrambled data, one can simply choose the probabilities appropriately in order to minimize the mean value of the witness.

As a remark, for $\alpha, \beta, \gamma < 0$, the witnesses are additionally related to an entropic uncertainty relation for the min-entropy $S^\infty(\boldsymbol{p}) = -\log \max_j p_j$ since the smallest expectation value of the corresponding family of entanglement witnesses can be written as

$$\langle W \rangle = 1 + \alpha e^{-S_{xx}^{(\infty)}(\rho)} + \beta e^{-S_{yy}^{(\infty)}(\rho)} + \gamma e^{-S_{zz}^{(\infty)}(\rho)}. \tag{5.40}$$

We want to find optimized α, β and γ such that W is an entanglement witness tangent to the space of separable states, i.e., there exists a separable state with $\langle W \rangle = 0$. In other words, we want W to be weakly optimal as described in Section 2.5. In the following analysis, we restrict ourselves to only two measurements and witnesses of the form

$$W = \mathbb{1} + \alpha \left|{+}{+}\right\rangle \left\langle{+}{+}\right| + \gamma \left|00\right\rangle \left\langle 00\right|. \tag{5.41}$$

First of all, we need to ensure that $\langle W \rangle \geqslant 0$ for all separable states. In order to obtain an optimal witness, we further need to adjust α and γ such that for some separable state $\langle W \rangle = 0$.

The optimal values for α and γ are found by optimizing $\min_{\rho_s} \mathrm{Tr}(\rho_s W)$ over separable states ρ_s for all α and γ. Because of linearity, we only need to consider general pure product states $|\psi_A\rangle \otimes |\psi_B\rangle$ where

$$|\psi_{A/B}\rangle = \cos \frac{\theta_{A/B}}{2} |0\rangle + e^{i\phi_{A/B}} \sin \frac{\theta_{A/B}}{2} |1\rangle \tag{5.42}$$

with $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \phi < 2\pi$. It turns out that for $\frac{\gamma}{\alpha} \geqslant -3 - 2\sqrt{2}$, the optimal state is given by $\phi_A = \phi_B = 0$ and $\theta_A = \theta_B$, while $\phi_A = \phi_B = 0$ and $\theta_A - \frac{3\pi}{4} = \frac{3\pi}{4} - \theta_B$ needs to be considered in the case of $\frac{\gamma}{\alpha} \leqslant -3 - 2\sqrt{2}$.

Finally, we have to ensure that there exist entangled states with $\langle W \rangle < 0$. Since $\langle W \rangle = 1 + \alpha p_{++} + \gamma p_{00}$ and the probabilities are nonnegative, either α or γ must necessarily be negative. In that case, the eigenvector of W corresponding to the smallest eigenvalue is indeed given by the entangled state $|\psi_t\rangle = \frac{1}{\sqrt{3+t^2}}(t, 1, 1, 1)^T$ with $t = -(\alpha - 2\gamma + 2\sqrt{\alpha^2 - \alpha\gamma + \gamma^2})/\alpha$.

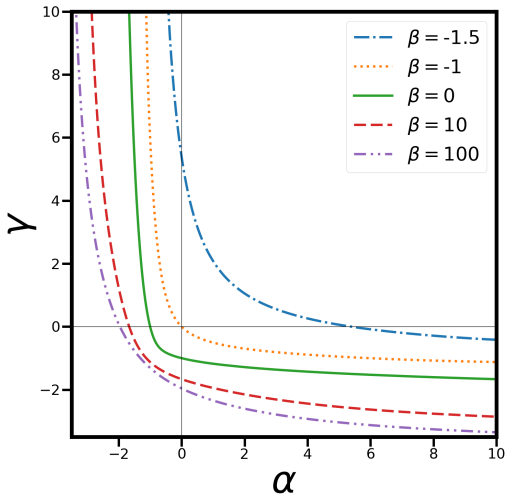


FIGURE 5.5: [199] Optimized values for the parameters α and γ for different β in entanglement witnesses of the form $W = \mathbb{1} + \alpha |x_1x_2\rangle\langle x_1x_2| + \beta |y_1y_2\rangle\langle y_1y_2| + \gamma |z_1z_2\rangle\langle z_1z_2|$. Here, optimized means that for some separable state $\langle W\rangle = 0$.

The resulting curve of optimal α and γ in the case of $\beta = 0$ can thus be obtained analytically and is shown in Fig. 5.5. More generally, for witnesses of the form in Eq. (5.38) where $\beta \neq 0$, we find the optimal parameters numerically.

5.5 Nonconvex structure of the nondetectable state space

For many methods of entanglement detection, it is crucial that the set of separable states is convex. For instance, the existence of a witness for any entangled state ρ relies on this fact. This convexity is also present in the case of restricted measurements, which are not tomographically complete. If there is a way to detect the entanglement from a restricted set of measurements, it can be done with an entanglement witness [220]. In this section, we show that this is not the case when only scrambled data is available.

In order to test whether there would in principle be a method to detect the entanglement of a specific state using only scrambled data from local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, we use the fact that the PPT criterion is necessary and sufficient in the two-qubit case [57]. Thus, we can formulate the problem as a family of semi-definite

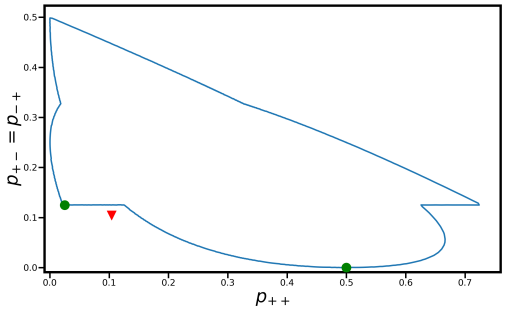


FIGURE 5.6: [199] Projection of the set of possibly separable states (blue line) for local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, where $p_{++} = p_{00}$, $p_{+-} = p_{-+} = p_{01} = p_{10}$, and $p_{--} = p_{11}$, onto the coordinates (p_{++}, p_{+-}). Clearly, this set is nonconvex. The green dots and the red triangle correspond to an explicit counterexample to the convexity, as explained in the main text.

programs [90]. We consider the problem

$$\begin{aligned} &\min_{\rho} 0 \\ &\text{s.t. } \operatorname{Tr}\rho = 1, \\ &\quad \rho \geqslant 0, \\ &\quad \rho^{T_B} \geqslant 0, \\ &\quad \rho \text{ realizes one of the } (4!)^2 \text{ permutations} \\ &\qquad \text{of the given probability distribution} \\ &\qquad \text{for measurements } \sigma_x \otimes \sigma_x \text{ and } \sigma_z \otimes \sigma_z. \end{aligned} \tag{5.43}$$

This is a so called feasibility problem: If a state ρ with the desired properties exist, the output of the SDP is zero, and ∞ otherwise. If this family of SDPs fails for all permutations, then there is no separable state that realizes the same scrambled data as the original state. Hence, the entanglement of such a state is detected. Otherwise, we call the state possibly separable.

In practice, without scrambled data, around 1.2% of all random states according to the Hilbert-Schmidt measure can be shown to be entangled using only local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$. In the case of scrambled data, we tested approximately $130,000,000$ random mixed states and found around 3000 detectable states using the corresponding scrambled data. Note that for the implementation it is possible to reduce the number of permutations that need to be considered to just 18, as local relabeling of the outcomes or the exchange of qubits can be neglected. Out of the 3000 states, only six can be detected using the scrambling-invariant entanglement witnesses and none using the entropic uncertainty relations where $q = \tilde{q}$. The reason for this poor performance is the nonconvex structure of the set of nondetectable states, as we discuss now.

First, we note that the set of possibly separable states is star-convex around the maximally mixed state $\frac{\mathbb{1}}{4}$. This can be seen as follows: If a state ρ is part of the set of possibly separable states, there is a separable state σ that realizes the same probability distribution as ρ up to a permutation. Then, $\lambda\sigma + (1-\lambda)\frac{\mathbb{1}}{4}$ is still separable for $0 \leqslant \lambda \leqslant 1$ and realizes the same probability distribution as $\lambda\rho + (1-\lambda)\frac{\mathbb{1}}{4}$ up to the same permutation as before.

This fact can be used to characterize the boundary of the possibly-separable state space by starting with the maximally mixed state and mixing it with detectable states until the mixture becomes detectable. To illustrate the nonconvexity of this set, we assume first that it is convex. Then, the intersection with any convex set, for example the set of states with probabilities $p_{++} = p_{00}$, $p_{+-} = p_{-+} = p_{01} = p_{10}$, and $p_{--} = p_{11}$, would again form a convex set. Furthermore, the projection onto the coordinates (p_{++}, p_{+-}) would be convex. This projection is shown in Fig. 5.6. Clearly, it is nonconvex, and hence, the initial assumption is incorrect. To make this statement independent of numerical analysis, we provide an explicit counterexample.

Observation 5.6. The set of possibly separable states for local measurements $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ is nonconvex.

Proof. The states

$$\begin{aligned}\rho_1 =& \frac{1}{4}\Big[\mathbb{1}\otimes\mathbb{1} - \frac{7}{10}(\mathbb{1}\otimes\sigma_x + \sigma_x\otimes\mathbb{1} + \mathbb{1}\otimes\sigma_z + \sigma_z\otimes\mathbb{1}) \\ &+ \frac{1}{2}(\sigma_x\otimes\sigma_x + \sigma_z\otimes\sigma_z + \sigma_x\otimes\sigma_z + \sigma_z\otimes\sigma_x)\Big]\end{aligned} \tag{5.44}$$

and $\rho_2 = |\Phi^+\rangle\langle\Phi^+|$ where $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ realize probability distributions corresponding to the left and right green dot in Fig. 5.6, respectively. While ρ_1 is separable, the product state $|+\rangle|0\rangle$ realizes the same scrambled data as ρ_2 and hence, ρ_2 is possibly separable. Thus, both are part of the possibly separable state space. However, the mixture $\rho = \frac{5}{6}\rho_1 + \frac{1}{6}\rho_2$, shown as a red triangle in Fig. 5.6, is detectable. The scrambled data of the corresponding probability distribution $p_{++} = p_{+-} = p_{-+} = p_{00} = p_{01} = p_{10} = \frac{5}{48}$ and $p_{--} = p_{11} = \frac{33}{48}$ cannot origin from a separable state. The witnesses $W = \mathbb{1} \pm \sigma_x\otimes\sigma_x \pm \sigma_z\otimes\sigma_z$ certify the entanglement for all permutations. □

5.6 Conclusion

We have introduced the concept of scrambled data, meaning that the assignment of probabilities to outcomes of the measurements is lost. Clearly, this restriction limits

the possibilities of entanglement detection. Nevertheless, we have shown that using entropies and entanglement witnesses one can still detect the entanglement in some cases. These methods are limited, however, as the set of states whose scrambled data can be realized by separable states is generally not convex.

There are several directions in which our work may be extended or generalized. First, one may consider more general scenarios than the two-qubit situation considered here, such as the case of three or more particles. Second, it would be interesting to study our results on entropies further, in order to derive systematically entropic uncertainty relations for various entropies. Such entropic uncertainty relations find natural applications in quantum information theory, e.g., in the security analysis of quantum key distribution. Finally, it would be intriguing to connect our scenario to Bell inequalities. This could help to relax assumptions on the data for non-locality detection and device-independent quantum information processing.

6 Confident entanglement detection via numerical range

Prerequisites

2.2 Quantum mechanics
2.5 Entanglement
2.6 Numerical range
2.8 Semidefinite programming

6.1 Introduction

The main parts of this chapter have been published as Publication (E) [221]. We investigate the joint (separable) numerical range of multiple measurements, i.e., the regions of expectation values accessible with (separable) quantum states for given observables. This not only enables efficient entanglement detection, but also sheds light on the geometry of quantum states. More precisely, in an experiment, if the confidence region for the obtained data and the separable numerical range are disjoint, entanglement is reliably detected. Generically, the success of such an experiment is more likely the smaller the separable numerical range is compared to the general numerical range. We quantify this relation using the volume ratio and show that it cannot be arbitrarily small, giving analytical bounds for any number of particles, local dimensions as well as number of measurements. Moreover, we explicitly compute the volume of separable and general numerical range for two locally traceless two-qubit product observables, which are of particular interest as they are easier to measure in practice. Furthermore, we consider specific examples of extreme volume ratios and the relation between commutativity and entanglement detection.

Verifying entanglement in experiments is essential and many methods to do this have been developed [52]. Here, we consider multiple observables and the measurements of their expectation values. The accessible regions of expectation value vectors for (separable) quantum states is given by the (separable) numerical range. If, for a given state, the measurement results give a point outside the separable numerical range,

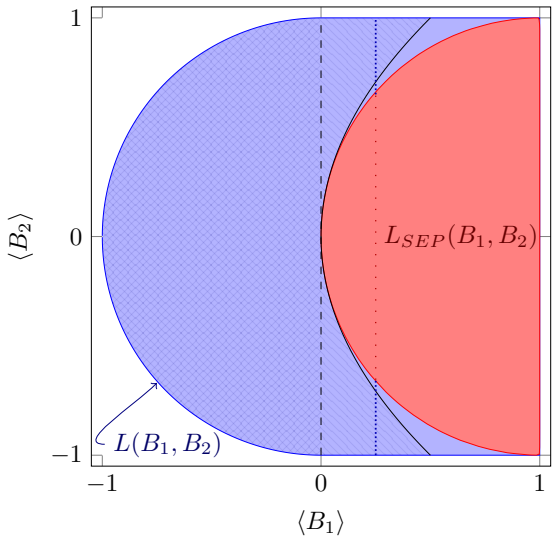


FIGURE 6.1: An illustration of different variants of entanglement witnesses. $L(B_1, B_2)$ and $L_{\text{SEP}}(B_1, B_2)$ are the (separable) numerical range of the two-qubit observables $B_1 = |00\rangle\langle 11| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 00|$ and $B_2 = -i|00\rangle\langle 11| + |01\rangle\langle 01| - |10\rangle\langle 10| + i|11\rangle\langle 00|$. The dashed line indicates the entanglement witness $W = B_1$, the dotted line the ultrafine entanglement witnesses with $\langle W_1 \rangle = \langle B_1 \rangle = 1/4 = \omega_1$ and $W_2 = \sqrt{7}/4 - B_2$ or $W_2 = (1 - \sqrt{7})/4 + B_2$, and the curved solid line the nonlinear entanglement witness $\langle W_{\text{NL}} \rangle = \langle B_1 \rangle - \langle B_2 \rangle^2$.

then that state must have been entangled. Thus, we first provide insight into how this approach can be used for entanglement detection. Second, since the (separable) numerical range is ultimately an affine transformation of a projection of the (separable) quantum state space, our investigation also sheds light on the geometry of quantum states, especially, on the relation between the separable and the general quantum state space geometry. Finally, in practical experiments, statistical and systematic errors lead to a confidence region instead of a single point contained in the numerical range of the measurements. We compare the volumes of separable and general numerical range to gain intuition on how useful the considered measurements are for entanglement detection in practical scenarios.

Entanglement witnesses are the standard tool for entanglement detection and employed ubiquitously [52]. Although a single measurement, that is repeated sufficiently many times, suffices to detect entanglement, this measurement, i.e. the entanglement witness, might be a highly entangled observable and hard to implement in practice. Ultrafine entanglement witnesses extend the concept of entanglement witnesses by taking into account multiple measurements for more reliable detection [222, 223]. Also, measurements that are easier to implement, such as product observables, can be combined to simplify the detection in the experiment. To do so, a first observable W_1 determines via measurement a subset of states that need to be considered contrained by $\operatorname{Tr} W_1 \rho = \omega_1$, where ω_1 is the obtained measurement result. Then, the

second observable W_2 only needs to satisfy that $\operatorname{Tr} W\rho_{\text{sep}} \geq 0$ for all separable states ρ_{sep} with $\operatorname{Tr} W_1\rho_{\text{sep}} = \omega_1$ allowing for more effective entanglement detection. Lastly, nonlinear entanglement witnesses combine multiple measurements in a nonlinear way to improve entanglement detection [224–227]. While ordinary and ultrafine entanglement witnesses only provide a polyhedral approximation to the separable numerical range, nonlinear witnesses observe the structure of the convex set generally better; see Fig. 6.1 for a schematic visualization of the different methods.

In contrast to the different variants of entanglement witnesses, for the general situation of multiple measurements, the joint (separable) numerical range provides a comprehensive framework to tackle the problem of entanglement detection as it contains all information accessible [75, 228–230]. Throughout the Chapter, we are going to use the notation introduced in Section 2.6. Furthermore, we define the volume of the (restricted) joint numerical range $L_X(A_1, \ldots, A_k)$ as $\operatorname{vol} L_X$ using the Euclidean norm of $\mathbb{R}^k$. This is the natural volume measure for experiments since it is the relevant measure for the confidence region obtained by measuring a given quantum state.

6.2 Experimental confidence region

An entangled state whose entanglement is in principle detectable by observables $A_1, \ldots, A_k$ corresponds to a point in the numerical range which is outside the separable numerical range. In a realistic experiment, however, only a finite number of measurement results can be collected which leads to an α-confidence region inside the numerical range that covers the exact point given by the underlying state with probability at least $1 - \alpha$. Such a confidence region can be obtained, e.g., using Hoeffding's tail inequality [231]. In this case, it holds that the probability of an estimator a_j for the expectation value of A_j having at least a distance of t from the actual expectation value is bounded by

$$P\left(\left|a_j - \operatorname{Tr}\rho A_j\right| \geq t\right) \leq 2\exp\left[-\frac{2mt^2}{\left(\lambda_1(A_j) - \lambda_{-1}(A_j)\right)^2}\right], \tag{6.1}$$

where the estimator a_j is obtained as the average from m measurements on the state ρ and $\lambda_1(A_j)$, $\lambda_{-1}(A_j)$ are the largest and smallest eigenvalues of A_j, respectively. Since the individual measurements are independent, so are the estimators a_j. Let us rescale the observables such that $\lambda_1(A_j) - \lambda_{-1}(A_j) = 1$. Then, we obtain a confidence region in the form of a hyperrectangle via

$$P\left(\exists j : \left|a_j - \operatorname{Tr}\rho A_j\right| \geq t_j\right) \leq 1 - \Pi_j\left[1 - 2\exp\left(-2mt_j^2\right)\right], \tag{6.2}$$

where each observable is measured m times, and hence there are km measurements done in total. The shape of the rectangle can be adjusted by choosing appropriate t_j. Independent from the specific shape and origin of the experimenter's confidence region, they can exclude a separable quantum state as cause of the data with statistic significance only if the separable numerical range and the confidence region are disjoint. Thus, generically, choosing observables such that the volume ratio between the separable and the general numerical range is small provides a higher statistical significance for entanglement detection. This is because the confidence region is more likely to lie outside the separable numerical range. Importantly, this question is different from maximizing the number, i.e. the volume, of entangled states that can be detected by infinite repetition of the measurements with infinite precision.

This reasoning motivates us to investigate the volume ratio of (separable) numerical range. As it turns out, it cannot be arbitrarily small and we provide bounds for any number of particles, local dimensions, and number of observables. Moreover, we focus on product observables since they are easier to implement in experiments. For two qubits and two locally traceless product observables, we provide explicit expressions for the volume of their general and separable numerical range.

6.3 Minimal volume ratio

We investigate minimal volume ratios of the separable numerical range compared to the general numerical range that can be reached for given number of particles, local dimensions, and number of measurements. More precisely, we define:

Definition 6.1. For k independent measurements on a quantum system consisting of n particles and local dimensions $\boldsymbol{d} = (d_1, \ldots, d_n)$, we denote the minimal volume ratio of the separable numerical range compared to the general numerical range as

$$\mu_{n,\boldsymbol{d},k} = \min_{A_1,\ldots,A_k} \frac{\operatorname{vol} L_{\mathrm{SEP}}(A_1, \ldots, A_k)}{\operatorname{vol} L(A_1, \ldots, A_k)}. \tag{6.3}$$

If $d_1 = \cdots = d_n = d$, we just write d instead of $\boldsymbol{d}$. Also, we denote the total dimension of the Hilbert space as $D = d_1 \cdots d_n$.

As we discussed in the introduction, measurements reaching the minimal volume ratio are in some sense optimal for entanglement detection in practical experiments. Thus, we find lower bounds for $\mu_{n,\boldsymbol{d},k}$ as well as measurements with low $\mu_{n,\boldsymbol{d},k}$, consequently also providing upper bounds.

First, we consider the case in which the observables $A_1, \ldots, A_k$ are not linearly independent. Then, the (separable) numerical range is contained in a lower-dimensional manifold and the volume in $\mathbb{R}^k$ vanishes. However, we can still define the relative volume comparing the volumes of the manifolds. More specifically, we have the following result.

Proposition 6.2. *For Hermitian observables $A_1, \ldots, A_k$, let $B_1, \ldots, B_{k'}$ be a maximal linearly independent subset of $\{A_1 - \frac{\mathbb{1}}{D} \operatorname{Tr} A_1, \ldots, A_{D-1} - \frac{\mathbb{1}}{D} \operatorname{Tr} A_{D-1}\}$. Then, it holds that*

$$\frac{\operatorname{vol} L_{\mathrm{SEP}}(A_1, \ldots, A_k)}{\operatorname{vol} L(A_1, \ldots, A_k)} = \frac{\operatorname{vol} L_{\mathrm{SEP}}(B_1, \ldots, B_{k'})}{\operatorname{vol} L(B_1, \ldots, B_{k'})}, \tag{6.4}$$

i.e., we can simply ignore observables that are linearly dependent.

Proof. Adding multiples of the identity to the observables corresponds to a translation of the (separable) numerical range, and hence, does not change the volume ratio. More generally, affine transformations also do not change the relative volume; for details, see the proof of Proposition 6.5. Let $A_j - \frac{\mathbb{1}}{D}$ depend linearly on the $B_1, \ldots, B_k$, i.e., $A_j - \frac{\mathbb{1}}{D} = \sum_l x_l B_l$. Then, we apply the transformation $a_j \to \tilde{a}_j = a_j - \sum_l x_l b_l$, where a_j is the variable corresponding to the observable $A_j - \frac{\mathbb{1}}{D}$ and similarly for b_j corresponding to B_j. That means, we have

$$\begin{aligned} \tilde{a}_j &= a_j - \sum_l x_l b_l \\ &= \operatorname{Tr}\left[(A_j - \frac{\mathbb{1}}{D}) - \sum_l x_l B_l \right] \rho = 0, \end{aligned} \tag{6.5}$$

for any state ρ. Hence, we obtain the volume in the subspace given by $\tilde{a}_j = 0$ which simply lowers the dimension. Applying this procedure iteratively proves the observation. □

Thus, it suffices to restrict the observables to be linearly independent, traceless, and bounded, which we assume in the following.

The main idea to obtain a general lower bound for the volume ratio $\mu_{n,\boldsymbol{d},k}$ is to compute the volume of the (separable) numerical range via integration using polar coordinates. This idea is inspired by the approach used in Refs. [232, 233], however, while the authors of these works focus on simply finding all the boundary points, i.e., the extreme points whose convex hull forms the (separable) numerical range, we find the boundary point in a certain direction.

Lemma 6.3. *For Hermitian operators $A_1, \ldots, A_k$ and a star-convex state set X around the maximally mixed state, the k-dimensional volume of the restricted numerical range is given by*

$$\operatorname{vol} L_X = \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{k-2} \, \frac{1}{k} R^k \Pi_{j=1}^{k-2} \sin^j \vartheta_j, \tag{6.6}$$

where the radius $R(\varphi, \vartheta_1, \ldots, \vartheta_{k-2})$ is given by

$$\begin{aligned} R = \max_{\rho \in X} \quad & \operatorname{Tr} \rho[\hat{\boldsymbol{r}} \cdot \boldsymbol{A}] \\ \text{s.t.} \quad & \operatorname{Tr} \rho[\hat{\boldsymbol{\varphi}} \cdot \boldsymbol{A}] = 0, \\ & \operatorname{Tr} \rho[\hat{\boldsymbol{\vartheta}}_j \cdot \boldsymbol{A}] = 0 \text{ for } j = 1, \ldots, k-2, \end{aligned} \tag{6.7}$$

and $\boldsymbol{A} = (A_1 - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_1, \ldots, A_k - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_k)^T$. The vectors $\hat{\boldsymbol{r}}, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\vartheta}}_{\mathbf{1}}, \ldots, \hat{\boldsymbol{\vartheta}}_{\boldsymbol{k-2}}$ are the unit vectors of the k-dimensional polar coordinates.

Proof. The volume V of a k-dimensional (star)-convex (around the origin) region that includes the origin is given via integration as

$$\begin{aligned} V &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{k-2} \int_0^{R} dr \, r^{k-1} \, \Pi_{j=1}^{k-2} \sin^j \vartheta_j \\ &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{k-2} \, \frac{1}{k} R^k \, \Pi_{j=1}^{k-2} \sin^j \vartheta_j, \end{aligned} \tag{6.8}$$

using k-dimensional polar coordinates. Obviously, we have $\operatorname{vol} L(A_1, \ldots, A_k) = \operatorname{vol} L(A_1 - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_1, \ldots, A_k - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_k)$ since the transformation merely translates the restricted joint numerical range. The restricted numerical range $L_X(A_1 - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_1, \ldots, A_k - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_k)$ is a star-convex set around the origin as X is star-convex around the maximally mixed state. Hence, we can calculate its volume using the above geometric formula. The radius $R(\varphi, \vartheta_1, \ldots, \vartheta_{k-2})$ is apparently given by

$$\begin{aligned} R(\varphi, \vartheta_1, \ldots, \vartheta_{k-2}) = \max_{\rho \in X} \quad & \operatorname{Tr} \rho[\hat{\boldsymbol{r}}(\varphi, \vartheta_1, \ldots, \vartheta_{k-2}) \cdot \boldsymbol{A}] \\ \text{s.t.} \quad & \operatorname{Tr} \rho[\hat{\boldsymbol{\varphi}}(\varphi, \vartheta_1, \ldots, \vartheta_{k-2}) \cdot \boldsymbol{A}] = 0, \\ & \operatorname{Tr} \rho[\hat{\boldsymbol{\vartheta}}_j(\varphi, \vartheta_1, \ldots, \vartheta_{k-2}) \cdot \boldsymbol{A}] = 0 \text{ for } j = 1, \ldots, k-2, \end{aligned} \tag{6.9}$$

where $\boldsymbol{A} = (A_1 - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_1, \ldots, A_k - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_k)^T$ and $\hat{\boldsymbol{r}}, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\vartheta}}_{\mathbf{1}}, \ldots, \hat{\boldsymbol{\vartheta}}_{\boldsymbol{k-2}}$ are the unit vectors of the k-dimensional polar coordinates. The constraints of the optimization make sure that it yields the distance of the boundary to the origin in a certain direction given by the angles $\varphi, \vartheta_1, \ldots, \vartheta_{k-2}$. □

This result not only gives an interesting characterization of the volume, it can also be directly implemented using semi-definite programming [90] to efficiently approximate

the volume if the set X can be characterized using semi-definite and linear constraints which, e.g., is the case if X is the set of all quantum states or the set of quantum states with positive partial transpose. In the case of two qubits, the separable states are exactly those with a positive partial transpose, which allows for efficient numerical treatment via semi-definite programming [57].

When we are only interested in the relative volume $\operatorname{vol} L_X / \operatorname{vol} L_Y$, we can obtain a lower bound by comparing the optimizations of the radii. In particular, for X being the set of separable states and Y being the set of all quantum states, we obtain:

Theorem 6.4. *For k observables and n-partite quantum systems with local dimensions $\boldsymbol{d} = (d_1, \ldots, d_n)$ and total dimension $D = d_1 \cdots d_n$, the relative volume of the numerical range restricted to separable states compared to all quantum states is lower bounded by*

$$\mu_{n,\boldsymbol{d},k} \geq \left[\frac{b}{D} \sqrt{\frac{D-1}{D-b^2}} \right]^k, \tag{6.10}$$

where

$$b = \sqrt{\frac{D^n}{(2D-1)^{n-2}(D^2-1)+1}}. \tag{6.11}$$

Moreover, for a bipartite $d \times d$-system, i.e., $n = 2$ and $d_1 = d_2 = d$, it holds that

$$\mu_{2,d,k} \geq \frac{1}{(d^2-1)^k}. \tag{6.12}$$

Proof. For an n-partite quantum system ρ with local dimensions $d_1, \ldots, d_n$ and total dimension $D = d_1 \cdots d_n$, the state $\rho = (1-\epsilon)\frac{\mathbb{1}}{D} + \epsilon\sigma$ is fully separable for any state σ if $\epsilon \leq \frac{b}{D}\sqrt{\frac{D-1}{D-b^2}}$, where $b = \sqrt{\frac{D^n}{(2D-1)^{n-2}(D^2-1)+1}}$ [234, 235]. Also, for bipartite systems with $d_1 = d_2 = d$, the same is true if $\epsilon \leq 1/(d^2-1)$ [236]. Let $\rho^*(\varphi, \vartheta_1, \ldots, \vartheta_{k-2})$ be the optimal state in the maximization that determines $R(\varphi, \vartheta_1, \ldots, \vartheta_{k-2})$ in Lemma 6.3 for X being the set of all quantum states with objective value $R^*(\varphi, \vartheta_1, \ldots, \vartheta_{k-2})$. Then, the state $\tilde{\rho} = \epsilon\rho^* + (1-\epsilon)\frac{\mathbb{1}}{D}$ with maximal ϵ such that full separability can be guaranteed with the above results is a feasible point of the corresponding optimization with X being the set of fully separable states with objective value ϵR^* since $\operatorname{Tr}(A_i - \frac{\mathbb{1}}{\operatorname{Tr}\mathbb{1}} \operatorname{Tr} A_i)\frac{\mathbb{1}}{D} = 0$. Together with Lemma 6.3, this proves the theorem. □

In the simplest example, i.e., a single Hermitian operator A and a two-qubit system ρ, Theorem 6.4 gives a lower bound of $1/3$ for the relative volume $\mu_{2,2,1}$.

A special case is the scenario in which the measurement results determine the underlying quantum state uniquely. Although the volume ratio of the (separable) numerical

range is not directly related to the share of nondetectable states, in the case of quantum state tomography, they coincide.

Proposition 6.5. *Let $A_1, \ldots, A_{D-1}$ be Hermitian observables such that the translated operators $A_1 - \frac{\mathbb{1}}{D} \operatorname{Tr} A_1, \ldots, A_{D-1} - \frac{\mathbb{1}}{D} \operatorname{Tr} A_{D-1}$ are linearly independent. Then, the relative volume of the (separable) numerical range is given by*

$$\frac{\operatorname{vol} L_{\mathrm{SEP}}(A_1, \ldots, A_{D-1})}{\operatorname{vol} L(A_1, \ldots, A_{D-1})} = \frac{\operatorname{vol}_{HS} SEP}{\operatorname{vol}_{HS} ALL}, \tag{6.13}$$

where $\operatorname{vol}_{HS} SEP$ and $\operatorname{vol}_{HS} ALL$ denote the volume of separable and all states, respectively, w.r.t. the Hilbert-Schmidt norm.

Proof. We can express any D-dimensional quantum state ρ in terms of the generalized Gell-Mann matrices [237], i.e.,

$$\rho = \frac{\mathbb{1}}{D} + \sum_{j=1}^{D-1} \mu_j G_j, \tag{6.14}$$

where the $\mu_j \in \mathbb{R}$ are real coefficients and the G_j together with $\frac{\mathbb{1}}{D}$ form a Hermitian ($G_j^\dagger = G_j$), orthonormal ($\operatorname{Tr} G_i G_j = \delta_{ij}$) and traceless ($\operatorname{Tr} G_j = 0$) basis. The distance between two quantum states can be measured using the Hilbert-Schmidt norm $\|A\|_{HS} = \sqrt{\operatorname{Tr} A^\dagger A}$. We obtain

$$\begin{aligned} \|\rho - \sigma\|_{HS} &= \sqrt{\operatorname{Tr}\left[(\rho - \sigma)^2\right]} \\ &= \sqrt{\sum_{i,j=1}^{D-1} (\mu_i - \eta_i)(\mu_j - \eta_j) \operatorname{Tr} G_i G_j} \\ &= \sqrt{\sum_{j=1}^{D-1} (\mu_j - \eta_j)^2} \\ &= |\boldsymbol{\mu} - \boldsymbol{\eta}| \end{aligned} \tag{6.15}$$

for the distance between the quantum states $\rho = \frac{\mathbb{1}}{D} + \sum_{j=1}^{D-1} \mu_j G_j$ and $\sigma = \frac{\mathbb{1}}{D} + \sum_{j=1}^{D-1} \eta_j G_j$. This is the same as the Euclidean distance when we consider the μ_j and η_j as coordinates in $\mathbb{R}^{D-1}$.

Because the $\tilde{A}_j = A_j - \frac{\mathbb{1}}{D} \operatorname{Tr} A_j$ are linearly independent, there exists a dual basis [188] $\{B_j\}_{j=1,\ldots,D-1}$, i.e., $\operatorname{Tr} A_i B_j = \delta_{ij}$, and an invertible matrix Λ such that $G_i = \sum_j \Lambda_{ij} B_j$.

Furthermore, we have that

$$\begin{aligned}\rho_{\boldsymbol{\mu}} &= \frac{\mathbb{1}}{D} + \sum_{j=1}^{D-1} \mu_j G_j = \frac{\mathbb{1}}{D} + \boldsymbol{\mu} \cdot \boldsymbol{G} \\ &= \frac{\mathbb{1}}{D} + \boldsymbol{\mu} \cdot \Lambda \boldsymbol{B} \\ &= \frac{\mathbb{1}}{D} + (\Lambda^T \boldsymbol{\mu}) \cdot \boldsymbol{B},\end{aligned} \tag{6.16}$$

where Λ^T is the tranpose of Λ, which is also invertible. The coefficients $(\Lambda^T \boldsymbol{\mu})_j$ are the coordinates of ρ in the space $L(\tilde{A}_1, \ldots, \tilde{A}_{D-1})$ because we used the dual basis.

In general, a coordinate transformation $(v_1, \ldots, v_n) = \varphi(u_1, \ldots, u_n)$ leads to a change of a volume integral

$$\int_{\boldsymbol{v} \in \varphi(U)} f(\boldsymbol{v}) d^n v = \int_{\boldsymbol{u} \in U} f(\varphi(\boldsymbol{u})) |\det(D\varphi)(\boldsymbol{u})| d^n u, \tag{6.17}$$

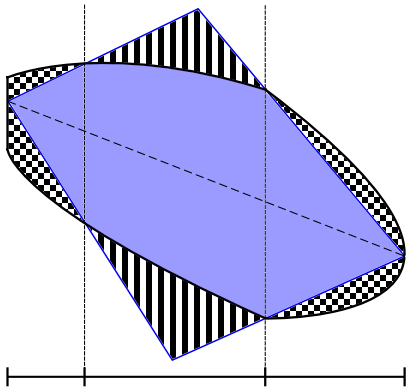
where $\det(D\varphi)$ is the determinant of the Jacobi matrix of φ. In this case, $\varphi(\boldsymbol{u}) = (\Lambda^T)^{-1} \boldsymbol{v}$, which means that

$$\begin{aligned}\mathrm{vol}_{HS}\, X &= \int_{\rho_{\boldsymbol{\mu}} \in X} d^n \mu \\ &= |\det(\Lambda^T)^{-1}| \int_{\rho_{\boldsymbol{\xi}} \in Y(X)} d^n \xi \\ &= |\det(\Lambda^T)^{-1}| \,\mathrm{vol}\, L_X(A_1, \ldots, A_{D-1}),\end{aligned} \tag{6.18}$$

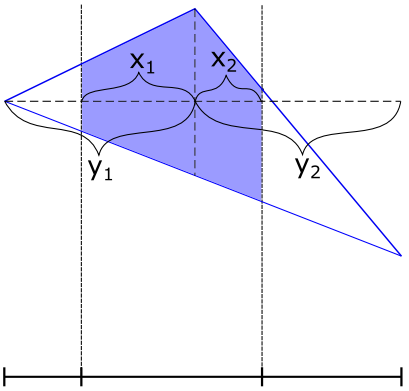
where $\boldsymbol{\xi} = \Lambda^T \boldsymbol{\mu}$ and $Y(X) = \{\rho_{\Lambda^T \boldsymbol{\mu}} \mid \rho_{\boldsymbol{\mu}} \in X\}$. Note that $|\det(\Lambda^T)^{-1}| = \mathrm{const.} > 0$ because the transformation is invertible. Also, since the transformation is linear, it is independent of the variables. Hence, the relative volume does not change as we apply the same transformation independently of X. □

6.4 Geometric considerations

In the case of two qubits and $n = 15$, the volume ratio is strongly believed to be $\frac{\mathrm{vol}_{HS}\, SEP}{\mathrm{vol}_{HS}\, ALL} = 8/33$ [238–241]. In general, we can find a lower bound on the volume ratio considering the ϵ-ball around the maximally mixed state [234, 235] compared to the volume of all states [242]. From this result for $n = D - 1$, it might be possible to obtain bounds for a lower number of observables as well since this corresponds to some projection of convex bodies. Starting from $D - 1$ observables and removing some of them corresponds to taking projections onto lower dimensional subspaces. Then, the idea is that the volume ratio of the projections cannot be arbitrarily small compared



(A) To maximize the volume ratio, the best choice for the inner convex body is an infinite rectangle bounded by the outer body. Furthermore, it is shown how the outer convex body can be transformed to a convex quadrangle with larger volume ratio while not changing the volume ratio of the projection by removing checkerboard areas and adding striped areas.



(B) The volume ratio of the resulting shape can be optimized for both triangles separately. It is given by $\frac{\operatorname{vol} P_u K}{\operatorname{vol} P_u L} = 1 - \frac{(y_1 - x_1)^2}{y_1 y} - \frac{(y_2 - x_2)^2}{y_2 y}$, independent from the vertical positions of the corner points, and direct optimization gives $\frac{\operatorname{vol} P_u K}{\operatorname{vol} P_u L} = 1 - (1 - \frac{x}{y})^2$ with $x = x_1 + x_2$ and $y = y_1 + y_2$ as the largest possible volume ratio for fixed volume ratio $\frac{x}{y}$ of the projections.

FIGURE 6.2: Illustration of the proof of Proposition 6.6.

to the volume ratio of the original convex bodies. The (very much mathematical) geometric question is:

> Given two convex bodies, i.e. convex, compact sets, $K, L \subset \mathbb{R}^n$ such that $K \subset L$ with fixed volume ratio $\frac{\operatorname{vol} K}{\operatorname{vol} L}$. What is the minimal volume ratio of projections $\min_{u_1,\dots,u_l} \frac{\operatorname{vol} P_{u_1,\dots,u_l} K}{\operatorname{vol} P_{u_1,\dots,u_l} L}$, where $P_{u_1,\dots,u_l}$ denotes the projection onto the subspace orthogonal to the vectors $u_1, \dots, u_l$?

The question might be easier when we only consider $l = 1$. The following observation finds the optimal bound in the simplest case, i.e. $n = 2$.

Proposition 6.6. *For two convex sets $K \subset L \subset \mathbb{R}^2$ with fixed volume ratio $\frac{\operatorname{vol} K}{\operatorname{vol} L}$, it holds that*

$$\min_{u \in \mathbb{R}^2} \frac{\operatorname{vol} P_u K}{\operatorname{vol} P_u L} \geq 1 - \sqrt{1 - \frac{\operatorname{vol} K}{\operatorname{vol} L}}. \tag{6.19}$$

This bound is tight.

Proof. The proof is visualized in Fig. 6.2. Instead of minimizing the volume ratio of projections for given convex bodies, we consider two projections $PK \subset PL \subset \mathbb{R}^1$ with given volume ratio and maximize the volume ratio of possible convex bodies $K \subset L$. That means, we maximize the volume of K while minimizing that of L under the constraints. Since PK is some line segment, the maximal volume is reached by the infinite

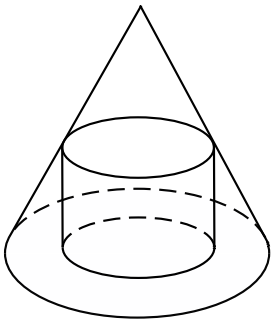


FIGURE 6.3: The outer convex body is the large cone while the inner body consists of the cylinder in addition to the small cone on top of the cylinder.

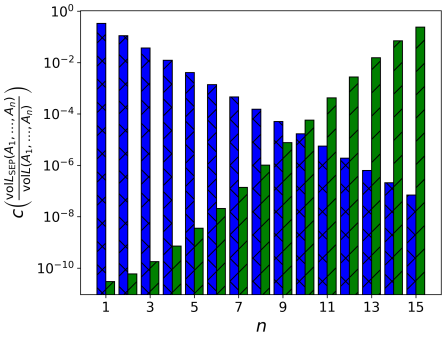


FIGURE 6.4: Shows the lower bounds from Theorem 6.4 (blue, cross pattern) and Conjecture 6.7 (green, line pattern). The minimum is reached for $n = 9$ and is given by $\frac{1}{3^9} \approx \frac{1}{2}10^{-5}$.

rectangle with this projection bounded by L as we have $K \subset L$. An arbitrary body L can be made smaller without changing K by considering two points that project onto the two end points of PL and their straight connections to the points that lie on the boundary of both K and L. Removing all points from L that are not within the set constrained by these straight lines, makes the volume of L smaller while not changing the volume of K, thus increasing the volume ratio. These points must have been inside L before because of convexity. Points that lie within the infinite rectangle with projection PK and the area constrained by the straight lines can be added to both L and K, increasing both volumes by the same constant, and hence, increasing the volume ratio; see Fig. 6.2a. Thus, we are left with optimizing the position of a quadrangle. Indeed, it is sufficient to separately optimize the triangles above and below the connecting line of the points that are projected onto the end points of PL. Using the intercept theorem, it is easy to show that the relative volume is independent of the vertical position (compare Fig. 6.2b) of the corner points. Straightforward optimization proofs that the optimal form is a triangle and PK which are symmetric w.r.t. the center of PL. Then, volume ratio is given by $\frac{\mathrm{vol}\,K}{\mathrm{vol}\,L} = 1 - (1 - \frac{\mathrm{vol}\,PK}{\mathrm{vol}\,PL})^2$. □

This proof motivates us to formulate the following conjecture.

Conjecture 6.7. For two convex sets $K \subset L \subset \mathbb{R}^n$ with fixed volume ratio $\frac{\mathrm{vol}\,K}{\mathrm{vol}\,L}$, it holds that

$$\min_{u \in \mathbb{R}^n} \frac{\mathrm{vol}\,P_u K}{\mathrm{vol}\,P_u L} \geq c\left(\frac{\mathrm{vol}\,K}{\mathrm{vol}\,L}\right), \tag{6.20}$$

where $0 \leq c\left(\frac{\mathrm{vol}\,K}{\mathrm{vol}\,L}\right) \leq 1$ is the solution to the equation

$$\frac{\mathrm{vol}\,K}{\mathrm{vol}\,L} = c\left[1 + (n-1)\left(1 - \sqrt[n-1]{c}\right)\right]. \tag{6.21}$$

This bound is tight and can be reached by two concentric $(n-1)$-dimensional balls PK and PL that serve as the base for a cylindrical object and a symmetrically positioned cone, respectively; compare Fig. 6.3.

If this conjecture is true, together with Theorem 6.4, we would have for two qubits the bounds shown in Fig. 6.4. We obtain a bound independent of n by taking the limit $n \to \infty$, since

$$\begin{aligned} \frac{\partial}{\partial n} c\left[1+(n-1)\left(1-\sqrt[n-1]{c}\right)\right] \geq 0, \\ \frac{\partial}{\partial c} c\left[1+(n-1)\left(1-\sqrt[n-1]{c}\right)\right] \geq 0. \end{aligned} \tag{6.22}$$

Then, we have that

$$\frac{\operatorname{vol} K}{\operatorname{vol} L} = c\,(1-\log c) \tag{6.23}$$

This configuration, however, is most likely suboptimal when $l > 1$. Thus, better bounds could be obtained by considering the general geometrical question.

6.5 Two qubits

We extensively study the case of two qubits which is the most basic system for entanglement detection using numerical range. Hence, it serves as a testbed for future investigations.

6.5.1 Single observable

In the simplest scenario, i.e., a single measurement of a two-qubit quantum system, Theorem 6.4 gives a bound $\mu_{2,2,1} \geq \frac{1}{3}$ for the minimal volume ratio. To obtain a better bound, we use so-called absolutely separable states, whose separability can be inferred from the eigenvalues of the density matrix [243–245].

Proposition 6.8. *It holds that $\mu_{2,2,1} \geq \sqrt{2}-1 \approx 0.41$. Moreover, this is the best bound achievable when only absolutely separable states are considered.*

Proof. Let us translate and scale the observable A such that its smallest and largest eigenvalue are 0 and 1. We denote by a and b the two other eigenvalues such that $0 \leq a \leq b \leq 1$ and we have that

$$A(a,b) = |\psi_1\rangle\,|\psi_1\rangle + b\,|\psi_b\rangle\,\langle\psi_b| + a\,|\psi_a\rangle\,\langle\psi_a| + 0\,|\psi_0\rangle\,\langle\psi_0|\,. \tag{6.24}$$

Now, let ρ be the quantum state

$$\rho = \lambda_1 |\psi_1\rangle |\psi_1\rangle + \lambda_2 |\psi_b\rangle \langle\psi_b| + \lambda_3 |\psi_a\rangle \langle\psi_a| + \lambda_4 |\psi_0\rangle \langle\psi_0| , \tag{6.25}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 = 1 - \lambda_1 - \lambda_2 - \lambda_3 \geq 0$. This state is absolutely separable if and only if $(\lambda_1 - \lambda_3)^2 \leq 4\lambda_2\lambda_4$ [245]. More specifically, we consider

$$\lambda_1(a,b) = \frac{1}{4}\left(\frac{8+\delta(a,b)}{\gamma(a,b)} - 1\right), \tag{6.26}$$

$$\lambda_2(a,b) = \lambda_3(a,b) = \frac{1}{4}\left(1 - \frac{\delta(a,b)}{\gamma(a,b)}\right), \tag{6.27}$$

$$\lambda_4(a,b) = 1 - \lambda_1(a,b) - \lambda_2(a,b) - \lambda_3(a,b) = \frac{1}{4}\left(3 - \frac{8-\delta(a,b)}{\gamma(a,b)}\right), \tag{6.28}$$

$$\tag{6.29}$$

where $\delta(a,b) = 1 - a - b$ and $\gamma(a,b) = \sqrt{8 + \delta^2(a,b)}$. From $a, b \leq 1$ and $\gamma \leq 3$, it follows that, indeed, $\lambda_1 \geq \lambda_2 = \lambda_3 \geq \lambda_4$. Furthermore, $\lambda_4 \geq 0$ is equivalent to $3\gamma \geq 8 - \delta$ which holds due to $8 - \delta \geq 0$ and $(3\gamma)^2 - (8-\delta)^2 = 8(\delta+1)^2 \geq 0$. Finally, we have that $(\lambda_1 - \lambda_3)^2 = 4\lambda_2\lambda_4$ and hence, $\rho(a,b)$ is an absolutely separable state for any a, b. Additionally, let us consider the state σ with

$$\sigma = \nu_4 |\psi_1\rangle |\psi_1\rangle + \nu_3 |\psi_b\rangle \langle\psi_b| + \nu_2 |\psi_a\rangle \langle\psi_a| + \nu_1 |\psi_0\rangle \langle\psi_0| , \tag{6.30}$$

where $\nu_j = \lambda_j(1-b, 1-a)$. Since ρ is an absolutely separable state for any a, b, so is σ. For the distance between their respective expectation values when measuring the observable A, we obtain

$$\operatorname{Tr} A\rho - \operatorname{Tr} A\sigma = -1 + \frac{2}{\gamma}\left(2 + \delta^2\right), \tag{6.31}$$

whose minimum $\sqrt{2} - 1$ is obtained at $\delta = 0$ and implies the bound $\mu_{2,2,1} \geq \sqrt{2} - 1$.

To prove that this is the best we can do with absolutely separable states, we consider $A(\frac{1}{2}, \frac{1}{2})$. It is well known that for a state ρ, there exists a von Neumann measurement with probability vector $p_j = \langle\psi_j| \rho |\psi_j\rangle$ if, and only if, $\boldsymbol{p} \prec \boldsymbol{\lambda}$, i.e., the probability vector is majorized by the eigenvalue vector [20]. Thus, the maximal expectation value with

$A(\frac{1}{2},\frac{1}{2})$ for an absolutely separable state is given by the following optimization

$$\begin{aligned}
\max_{\boldsymbol{p},\boldsymbol{\lambda}} \quad & p_1 + \frac{p_2+p_3}{2} \\
\text{s.t.} \quad & p_1 \leq \lambda_1, \quad p_1+p_2 \leq \lambda_1+\lambda_2, \quad p_1+p_2+p_3 \leq \lambda_1+\lambda_2+\lambda_3, \\
& p_1 \geq p_2 \geq p_3 \geq 1-p_1-p_2-p_3 \geq 0, \\
& \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 1-\lambda_1-\lambda_2-\lambda_3 \geq 0, \\
& (\lambda_1-\lambda_3)^2 \leq 4\lambda_2(1-\lambda_1-\lambda_2-\lambda_3).
\end{aligned} \tag{6.32}$$

Note that for absolutely separable states, the eigenvectors of A are irrelevant. Clearly, for a given feasible point, we can increase p_3 while decreasing p_2 such that $p_2 + p_3 = \text{const}$. since it leaves the objective value invariant and only relaxes the second constraint. Thus, there exists an optimal solution with $p_2 = p_3$. Now increasing p_1 while decreasing $p_2 = p_3$ such that $p_1 + p_2 = \text{const}$. leads to a relaxation of the constraints, and hence, we can require the optimum to satisfy $p_1 = \lambda_1$. In turn, we are left with maximizing p_2 depending on $\boldsymbol{\lambda}$ which gives $p_2 = p_3 = \frac{\lambda_2+\lambda_3}{2}$ and leads to the simplified optimization

$$\begin{aligned}
\max_{\boldsymbol{\lambda}} \quad & \lambda_1 + \frac{\lambda_2+\lambda_3}{2} \\
\text{s.t.} \quad & \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 1-\lambda_1-\lambda_2-\lambda_3 \geq 0, \\
& (\lambda_1-\lambda_3)^2 \leq 4\lambda_2(1-\lambda_1-\lambda_2-\lambda_3).
\end{aligned} \tag{6.33}$$

It is certainly optimal to maximize λ_1 until either $1-\lambda_1-\lambda_2-\lambda_3 = 0$ or $(\lambda_1-\lambda_3)^2 = 4\lambda_2(1-\lambda_1-\lambda_2-\lambda_3)$. In the first case, we can replace λ_3 by $\lambda_3 = 1-\lambda_1-\lambda_2$. Hence, we obtain the constraint $(2\lambda_1+\lambda_2-1)^2 \leq 0$ implying $\lambda_2 = 1-2\lambda_1$. Then, it must hold that $\lambda_1 \geq 1-2\lambda_1 \geq \lambda_1$ and hence, $\lambda_1 = \frac{1}{3}$ leading to an optimal value of $2/3$. In the other case, we use the method of Lagrange multipliers for the equality constraint and obtain $\lambda_1 = \frac{1}{2}$, $\lambda_2 = 0$, and $\lambda_3 = \frac{1}{2}$ which does not satisfy the inequalities. Hence, we look for the optimal solution at one of the boundaries $\lambda_1 = \lambda_2$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 1-\lambda_1-\lambda_2-\lambda_3$. For $\lambda_1 = \lambda_2$, we obtain via the method of Lagrange multipliers a single feasible point $\lambda_1 = \lambda_2 = (3+\sqrt{3})/12$, $\lambda_3 = -1+(3+\sqrt{3})/4$ with objective value $(1+\sqrt{3})/4 \approx 0.68$. In the case of $\lambda_2 = \lambda_3$, the same analysis gives the feasible point $\lambda_1 = (-1+2\sqrt{2})/4$, $\lambda_2 = \lambda_3 = 1/4$ with objective value $1/\sqrt{2} \approx 0.71$. The last case, $\lambda_3 = 1-\lambda_1-\lambda_2-\lambda_3$, leads to an objective value of $(1+\sqrt{3})/4 \approx 0.68$ at $\lambda_1 = (-3+5\sqrt{3})/12$, $\lambda_2 = (3-\sqrt{3})/4$, and $\lambda_3 = (3-\sqrt{3})/12$. Thus, the optimum is $1/\sqrt{2}$. For the minimal expectation value, we use the fact that

$$\min_{\rho} \operatorname{Tr} A\rho = 1 - \max_{\rho} \operatorname{Tr}(\mathbb{1}-A)\rho. \tag{6.34}$$

Since $\mathbb{1} - A(\frac{1}{2}, \frac{1}{2})$ is equivalent to $A(\frac{1}{2}, \frac{1}{2})$ when only considering eigenvalues, the minimal expectation value is given by $1 - 1/\sqrt{2}$ and the volume between maximal and minimal expectation value for absolutely separable states is $1/\sqrt{2} - (1 - 1/\sqrt{2}) = \sqrt{2} - 1$. □

However, extensive numerical investigation suggests that this bound is not tight and leads us to formulate the following conjecture.

Conjecture 6.9. For a single measurement on a two-qubit system, the minimal volume ratio is $\mu_{2,2,1} = \frac{1}{2}$.

This value is for example reached by $A = |\phi^+\rangle \langle\phi^+|$ with eigenvalues 0 and 1, being the projector onto the maximally entangled state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. From the given Schmidt decomposition, it is obvious that the maximal overlap of a product state with the maximally entangled state is $\frac{1}{2}$; also, the overlap of the product state $|01\rangle$ with $|\phi^+\rangle$ is 0. Thus, it follows that $\mu_{2,2,1} \leq \frac{1}{2}$.

To prove this conjecture, we try to find a proof for the following statement.

Conjecture 6.10. The minimal volume ratio $\mu_{2,d,1}$, i.e., a single measurement of a bipartite locally d-dimensional system, is reached by a projector.

This result would immediately imply Conjecture 6.9 as well as $\mu_{2,d,1} = \frac{1}{3}$, since for any projector P, either the subspace P or $\mathbb{1} - P$ has at least dimension $\frac{d^2}{2}$ and hence, contains a product vector if $d = 2, 3$ [246]. Then, from the Schmidt decomposition of the maximally entangled state, it follows that there exists a product state having an overlap with the other subspace of at least $\frac{1}{d}$. Furthermore, a projector onto the maximally entangled state reaches a volume ratio of $\frac{1}{d}$. In higher dimensions, there exist partitions of the Hilbert space into two subspaces which neither contain a product state [246]. Indeed, almost all partitions do have this property [247].

One idea to prove Conjecture 6.10, that seems to fail, is to use perturbation theory. Nevertheless, we present it here as it might inspire further investigations. To make things easy, we (initially) only show evidence that either the ground state or the most excited state of the observable A must be degenerate (which is not true!). Let us assume that we found an optimal observable

$$H = \lambda_{\min} |\psi_{\min}\rangle \langle\psi_{\min}| + \sum_{j=2}^{d^2-1} \lambda_j |\psi_j\rangle \langle\psi_j| + \lambda_{\max} |\psi_{\max}\rangle \langle\psi_{\max}| \tag{6.35}$$

where $\lambda_{\max} = 1$ and $\lambda_{\min} = 0$ as well as $0 < \lambda_j < 1$. Thus, we consider optimal solutions with nondegenerate ground and most excited states, which is supposed to

lead to a contradiction. Also, let us assume that the lower and upper boundary of the separable numerical range are given by α and β, and realized by product states $|\alpha\rangle$ and $|\beta\rangle$, respectively. Then, we introduce a small perturbation such that $\tilde{H} = H + \epsilon |\chi\rangle \langle\chi|$. For the extremal eigenvalues of H, nondegenerate pertubation theory applies and we obtain up to first order in ϵ,

$$\tilde{\lambda}_{\min} = \epsilon |\langle\psi_{\min}|\chi\rangle|^2, \tag{6.36}$$

$$\tilde{\lambda}_{\max} = 1 + \epsilon |\langle\psi_{\max}|\chi\rangle|^2. \tag{6.37}$$

For the perturbed boundaries of the separable numerical range, the analysis is more involved. We only consider $|\alpha\rangle = |\alpha_1\alpha_2\rangle$ because it works analogously for $|\beta\rangle = |\beta_1\beta_2\rangle$. Up to first order in ϵ, we have that

$$\begin{aligned} |\tilde{\alpha}\rangle = |\tilde{\alpha}_1\rangle \otimes |\tilde{\alpha}_2\rangle &= \left(|\alpha_1\rangle + \epsilon |\alpha_1^{(1)}\rangle\right) \otimes \left(|\alpha_2\rangle + \epsilon |\alpha_2^{(1)}\rangle\right) \\ &= |\alpha_1\alpha_2\rangle + \epsilon \left(|\alpha_1\rangle |\alpha_2^{(1)}\rangle + |\alpha_1^{(1)}\rangle |\alpha_2\rangle\right). \end{aligned} \tag{6.38}$$

Furthermore, for the perturbation of α, we obtain

$$\tilde{\alpha} = \langle\tilde{\alpha}| \tilde{H} |\tilde{\alpha}\rangle = \alpha + \epsilon |\langle\alpha|\chi\rangle|^2 + \epsilon \left(\langle\alpha_1\alpha_2| H |\alpha_1^{(1)}\alpha_2\rangle + \langle\alpha_1\alpha_2| H |\alpha_1\alpha_2^{(1)}\rangle + h.c.\right). \tag{6.39}$$

Since $|\alpha_1\rangle$ is an eigenvector with minimal eigenvalue α of $\langle\alpha_2| H |\alpha_2\rangle$ — otherwise there would be, by definition, a smaller α — it holds that $\langle\alpha_1\alpha_2| H |\alpha_1^{(1)}\alpha_2\rangle = \alpha \langle\alpha_1|\alpha_1^{(1)}\rangle$ and similar for the other terms. Moreover, normalization requires due to Eq. (6.38) that $|\alpha_1\rangle |\alpha_2^{(1)}\rangle + |\alpha_1^{(1)}\rangle |\alpha_2\rangle$ is orthogonal to $|\alpha_1\alpha_2\rangle$. Thus, $\langle\alpha_1|\alpha_1^{(1)}\rangle + \langle\alpha_2|\alpha_2^{(1)}\rangle = 0$ implying

$$\tilde{\alpha} = \alpha + \epsilon |\langle\alpha|\chi\rangle|^2. \tag{6.40}$$

However, in contrast to $|\psi_{\min}\rangle$ for nondegenerate H, the state $|\alpha\rangle$ might not be unique and hence, in Eq. (6.40), $|\alpha\rangle$ depends on $|\chi\rangle$. More precisely,

$$\tilde{\alpha} = \alpha + \epsilon \min_{|\alpha\rangle} |\langle\alpha|\chi\rangle|^2, \tag{6.41}$$

$$\tilde{\beta} = \beta + \epsilon \max_{|\beta\rangle} |\langle\beta|\chi\rangle|^2. \tag{6.42}$$

In general, we can approximate the optimizations by choosing any $|\alpha\rangle$ and $|\beta\rangle$, respectively,

$$\tilde{\alpha} \leq \alpha + \epsilon |\langle\alpha|\chi\rangle|^2, \tag{6.43}$$

$$\tilde{\beta} \geq \beta + \epsilon |\langle\beta|\chi\rangle|^2. \tag{6.44}$$

Now, let $f : H \mapsto \frac{\beta(H)-\alpha(H)}{\lambda_{\max}(H)-\lambda_{\min}(H)}$ be the function that gives the volume ratio for a single observable. Again, up to first order in ϵ, it holds that

$$\begin{aligned} f(\tilde{H}) - f(H) &= \left[1 - \epsilon\left(|\langle\psi_{\max}|\chi\rangle|^2 - |\langle\psi_{\min}|\chi\rangle|^2\right)\right](\tilde{\beta}-\tilde{\alpha}) - (\beta-\alpha) \\ &\geq \epsilon\left[-(\beta-\alpha)\left(|\langle\psi_{\max}|\chi\rangle|^2 - |\langle\psi_{\min}|\chi\rangle|^2\right) + \left(|\langle\beta|\chi\rangle|^2 - |\langle\alpha|\chi\rangle|^2\right)\right] \end{aligned} \tag{6.45}$$

If the function f were differentiable, then for an optimal observable H, i.e., at a minimum of f, the first derivative and hence, the difference $f(\tilde{H}) - f(H)$ up to first order in ϵ would be 0. Since this would hold for positive and for negative ϵ, the above inequality would imply that, for all $|\chi\rangle$,

$$(\beta-\alpha)\left(|\langle\psi_{\max}|\chi\rangle|^2 - |\langle\psi_{\min}|\chi\rangle|^2\right) = \left(|\langle\beta|\chi\rangle|^2 - |\langle\alpha|\chi\rangle|^2\right), \tag{6.46}$$

equivalent to

$$\langle\chi|(\beta-\alpha)\left(|\psi_{\max}\rangle\langle\psi_{\max}| - |\psi_{\min}\rangle\langle\psi_{\min}|\right)|\chi\rangle = \langle\chi|\left(|\beta\rangle\langle\beta| - |\alpha\rangle\langle\alpha|\right)|\chi\rangle, \tag{6.47}$$

or as an operator equation

$$(\beta-\alpha)|\psi_{\max}\rangle\langle\psi_{\max}| - |\psi_{\min}\rangle\langle\psi_{\min}| = |\beta\rangle\langle\beta| - |\alpha\rangle\langle\alpha|. \tag{6.48}$$

Finally, consider the projector $P = |\psi_{\max}\rangle\langle\psi_{\max}| + |\psi_{\min}\rangle\langle\psi_{\min}|$ and apply it to the equation from both sides, which clearly leaves the left-hand side of the equation invariant and hence, the same must be true for the right-hand side, i.e.,

$$P|\beta\rangle\langle\beta|P - P|\alpha\rangle\langle\alpha|P = |\beta\rangle\langle\beta| - |\alpha\rangle\langle\alpha|. \tag{6.49}$$

Let $|\alpha\rangle = |\alpha^{\parallel}\rangle + |\alpha^{\perp}\rangle$ such that $P|\alpha^{\parallel}\rangle = |\alpha^{\parallel}\rangle$ and $P|\alpha^{\perp}\rangle = 0$, similarly for $|\beta\rangle$. Inserting and applying P from the left and $\mathbb{1} - P$ from the right gives

$$|\alpha^{\parallel}\rangle\langle\alpha^{\perp}| = |\beta^{\parallel}\rangle\langle\beta^{\perp}|, \tag{6.50}$$

which by comparing the range and kernel of the matrices yields $|\alpha^{\parallel}\rangle = |\beta^{\parallel}\rangle$ as well as $\langle\alpha^{\perp}| = \langle\beta^{\perp}|$, and hence, $|\alpha\rangle = |\beta\rangle$. Thus, $\alpha = \beta$ which is impossible due to Theorem 6.4.

For the contradiction, two assumptions were crucial. First, we started with an optimal H with nondegenerate ground and most excited state and, second, we assumed that the function f yielding the volume ratio is differentiable. This argument can be extended using degenerate perturbation theory to, in principle, prove the conjecture, however, it is flawed in the very same way.

To see that f is indeed not differentiable at all points, consider the following example $H = |\phi^+\rangle\langle\phi^+| + \frac{1}{2}|\phi^-\rangle\langle\phi^-| + \frac{1}{2}|\psi^-\rangle\langle\psi^-|$, or scaled and translated,

$$H' = |\phi^+\rangle\langle\phi^+| - |\psi^+\rangle\langle\psi^+|, \tag{6.51}$$

where the $|\phi^\pm\rangle$, $|\psi^\pm\rangle$ are the Bell states. From this it is clear that for H, $\alpha = \frac{1}{4}$ and $\beta = \frac{3}{4}$ with, e.g., $|\alpha\rangle = |01\rangle$ and $|\beta\rangle = |00\rangle$. Now, the perturbation $\tilde{H} = H + \epsilon|\phi^-\rangle\langle\phi^-|$ behaves differently depending on whether ϵ is positive or negative. Note that the maximal and minimal eigenvalue of H do not change with this perturbation. Let first $\epsilon > 0$, then $\tilde{\alpha} = \alpha$ since the perturbation is positive semidefinite and $\langle 01|\phi^-\rangle = 0$. Also, $\tilde{\beta} = \beta + \frac{\epsilon}{2}$ since there is no product state with larger overlap than $|\langle 00|\phi^-\rangle|^2 = \frac{1}{2}$. Thus,

$$\lim_{\epsilon^+\to 0} \frac{f(\tilde{H}) - f(H)}{\epsilon} = \frac{1}{2}, \tag{6.52}$$

Second, let $\epsilon < 0$ and consider $|\alpha\rangle = |i_+\rangle|i_+\rangle$ and $|\beta\rangle = |i_+\rangle|i_-\rangle$ where $|i_\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$ with the properties

$$|\langle\beta|\phi^+\rangle|^2 = \frac{1}{2}, \quad |\langle\beta|\psi^+\rangle|^2 = 0, \quad |\langle\beta|\phi^-\rangle|^2 = 0, \tag{6.53}$$

$$|\langle\alpha|\psi^+\rangle|^2 = \frac{1}{2}, \quad |\langle\alpha|\phi^+\rangle|^2 = 0, \quad |\langle\alpha|\phi^-\rangle|^2 = \frac{1}{2}. \tag{6.54}$$

This shows that $\tilde{\alpha} = \alpha + \frac{\epsilon}{2} < \alpha$ as well as $\tilde{\beta} = \beta$ and hence,

$$\lim_{\epsilon^-\to 0} \frac{f(\tilde{H}) - f(H)}{\epsilon} = \frac{-\epsilon}{2\epsilon} = -\frac{1}{2}. \tag{6.55}$$

Thus, the limits do not coincide and hence, f is not differentiable at H.

Note that, with $|\alpha\rangle = |01\rangle$ and $|\beta\rangle = |00\rangle$, Eq. (6.45) implies that $f(\tilde{H}) - f(H) \geq \frac{\epsilon}{2}$ which is indeed satisfied since $f(\tilde{H}) - f(H) = \frac{|\epsilon|}{2}$.

6.5.2 Multiple observables

Concerning multiple observables, the bound in Theorem 6.4 decreases exponentially with the number of measurements. From Proposition 6.5, it is clear that, in contrast to the bounds found, the actual minimal volume ratio $\mu_{2,2,k}$ does not decrease exponentially with k. Indeed, we find an instance of three measurements proving that $\mu_{2,2,3} \leq 1/5 < 8/33 \approx \mu_{2,2,15}$, where the approximation is due to Ref. [238–241], and hence, $\mu_{2,2,k}$ is a nonmonotonic function of k. This example is given by the observables $A_1 = 0 \oplus X \oplus 0$, $A_2 = 0 \oplus Y \oplus 0$, and $A_3 = 0 \oplus Z \oplus 0$, where X, Y, and Z are the

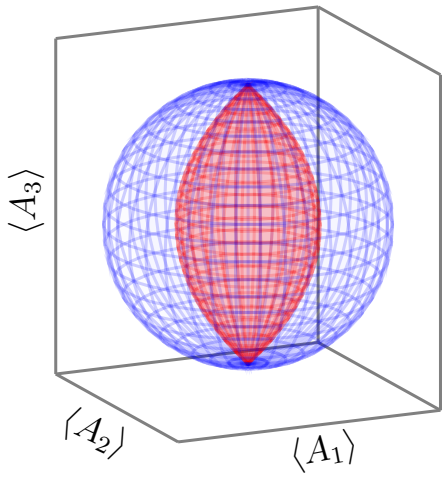


FIGURE 6.5: This figure shows the (separable) numerical range for a two-qubit quantum system and observables $A_1 = 0 \oplus X \oplus 0$, $A_2 = 0 \oplus Y \oplus 0$, and $A_3 = 0 \oplus Z \oplus 0$. The relative volume is given by $\frac{1}{5}$ and hence, $\mu_{2,2,3} \leq \frac{1}{5}$. Also, when only the measurements A_1 and A_2 are considered, the relative volume is given by $\frac{1}{4}$ and hence, $\mu_{2,2,2} \leq \frac{1}{4}$.

Pauli matrices and the symbol $\oplus$ denotes the direct sum of matrices where the number 0 is understood as a 1-by-1 matrix. The resulting (separable) numerical range is visualized in Fig. 6.5. From the structure of the observables, it is clear that the joint numerical range is a Bloch ball since only the subspace of nonzero eigenvalues of the A_j is relevant. For separable states, the local unitaries $U_1 \otimes U_2$, where

$$U_1 = |0\rangle\langle 0| + e^{i\varphi/2}|1\rangle\langle 1|, \tag{6.56}$$

$$U_2 = |0\rangle\langle 0| + e^{-i\varphi/2}|1\rangle\langle 1|, \tag{6.57}$$

leave A_3 invariant while continuously transforming A_1 and A_2 as $A_1 \to \cos\varphi A_1 + \sin\varphi A_2$ and $A_2 \to -\sin\varphi A_1 + \cos\varphi A_2$, respectively. Thus, the separable numerical range is symmetric w.r.t. rotations around the axis of the third measurement, but it is not symmetric w.r.t. other rotations. As we are considering the three Pauli matrices, this asymmetry might seem counter-intuitive at first glance, however, while the eigenvectors with corresponding nonzero eigenvalues are product states for A_3, this is not true for A_1 and A_2, which explains the difference in symmetries for separable and all quantum states. Also, restricting just to measurements A_1 and A_2 gives two concentric circles for the (separable) numerical range with a volume ratio of $\frac{1}{4}$ and hence, $\mu_{2,2,2} \leq \frac{1}{4}$.

More precisely, the joint numerical range is given by a Bloch ball on the subspace spanned by $|01\rangle$ and $|10\rangle$, i.e., a ball of radius 1. Thus, the two- and three-dimensional volumes are given by π and $\frac{4}{3}\pi$, respectively. Because for local unitaries U, $\operatorname{Tr}\rho_{\text{sep}}UA_iU^\dagger = \operatorname{Tr}U^\dagger\rho_{\text{sep}}UA_i = \operatorname{Tr}\sigma_{\text{sep}}A_i$ for any separable state ρ_{sep} and separable $\sigma_{\text{sep}} = U^\dagger\rho_{\text{sep}}U$,

the transformations with the local unitaries $U_1 \otimes U_2$ impliy the rotational symmetry. Thus, it is sufficient to solve the parametric optimization

$$\begin{aligned} \max_{\rho \in \mathrm{SEP}} \quad & \operatorname{Tr} \rho A_1 \\ \text{s.t.} \quad & \operatorname{Tr} \rho A_2 = 0, \\ & \operatorname{Tr} \rho A_3 = c, \end{aligned} \tag{6.58}$$

where $-1 \leq c \leq 1$. As the separable numerical range is the convex hull of the pure-product numerical range, coming from pure product states, we consider general product states $|\alpha\rangle |\beta\rangle$ with

$$|\alpha\rangle = \cos\frac{\alpha}{2} |0\rangle + e^{i\phi} \sin\frac{\alpha}{2} |1\rangle , \tag{6.59}$$

$$|\beta\rangle = \cos\frac{\beta}{2} |0\rangle + e^{i\psi} \sin\frac{\beta}{2} |1\rangle . \tag{6.60}$$

We have that $\langle \alpha\beta | A_1 | \alpha\beta \rangle = \frac{1}{2}\cos(\phi - \psi)\sin\alpha \sin\beta$ and $\langle \alpha\beta | A_3 | \alpha\beta \rangle = \frac{1}{2}(\cos\alpha - \cos\beta)$. Hence, to maximize $\langle A_1 \rangle$, certainly $\cos(\phi - \psi) = 1$ since we can choose the signs of $\sin\alpha$ and $\sin\beta$ independently from those of $\cos\alpha$ and $\cos\beta$. The choice $\phi = \psi = 0$ not only gives $\cos(\phi - \psi) = 1$, but also makes sure $|\alpha\beta\rangle$ satisfies $\langle A_2 \rangle = 0$ as A_2 is a skew-symmetric matrix and, in this case, $|\alpha\beta\rangle$ is a real-valued vector in the computational basis. Actually, it is clear from the rotational symmetry that minimal and maximal $\langle A_1 \rangle$ are reached for $\langle A_2 \rangle = 0$. Thus, we are left with optimizing

$$\begin{aligned} \max_{\alpha,\beta} \quad & \sin\alpha \sin\beta \\ \text{s.t.} \quad & \cos\alpha - \cos\beta = c'. \end{aligned} \tag{6.61}$$

To solve this, we write $x = \cos\alpha$, $y = -\cos\alpha$, and $\sin\alpha \sin\beta = \sqrt{(1-x^2)(1-y^2)}$ since we can always choose the positive solutions for the sines for given cosines. Because the square root is a monotonic function, it is equivalent to maximize $(1-x^2)(1-y^2) = 1 - c'^2 + xy(xy+2)$ and, as $-1 \leq xy$, also equivalent to maximize xy, leaving us with

$$\begin{aligned} \max_{x,y} \quad & xy \\ \text{s.t.} \quad & x + y = c'. \end{aligned} \tag{6.62}$$

By changing both the signs of x and y, c' can always be chosen nonnegative, and we have to find the rectangle with largest volume for given circumference $\frac{c'}{2}$, which is known to be a square. Hence, $x = y$, or equivalently $\alpha = \beta$, provides the maximum with $\langle A_1 \rangle = \frac{1}{2}(1 - c^2)$. As this line provides the boundary of a convex set, the pure-product and separable numerical range coincide; see Fig. 6.5 for a visualization. The

corresponding volume is given by

$$\operatorname{vol} L_{\mathrm{SEP}}(A_1, A_2, A_3) = \pi \int_{-1}^{1} dc \left[\frac{1}{2}(1-c^2)\right]^2 = \frac{4}{15}\pi, \tag{6.63}$$

and the relative volume is $\frac{1}{5}$. In the two-dimensional case restricted to observables A_1 and A_2, maximal $\langle\alpha\beta| A_1 |\alpha\beta\rangle$ independent from c gives us the relevant volume as

$$\operatorname{vol} L_{\mathrm{SEP}}(A_1, A_2) = \left(\frac{1}{2}\right)^2 \pi = \frac{\pi}{4}, \tag{6.64}$$

which leads to a relative volume of $\frac{1}{4}$.

6.5.3 Product observables

In practice, it is much harder to implement highly entangled measurements compared to local, i.e. product, observables. That is why, in the following, we focus on such simpler observables. Product observables, i.e. $A = B_1 \otimes B_2$, are easy to measure in spatially separated laboratories or on a composite quantum system consisting of separate particles, and therefore a physically well motivated subset of all possible observables. For a single product observable, separable and general numerical range are obviously identical as the eigenvectors are product states. Thus, let us consider the simplest non-trivial case which is a two-qubit system and two product observables A_1, A_2. Moreover, we restrict ourselves to locally traceless observables, i.e., after applying suitable local unitaries, which apparently do not affect the relative volume, we have

$$A_1 = X \otimes X, \tag{6.65}$$

$$A_2 = (\cos\theta_A X + \sin\theta_A Z) \otimes (\cos\theta_B X + \sin\theta_B Z)\,. \tag{6.66}$$

For local implementations of the measurements, the restriction to locally traceless observables corresponds to a constant translation of the measurement results which is trivial from an experimenter's point of view. Note that instead of ρ, we can consider $\rho' = \frac{1}{2}(\rho + \rho^T)$ which gets rid of Pauli terms with a single Pauli-Y operator.

We denote the minimal volume ratio for (locally traceless) product measurements as $\mu^{\otimes}_{n,\boldsymbol{d},k}$ ($\mu^{\otimes,\mathrm{LT}}_{n,\boldsymbol{d},k}$) where n is the number of particles, $\boldsymbol{d}$ the vector of local dimensions, and k the number of observables. We conjecture that $\mu^{\otimes}_{2,2,2} = \mu^{\otimes,\mathrm{LT}}_{2,2,2} = \frac{1}{2}$. A volume ratio of $\frac{1}{2}$ is for instance realized by the measurements $A_1 = X \otimes X$ and $A_2 = Z \otimes Z$, illustrated in Fig. 6.6.

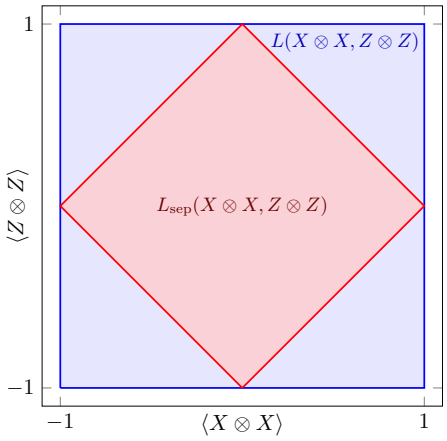


FIGURE 6.6: This figure shows the (separable) numerical range for observables $X \otimes X$ and $Z \otimes Z$.

Indeed, we can calculate the volume ratio for given locally traceless observables explicitly.

Theorem 6.11. *For two-qubit observables A_1 and A_2 given by $A_1 = B_1 \otimes B_2$, $A_2 = C_1 \otimes C_2$ with $\operatorname{Tr} B_j = \operatorname{Tr} C_j = 0$, where the ϑ_j are the angles between B_j and C_j in the Bloch sphere and $\vartheta_\pm = \vartheta_1 \pm \vartheta_2$, the volume ratio is*

$$\begin{aligned}\frac{\operatorname{vol} L_{\text{SEP}}}{\operatorname{vol} L} =& \frac{\pi}{8}\left[(|\sin\vartheta_-| + |\sin\vartheta_+|) - \frac{\tilde{F}(\vartheta_-,\vartheta_+)}{\tilde{T}(\vartheta_-,\vartheta_+)}\right] / \\ & \left[\cos\vartheta_- - \cos\vartheta_+ + G_-(\vartheta_-) + G_+(\vartheta_+)\right],\end{aligned} \tag{6.67}$$

where

$$G_-(\vartheta_-) = |\frac{\vartheta_-}{2}\sin\vartheta_-|, \tag{6.68}$$

$$G_+(\vartheta_+) = |\left(\frac{\vartheta_+}{2} - \frac{\pi}{2}\right)\sin\vartheta_+|, \tag{6.69}$$

$$\tilde{F}(\vartheta_-,\vartheta_+) = F(\vartheta_-,\vartheta_+) - F(\vartheta_+,\vartheta_-), \tag{6.70}$$

$$\tilde{T}(\vartheta_-,\vartheta_+) = T(\vartheta_-,\vartheta_+) - T(\vartheta_+,\vartheta_-), \tag{6.71}$$

and the functions F and T are given by

$$F(x,y) = \left|\sin\frac{x}{2}\cos\frac{y}{2}\right|\left[K\left(1 - T^2(x,y)\right) - E\left(1 - T^2(x,y)\right)\right], \tag{6.72}$$

$$T(x,y) = \left|\frac{\tan\frac{x}{2}}{\tan\frac{y}{2}}\right|, \tag{6.73}$$

where $K(\cdot)$ and $E(\cdot)$ are the elliptic integrals of the first and second kind, respectively.

Since rescaling as well as local unitary transformations preserve the volume ratio, it is sufficient to consider observables $A_1 = X \otimes X$ and $A_2 = (\cos\vartheta_A X + \sin\vartheta_A Z) \otimes$

$(\cos\vartheta_B X + \sin\vartheta_B Z)$ with $0 \leq \vartheta_A, \vartheta_B \leq \pi$. To prove Theorem 6.11, we compute the separable as well as the general numerical range for these observables explicitly via the following Lemmata.

Lemma 6.12. *For two-qubit observables $A_1 = X \otimes X$ and $A_2 = (\cos\vartheta_A X + \sin\vartheta_A Z) \otimes (\cos\vartheta_B X + \sin\vartheta_B Z)$ where $0 \leq \vartheta_A, \vartheta_B \leq \pi$, the volume of the separable numerical range is*

$$\operatorname{vol} L_{\mathrm{SEP}} = \frac{\pi}{4}\left(|\sin\vartheta_-| + |\sin\vartheta_+|\right) + \frac{2\left[F(\vartheta_-,\vartheta_+) - F(\vartheta_+,\vartheta_-)\right]}{T(\vartheta_+,\vartheta_-) - T(\vartheta_-,\vartheta_+)}, \tag{6.74}$$

where the functions F and T are given by

$$F(x,y) = \left|\sin\frac{x}{2}\cos\frac{y}{2}\right|\left[K\left(1-T^2(x,y)\right) - E\left(1-T^2(x,y)\right)\right], \tag{6.75}$$

$$T(x,y) = \left|\frac{\tan\frac{x}{2}}{\tan\frac{y}{2}}\right|, \tag{6.76}$$

$\vartheta_\pm = \vartheta_A \pm \vartheta_B$, and $K(\cdot)$ and $E(\cdot)$ are the elliptic integrals of the first and second kind, respectively.

Proof. For a general product state

$$|\psi_\otimes\rangle = \left(\cos\theta_1|0\rangle + e^{i\phi_1}\sin\theta_1|1\rangle\right) \otimes \left(\cos\theta_2|0\rangle + e^{i\phi_2}\sin\theta_2|1\rangle\right), \tag{6.77}$$

we have that

$$\langle A_1\rangle = \cos\theta_1\cos\theta_2, \tag{6.78}$$

$$\langle A_2\rangle = (\cos\theta_1\cos\vartheta_A + \cos\phi_1\sin\theta_1\sin\vartheta_A)(\cos\theta_2\cos\vartheta_B + \cos\phi_2\sin\theta_2\sin\vartheta_B). \tag{6.79}$$

Clearly, for fixed $\langle A_1\rangle$, the extremal points of $\langle A_2\rangle$ are reached for $\cos\phi_1 = \pm 1$ and $\cos\phi_2 = \pm 1$. Thus, allowing $\theta_1, \theta_2 \in [0, 2\pi)$, we can fix $\cos\theta_1 = \cos\theta_2 = 1$ to calculate the boundary of the separable numerical range. Hence, we obtain

$$\langle\psi_\otimes| A_1 |\psi_\otimes\rangle = \frac{1}{2}\left(\cos\theta_- + \cos\theta_+\right), \tag{6.80}$$

$$\langle\psi_\otimes| A_2 |\psi_\otimes\rangle = \frac{1}{2}\left[\cos(\theta_- - \vartheta_-) + \cos(\theta_+ - \vartheta_+)\right], \tag{6.81}$$

where $\theta_\pm = \theta_1 \pm \theta_2$ and $\vartheta_\pm = \vartheta_A \pm \vartheta_B$. To simplify the calculation, it is advantageous to perform a rotation by $-\frac{\pi}{4}$ such that $\tilde{A}_1 = \frac{1}{\sqrt{2}}(A_1 + A_2)$ and $\tilde{A}_2 = \frac{1}{\sqrt{2}}(-A_1 + A_2)$.

Again, this does not change the volume of the separable numerical range. Then,

$$\langle\psi_\otimes|\,\tilde{A}_1\,|\psi_\otimes\rangle = \frac{1}{\sqrt{2}}\left(\cos\frac{\vartheta_-}{2}\cos z_- + \cos\frac{\vartheta_+}{2}\cos z_+\right), \tag{6.82}$$

$$\langle\psi_\otimes|\,\tilde{A}_2\,|\psi_\otimes\rangle = \frac{1}{\sqrt{2}}\left(\sin\frac{\vartheta_-}{2}\sin z_- + \sin\frac{\vartheta_+}{2}\sin z_+\right), \tag{6.83}$$

where $z_\pm = \frac{\vartheta_\pm}{2} - \theta_\pm \in [0, 2\pi)$. Thus, the enclosed area is the Minkowski sum of two ellipses, and hence, it is convex. Indeed, in this case, the product numerical range and the separable numerical range coincide as varying $z_\pm$ continuously traces out the entire boundary. Therefore, to calculate the volume, we fix $\langle\psi_\otimes|\,\tilde{A}_2\,|\psi_\otimes\rangle$ while maximizing and minimizing $\langle\psi_\otimes|\,\tilde{A}_1\,|\psi_\otimes\rangle$. Because we can change the signs of $\cos z_\pm$ and $\sin z_\pm$ independently, the maximum and minimum actually coincide apart from the sign. Thus, it is sufficient to solve the optimization problem

$$\begin{aligned} \max_{z_\pm\in[0,\frac{\pi}{2}]} & \left|\cos\frac{\vartheta_-}{2}\right|\cos z_- + \left|\cos\frac{\vartheta_+}{2}\right|\cos z_+ \\ \text{s.t.} & \left|\sin\frac{\vartheta_-}{2}\right|\sin z_- + \left|\sin\frac{\vartheta_+}{2}\right|\sin z_+ = |c|, \end{aligned} \tag{6.84}$$

where $|c| \in [0, \left|\sin\frac{\vartheta_-}{2}\right| + \left|\sin\frac{\vartheta_+}{2}\right|]$. We can restrict ourselves to $z_\pm \in [0, \frac{\pi}{2}]$ since it is obviously optimal to have $\cos z_\pm \geq 0$, and if one of the $\sin z_\pm$ were negative, making it positive would reduce the other and hence, lead to a larger objective value. Since

$$\frac{d\left(\left|\cos\frac{\vartheta_\pm}{2}\right|\cos z_\pm\right)}{d\left(\left|\sin\frac{\vartheta_\pm}{2}\right|\sin z_\pm\right)} = -\frac{\tan z_\pm}{\left|\tan\frac{\vartheta_\pm}{2}\right|}, \tag{6.85}$$

a maximum is reached when $\frac{\tan z_-}{\left|\tan\frac{\vartheta_-}{2}\right|} = \frac{\tan z_+}{\left|\tan\frac{\vartheta_+}{2}\right|}$. This equation always describes a feasible point of the optimization as $\tan z_\pm$ approaches infinity when $\sin z_\pm$ goes to 1, and hence, continuously changing $\sin z_-$ from its minimal allowed value for a given $|c|$ to 1 leads to a continuous change of $\sin z_+$ from 1 to its minimal allowed value via the constraint, and somewhere along the way, the condition $\frac{\tan z_-}{\left|\tan\frac{\vartheta_-}{2}\right|} = \frac{\tan z_+}{\left|\tan\frac{\vartheta_+}{2}\right|}$ is satisfied. Thus, we obtain

$$\operatorname{vol} L_{\mathrm{SEP}}(A_1, A_2) = 2\int_0^{\left|\sin\frac{\vartheta_-}{2}\right|+\left|\sin\frac{\vartheta_+}{2}\right|} \operatorname{opt}(c)\,dc, \tag{6.86}$$

where $\mathrm{opt}(c)$ is the result of the optimization in Eq. 6.84. Since we have that $z_\pm \in [0, \frac{\pi}{2}]$,

$$\begin{aligned} c &= \left|\sin\frac{\vartheta_-}{2}\right| \sin z_- + \left|\sin\frac{\vartheta_+}{2}\right| \sin z_+ \\ &= \left|\sin\frac{\vartheta_-}{2}\right| \frac{\tan z_-}{\sqrt{1+\tan^2 z_-}} + \left|\sin\frac{\vartheta_+}{2}\right| T(\vartheta_+,\vartheta_-) \frac{\tan z_-}{\sqrt{1+\left(T(\vartheta_+,\vartheta_-)\tan z_-\right)^2}}, \end{aligned} \tag{6.87}$$

as well as

$$\begin{aligned} \mathrm{opt}(c) &= \left|\cos\frac{\vartheta_-}{2}\right| \cos z_- + \left|\cos\frac{\vartheta_+}{2}\right| \cos z_+ \\ &= \left|\cos\frac{\vartheta_-}{2}\right| \frac{1}{\sqrt{1+\tan^2 z_-}} + \left|\cos\frac{\vartheta_+}{2}\right| \frac{1}{\sqrt{1+\left(T(\vartheta_+,\vartheta_-)\tan z_-\right)^2}}, \end{aligned} \tag{6.88}$$

changing the integration variable from c to $u = \tan z_-$ leads to

$$\begin{aligned} \mathrm{vol}\, L_{\mathrm{SEP}} = \int_0^\infty du \Bigg[&\frac{|\sin\vartheta_-|}{(1+u^2)^2} + \frac{|\sin\vartheta_+|\, T(\vartheta_+,\vartheta_-)}{\left(1+T^2(\vartheta_+,\vartheta_-)u^2\right)^2} \\ &+ \frac{2\left|\sin\frac{\vartheta_-}{2}\cos\frac{\vartheta_+}{2}\right|}{(1+u^2)^{\frac{3}{2}}\left(1+T^2(\vartheta_+,\vartheta_-)u^2\right)^{\frac{1}{2}}} + \frac{2\left|\cos\frac{\vartheta_-}{2}\sin\frac{\vartheta_+}{2}\right| T(\vartheta_+,\vartheta_-)}{\left(1+T^2(\vartheta_+,\vartheta_-)u^2\right)^{\frac{3}{2}}(1+u^2)^{\frac{1}{2}}} \Bigg] \\ = \frac{\pi}{4}\left(|\sin\vartheta_-| + |\sin\vartheta_+|\right) &+ \frac{2\left[F(\vartheta_-,\vartheta_+) - F(\vartheta_+,\vartheta_-)\right]}{T(\vartheta_+,\vartheta_-) - T(\vartheta_-,\vartheta_+)}, \end{aligned} \tag{6.89}$$

□

To compute the volume of the general numerical range, we use a known procedure for the computation of the joint numerical range of two Hermitian matrices as described in Section 2.6.

Lemma 6.13. *For two-qubit observables* $A_1 = X \otimes X$ *and* $A_2 = (\cos\vartheta_A X + \sin\vartheta_A Z) \otimes (\cos\vartheta_B X + \sin\vartheta_B Z)$ *where* $0 \le \vartheta_A, \vartheta_B \le \pi$*, the volume of the joint numerical range is*

$$\mathrm{vol}\, L(A_1, A_2) = 2\left[\cos\vartheta_- - \cos\vartheta_+ + |\frac{\vartheta_-}{2}\sin\vartheta_-| + |\left(\frac{\vartheta_+}{2} - \frac{\pi}{2}\right)\sin\vartheta_+|\right] \tag{6.90}$$

where $\vartheta_\pm = \vartheta_A \pm \vartheta_B$.

Proof. We calculate the numerical range explicitly using its generating line $C(A_1, A_2)$ defined by the dual (line) equation

$$\det(uA_1 + vA_2 + w\mathbb{1}) = 0, \tag{6.91}$$

where $ux + vy + w = 0$ is the equation of a supporting line to $L(A_1, A_2)$ in the x-y-plane. Then, the numerical range is given by the convex hull of its generating line [76, 77]. We dehomogenize this equation by setting $v = 1$, replace w in Eq. 6.91 by $w = -ux - y$, and solve for the generating line with the resulting equations

$$F(u, x, y) = \det(uA_1 + A_2 - (ux + y)\mathbb{1}) = 0, \tag{6.92}$$

as well as $\partial F(u, x, y)/\partial u = 0$. To do so, we compute a Gröbner basis [198] of this system of polynomial equations that contains the polynomial

$$\begin{aligned} P(x, y) = \frac{1}{8} \big[& 2(-1 + 2x^2 + 2y^2)^2 + \cos(4\vartheta_A) + 4\cos(2\vartheta_A)(4xy(xy - \cos\vartheta_A \cos\vartheta_B)) \\ & + (-1 + 2x^2 + 2y^2)\cos(2\vartheta_B) + \cos(4\vartheta_B) \\ & + 16xy(xy\cos(2\vartheta_B) - \cos\vartheta_A \cos\vartheta_B(2(-1 + x^2 + y^2) + \cos(2\vartheta_B)))\big] \\ & \times (-1 + y^2)^2 \sin^2\vartheta_A \sin^2\vartheta_B. \end{aligned} \tag{6.93}$$

Thus, $P(x, y) = 0$ has to be satisfied for a solution of the system of polynomial equations determining the generating line.

The special cases $\sin\vartheta_A = 0$ or $\sin\vartheta_B = 0$ work analogously. Thus, let $\sin\vartheta_A = 0$ and hence, $A_2 = X \otimes (\cos\vartheta_B X + \sin\vartheta_B Z)$ which means that there is only a single measurement on the first particle. This can be simulated by a separable state in place of a possibly entangled state $|\psi\rangle$. Indeed, the state

$$\rho_{\text{sep}} = \rho_A \otimes \frac{\operatorname{Tr}_A[(X \otimes \mathbb{1}) |\psi\rangle \langle\psi|]}{\operatorname{Tr}[(X \otimes \mathbb{1}) |\psi\rangle \langle\psi|]}, \tag{6.94}$$

with $\operatorname{Tr} X\rho_A = \operatorname{Tr}[(X \otimes \mathbb{1}) |\psi\rangle \langle\psi|]$, yields the same $X \otimes M$ measurement statistics for any observable M. Hence, it is sufficient to consider product states $\rho = \rho_A \otimes \rho_B$ or Lemma 6.12. Then, the volume is

$$\operatorname{vol} L(X \otimes X, X \otimes (\cos\vartheta_B X + \sin\vartheta_B Z)) = \int_{-1}^{1} d\xi \; 2\sin\vartheta_B \sqrt{1 - \xi^2} = \pi \sin\vartheta_B. \tag{6.95}$$

Indeed, since the numerical range, and hence also its volume, changes continuously with the parameters ϑ_A and ϑ_B, this is the limit of the general volume function.

The other special case, i.e. $(1 - y^2) = 0$, describes the eigenspace with maximal or minimal eigenvalue of A_2. Because A_1 and A_2 behave the same relative to each other, we consider states $|\psi\rangle$ with $\langle\psi_\pm| A_1 |\psi_\pm\rangle = \pm 1$, which means that $|\psi_+\rangle = \alpha |{+}{+}\rangle +$

$\beta\,|--\rangle$ and $|\psi_-\rangle = \alpha\,|+-\rangle + \beta\,|-+\rangle$. Then, we obtain

$$\langle\psi_+|\,A_2\,|\psi_+\rangle = \cos\vartheta_A\cos\vartheta_B + \sin\vartheta_A\sin\vartheta_B(\alpha\beta^* + \alpha^*\beta), \tag{6.96}$$

$$\langle\psi_-|\,A_2\,|\psi_-\rangle = -\cos\vartheta_A\cos\vartheta_B + \sin\vartheta_A\sin\vartheta_B(\alpha\beta^* + \alpha^*\beta), \tag{6.97}$$

and hence, this solution yields line segments with $\pm\cos\vartheta_A\cos\vartheta_B - \sin\vartheta_A\sin\vartheta_B \leq \langle A_1\rangle \leq \pm\cos\vartheta_A\cos\vartheta_B + \sin\vartheta_A\sin\vartheta_B$ for $\langle A_2\rangle = \pm 1$, respectively. Let us keep this special case in mind for the final solution.

Finally, we consider the case

$$\begin{aligned}\tilde{P}(x,y) = {} & 2(-1+2x^2+2y^2)^2 + \cos(4\vartheta_A) + 4\cos(2\vartheta_A)(4xy(xy - \cos\vartheta_A\cos\vartheta_B)) \\ & + (-1+2x^2+2y^2)\cos(2\vartheta_B) + \cos(4\vartheta_B) \\ & + 16xy(xy\cos(2\vartheta_B) - \cos\vartheta_A\cos\vartheta_B(2(-1+x^2+y^2) + \cos(2\vartheta_B))) = 0.\end{aligned} \tag{6.98}$$

Rotating the coordinate system by $\frac{\pi}{4}$, i.e. setting $x = (\tilde{x}+\tilde{y})/\sqrt{2}$ and $y = (\tilde{x}-\tilde{y})/\sqrt{2}$, we find that the solution is given by two ellipses with semi-axes along the coordinate axes,

$$\begin{aligned}\tilde{P}(\tilde{x},\tilde{y}) = {} & 8\left[1 - \cos^2(\vartheta_A+\vartheta_B) - \tilde{x}^2(1-\cos(\vartheta_A+\vartheta_B)) - \tilde{y}^2(1+\cos(\vartheta_A+\vartheta_B))\right] \times \\ & \left[1 - \cos^2(\vartheta_A-\vartheta_B) - \tilde{x}^2(1-\cos(\vartheta_A-\vartheta_B)) - \tilde{y}^2(1+\cos(\vartheta_A-\vartheta_B))\right].\end{aligned} \tag{6.99}$$

To show that these curves are indeed reached by quantum states, we find those states explicitly. Rotating the coordinate system corresponds to considering the observables $A_\pm = (A_1 \pm A_2)/\sqrt{2}$. Thus, for the first ellipse with semi-axes of length $l_\pm^a = \sqrt{1 \pm \cos(\vartheta_A+\vartheta_B)}$, the states are given by

$$|\phi_a\rangle = \cos\varphi\,|\phi_+^a\rangle\,/\sqrt{\langle\phi_+^a|\phi_+^a\rangle} + \sin\varphi\,|\phi_-^a\rangle\,/\sqrt{\langle\phi_-^a|\phi_-^a\rangle}, \tag{6.100}$$

where

$$\begin{aligned}|\phi_+^a\rangle = {} & \sin(\vartheta_A+\vartheta_B)(|00\rangle - |11\rangle) \\ & + \left[1 + \cos(\vartheta_A+\vartheta_B) + \sqrt{2+2\cos(\vartheta_A+\vartheta_B)}\right](|01\rangle + |10\rangle),\end{aligned} \tag{6.101}$$

$$\begin{aligned}|\phi_-^a\rangle = {} & \sin(\vartheta_A+\vartheta_B)(|00\rangle - |11\rangle) \\ & - \left[1 - \cos(\vartheta_A+\vartheta_B) + \sqrt{2-2\cos(\vartheta_A+\vartheta_B)}\right](|01\rangle + |10\rangle)\end{aligned} \tag{6.102}$$

are eigenstates of $A_\pm$ with eigenvalues $l_\pm$, respectively. There is an exception when $\vartheta_A + \vartheta_B = \pi$ and hence, $|\phi_+^a\rangle = 0$. In this case, the ellipse is just a line segment

of length $\sqrt{2}$ along the y-axis. Since the respective eigenvectors of A_- with eigenvalues $\pm\sqrt{2}$ obey $\langle A_+\rangle = 0$, the line is traced out by mixtures of these states. Otherwise, it holds that $\langle \phi_+^a|\phi_-^a\rangle / \sqrt{\langle \phi_+^a|\phi_+^a\rangle \langle \phi_-^a|\phi_-^a\rangle} = -1/\sqrt{2}$, and thus, $\langle \phi_a|\phi_a\rangle = 1 - \sqrt{2}\sin\varphi\cos\varphi$, as well as $\langle \phi_\pm^a| A_\mp |\phi_\pm^a\rangle = 0$ and hence, $\langle \phi_a| A_+ |\phi_a\rangle = l_+^a(\cos^2\varphi - \sqrt{2}\sin\varphi\cos\varphi)$ and $\langle \phi_a| A_- |\phi_a\rangle = l_-^a(\cos^2\varphi - \sqrt{2}\sin\varphi\cos\varphi)$. Thus, the states $|\phi_a\rangle$ satisfy

$$\langle A_+\rangle = \frac{\langle \phi_a| A_+ |\phi_a\rangle}{\langle \phi_a|\phi_a\rangle} = l_+^a \frac{\cot\varphi(\cot\varphi - \sqrt{2})}{1 + \cot\varphi(\cot\varphi - \sqrt{2})}, \tag{6.103}$$

$$\langle A_-\rangle = \frac{\langle \phi_a| A_- |\phi_a\rangle}{\langle \phi_a|\phi_a\rangle} = l_-^a \frac{\tan\varphi(\tan\varphi - \sqrt{2})}{1 + \tan\varphi(\tan\varphi - \sqrt{2})}, \tag{6.104}$$

which indeed defines points on the desired ellipse because $(\langle A_+\rangle/l_+^a)^2 + (\langle A_-\rangle/l_-^a)^2 = 1$. Furthermore, with φ varying continuously from 0 to π, we observe the continuous variation of $(\langle A_+\rangle, \langle A_-\rangle)$ as $(1,0) \to (0,-1) \to (-1,0) \to (0,1) \to (1,0)$ where the change in-between each step is monotonic. Thus, the states indeed trace out the ellipse. For the second ellipse, an analogous argument holds for the states $|\phi_b\rangle = \cos\varphi\,|\phi_+^b\rangle / \sqrt{\langle \phi_+^b|\phi_+^b\rangle} + \sin\varphi\,|\phi_-^b\rangle / \sqrt{\langle \phi_-^b|\phi_-^b\rangle}$, where

$$\begin{aligned} |\phi_+^b\rangle &= \sin(\vartheta_A - \vartheta_B)(|00\rangle + |11\rangle) \\ &\quad - \left[1 + \cos(\vartheta_A - \vartheta_B) - \sqrt{2 + 2\cos(\vartheta_A - \vartheta_B)}\right](|01\rangle - |10\rangle), \end{aligned} \tag{6.105}$$

$$\begin{aligned} |\phi_-^b\rangle &= \sin(\vartheta_A - \vartheta_B)(|00\rangle + |11\rangle) \\ &\quad + \left[1 - \cos(\vartheta_A - \vartheta_B) - \sqrt{2 - 2\cos(\vartheta_A - \vartheta_B)}\right](|01\rangle - |10\rangle). \end{aligned} \tag{6.106}$$

The convex hull of these two ellipses is bounded partly by straight line segments. These line segments must indeed correspond to the maximal and minimal eigenvalues of A_1 and A_2 which we considered in the special case $y = \pm 1$ since there cannot be any state beyond $-1 \leq x, y \leq 1$. To compute the volume, we divide the convex hull of the ellipses into a polygon with eight vertices and the tips of the ellipses, whose volume can be computed via integration. This leads to the final result

$$\begin{aligned} \operatorname{vol} L &= 2\left[\cos\vartheta_- - \cos\vartheta_+ + |\sin\vartheta_-| \arccos|\cos\frac{\vartheta_-}{2}| + |\sin\vartheta_+| \arccos|\sin\frac{\vartheta_+}{2}|\right] \\ &= 2\left[\cos\vartheta_- - \cos\vartheta_+ + |\frac{\vartheta_-}{2}\sin\vartheta_-| + |\left(\frac{\vartheta_+}{2} - \frac{\pi}{2}\right)\sin\vartheta_+|\right], \end{aligned} \tag{6.107}$$

where $\vartheta_\pm = \vartheta_A \pm \vartheta_B$. □

Together, Lemma 6.12 and Lemma 6.13 prove Theorem 6.11. Using this result, numerical optimization shows that $\mu_{2,2,2}^{\otimes,\mathrm{LT}} = \frac{1}{2}$. Furthermore, Theorem 6.11 can be slightly

generalized to operators $A_j = \alpha_j B_1 \otimes B_2 + \beta_j C_1 \otimes C_2 + \gamma_j \mathbb{1} \otimes \mathbb{1}$ via affine transformations.

We also compute the volume ratio for certain instances of more than two measurements. More precisely, the observables $A_1 = X \otimes X$, $A_2 = X \otimes Y$, and $A_3 = Z \otimes Z$ yield a ratio of $1/3$, adding also the observable $A_4 = Y \otimes Z$ yields $1/6$. For the computation, we again make use of the fact that simultaneous local unitary transformations of the observables do not alter the numerical range. In this case, we consider $U = \mathbb{1} \otimes \exp(-iZ\varphi/2)$ which leaves A_3 invariant and transforms A_1 and A_2 as

$$A_1(\varphi) = \cos\varphi X \otimes X + \sin\varphi X \otimes Y, \tag{6.108}$$

$$A_2(\varphi) = -\sin\varphi X \otimes X + \cos\varphi X \otimes Y. \tag{6.109}$$

Thus, it corresponds to a rotation around the A_3-coordinate axis and hence, the numerical range is rotationally invariant around this axis and it suffices to consider A_1 and A_3. The well known entanglement witnesses $W = \mathbb{1} \pm X \otimes X \pm Z \otimes Z$ bound the separable numerical range, while with general quantum states any $\langle A_1 \rangle$ can be reached independently of $\langle A_3 \rangle$ and vice verse. This leaves us with a square inside a square as shown in Fig. 6.6. The volume ratio of the two solids of revolution generated by rotating around the A_3-axis is $1/3$, given by two cones glued to each other at the base inside a cylinder.

Furthermore, additionally measuring the observable $A_4 = Y \otimes Z$ yields a volume ratio of $1/6$. Similar to before, we, in addition, consider the local unitary transformation $V = \exp(-iX\phi/2) \otimes \mathbb{1}$, which leaves A_1 and A_2 invariant, while transforming A_3 and A_4 as

$$A_3(\phi) = \cos\phi Z \otimes Z + \sin\phi Y \otimes Z, \tag{6.110}$$

$$A_4(\phi) = -\sin\phi Z \otimes Z + \cos\phi Y \otimes Z. \tag{6.111}$$

The resulting four-dimensional solids of revolution have a volume of

$$\operatorname{vol} L = \int_0^1 dR \int_0^{2\pi} d\varphi R \int_0^1 dr \int_0^{2\pi} r = \pi^2, \tag{6.112}$$

$$\operatorname{vol} L_{\mathrm{SEP}} = \int_0^1 dR \int_0^{2\pi} d\varphi R \int_0^{1-R} dr \int_0^{2\pi} r = \frac{\pi^2}{6}, \tag{6.113}$$

and hence, we obtain a volume ratio of $1/6$.

In Table 6.1, we summarize our results for extremal volume ratios for the separable numerical range compared to the general numerical range in the case of two qubits,

k	$\mu_{2,2,k}$		$\mu_{2,2,k}^{\otimes,\mathrm{LT}}$	
	lower bound	upper bound	lower bound	upper bound
1	$\sqrt{2}-1$	$1/2$	1	
2	$1/9$	$1/4$	$1/2^*$	
3	$1/27$	$1/5$	$1/27$	$1/3$
4	$1/81$	$1/6$	$1/81$	$1/6$
15	$8/33$		$8/33$	

TABLE 6.1: This table shows lower and upper bounds for $\mu_{2,2,k}$ and $\mu_{2,2,k}^{\otimes,\mathrm{LT}}$ for different numbers of measurements k on a two-qubit system. The starred value is obtained partly via numerical two-parameter optimization.

bounding the minimal volume ratio $\mu_{2,2,k}$. The upper bounds for $\mu_{2,2,2}$ and $\mu_{2,2,3}$ are presumably tight.

6.6 Commutativity and entanglement detection

The commutator between observables A and B is defined as $[A, B] = AB - BA$, and A and B are simultaneously diagonalizable if, and only if, their commutator vanishes [19]. Operationally, this means that A and B are jointly measurable if, and only if, their commutator satisfies $[A, B] = 0$ [248]. In the proof of Lemma 6.13, we have seen that the measurement statistics of a single local measurement on one party and arbitrary measurements on the other party are always realizable by a separable state. Thus, for product observables $A_1 = B_1 \otimes B_2$ and $A_2 = C_1 \otimes C_2$ with $[B_1, C_1] = 0$ or $[B_2, C_2] = 0$, the separable and general numerical range coincide and hence, entanglement detection is impossible.

However, it is not obvious whether A_1 and A_2 always detect some entangled states if neither B_1 and C_1, nor B_2 and C_2 commute. In Ref. [223], the authors answer this question in the affirmative for two-qubit systems. In any other dimension, however, it is not the case [223, 229]. The authors of Ref. [229] still prove that entanglement detection is always possible if not all common eigenvectors of B_1 and C_1 as well as those of B_2 and C_2 correspond to a zero-eigenvalue. This sufficient criterion shows that almost all A_1 and A_2 allow for entanglement detection even in higher dimensions.

Nevertheless, we present six observables $A_1, \ldots, A_6$ of a qutrit-qutrit system that do

not commute locally but are still useless for entanglement detection if only their expectation values are observed. Namely, we consider

$$A_1 = (1 \oplus \mu X) \otimes \left(1 \oplus \mu \frac{Y+Z}{\sqrt{2}}\right), \quad A_2 = \left(1 \oplus \mu \frac{Y+Z}{\sqrt{2}}\right) \otimes (1 \oplus \mu X), \tag{6.114}$$

$$A_3 = (1 \oplus \mu Y) \otimes \left(1 \oplus \mu \frac{Z+X}{\sqrt{2}}\right), \quad A_4 = \left(1 \oplus \mu \frac{Z+X}{\sqrt{2}}\right) \otimes (1 \oplus \mu Y), \tag{6.115}$$

$$A_5 = (1 \oplus \mu Z) \otimes \left(1 \oplus \mu \frac{X+Y}{\sqrt{2}}\right), \quad A_6 = \left(1 \oplus \mu \frac{X+Y}{\sqrt{2}}\right) \otimes (1 \oplus \mu Z), \tag{6.116}$$

where 1 denotes a 1×1-matrix with its single element being 1. The observables do not commute pairwise for any $\mu > 0$. If entanglement detection were possible with these observables, then for some α_j, the largest or smallest eigenvalue of $\Omega = \sum_j \alpha_j A_j$ must not correspond to a product state. Because of the simple structure of Ω, we can consider different subspaces separately. In the subspace spanned by $|01\rangle$ and $|02\rangle$ and the subspace spanned by $|10\rangle$ and $|20\rangle$, there are product eigenvectors with eigenvalues

$$\lambda_\pm^\otimes = \pm\mu \sqrt{(\alpha_2 + \frac{\alpha_3}{\sqrt{2}} + \frac{\alpha_5}{\sqrt{2}})^2 + (\frac{\alpha_1}{\sqrt{2}} + \alpha_4 + \frac{\alpha_5}{\sqrt{2}})^2 + (\frac{\alpha_1}{\sqrt{2}} + \frac{\alpha_3}{\sqrt{2}} + \alpha_6)^2}, \tag{6.117}$$

$$\eta_\pm^\otimes = \pm\mu \sqrt{(\alpha_1 + \frac{\alpha_4}{\sqrt{2}} + \frac{\alpha_6}{\sqrt{2}})^2 + (\frac{\alpha_2}{\sqrt{2}} + \alpha_3 + \frac{\alpha_6}{\sqrt{2}})^2 + (\frac{\alpha_2}{\sqrt{2}} + \frac{\alpha_4}{\sqrt{2}} + \alpha_5)^2}, \tag{6.118}$$

respectively. Let us make some crude approximations. First, the above eigenvalues are bounded by the terms in the square root, for instance, we have that $|\lambda_\pm^\otimes| \geq \mu|\alpha_2 + \frac{\alpha_3}{\sqrt{2}} + \frac{\alpha_5}{\sqrt{2}}|$. Second, the entangled eigenvectors must lie within the subspace spanned by $|11\rangle$, $|12\rangle$, $|21\rangle$, and $|22\rangle$. Let us denote the projector onto this subspace as P_{ent}, then the eigenvalues of the matrices $P_{\text{ent}} A_j P_{\text{ent}}$ are bounded from above and below by $\pm\mu^2$, respectively. Hence, the eigenvalues of $P_{\text{ent}} \Omega P_{\text{ent}}$ are bounded by $\pm\mu^2 \sum_j |\alpha_j|$. Finally, we can express the α_j in terms of the bounds for $\lambda_\pm^\otimes/\mu$ and $\eta_\pm^\otimes/\mu$, e.g.,

$$\begin{aligned} \alpha_1 =& \sqrt{2}(\alpha_2 + \frac{\alpha_3}{\sqrt{2}} + \frac{\alpha_5}{\sqrt{2}}) + 1(\alpha_1 + \frac{\alpha_4}{\sqrt{2}} + \frac{\alpha_6}{\sqrt{2}}) \\ & - 1(\frac{\alpha_2}{\sqrt{2}} + \alpha_3 + \frac{\alpha_6}{\sqrt{2}}) - 1(\frac{\alpha_2}{\sqrt{2}} + \frac{\alpha_4}{\sqrt{2}} + \alpha_5), \end{aligned} \tag{6.119}$$

implying that, if the absolute values of those bounds do not exceed x, then $|\alpha_1| \leq (3+\sqrt{2})x$. Furthermore, if $\sum_j |\alpha_j| = x$, then at least one of the above bounds' is at least $x/(18+6\sqrt{2})$. Thus, choosing μ small enough, i.e., $\mu < 1/(18+6\sqrt{2})$ according to our crude approximations, makes sure that the eigenvalues of the entangled eigenstates of Ω do not correspond to the maximal or minimal eigenvalue because their absolute value is smaller than $|\lambda_\pm^\otimes|$ or $|\eta_\pm^\otimes|$. Hence, entanglement detection is impossible.

We conjecture that there might not be more than six locally noncommuting observables which cannot be used for entanglement detection. The reason is that, at least with this construction, the eigenvalues of $\lambda_{\pm}^{\otimes}$ and $\eta_{\pm}^{\otimes}$ would depend on more variables without yielding more information as the local Bloch vector is always three-dimensional. In other words, we conjecture that any set of seven locally noncommuting qutrit-qutrit observables can be used to detect some entangled state. This result would establish a new connection between commutativity and entanglement detection and it would be desirable to investigate it for systems of different dimensions.

6.7 Conclusion

Considering the (separable) numerical range is the most general concept for entanglement detection for given observables. The volume ratio of the separable numerical range compared to the general numerical range indicates how difficult it is to verify entanglement with statistical significance. More precisely, in an experiment, the measurement data imply a confidence region for the expectation values of the underlying state when the given observables are measured. To detect entanglement, this confidence region and the separable numerical range have to be disjoint, which is generically more likely for smaller volume ratios. We provide a general lower bound for any number of particles, local dimensions, and number of measurements and consider extreme cases. For two qubits and a single observable, we examine the numerical range generated by absolutely separable states and obtain a lower bound for the volume ratio of $\sqrt{2}-1$. Numerical investigations lead us to the conjecture that the minimal volume ratio is indeed $1/2$, thus it would be desirable to close this gap for the most basic case in the future. Furthermore, we focus on product observables which are easier to measure in experiments. For two locally traceless two-qubit observables, we explicitly provide the volumes for the (separable) numerical range. In addition, we provide an example of six locally noncommuting qutrit-qutrit observables that are insufficient to detect entanglement, i.e., their separable and general numerical range coincide. This sheds new light on the relation between commutativity and entanglement detection.

Besides experimental entanglement detection, our results also give insight into the geometry of (separable) quantum states and their relation. The (separable) numerical range is an affine transformation of a projection of the set of separable or general quantum states, i.e., a lower-dimensional shadow. This leads to interesting mathematical questions about volume ratios of shadows of convex bodies. In the future,

we hope to extend our results and consider more quantum phenomena such as genuine multipartite entanglement, steering, and nonlocality and their respective volume ratios.

7 Certifying quantum memories with coherence

Prerequisites

2.2 Quantum mechanics
2.3 Quantum channels
2.4 Coherence
2.5 Entanglement

7.1 Introduction

The main parts of this chapter have been published in Publication (B) [29]. Partly, the results from Ref. [29] have been published in [18]. This concerns the proofs to general properties of our memory performance measures as well as higher-dimensional systems, while results on single-qubit channels are covered here. We repeat the main ideas that motivate the definition of the introduced quantities to allow for a coherent treatment.

In order to work, quantum computers need reliable and well-characterized routines and devices. The loss of quantum coherence, however, is one of the major obstacles on the way to a scalable platform for quantum computing, and the suppression of decoherence is known as one of the DiVincenzo criteria for quantum computers [249]. In any computing architecture, the memory plays an essential role. Quantum computers are no exception and furthermore, quantum memories play a central role in the development of quantum repeaters [250–252]. Consequently, the search for reliable systems that store quantum states for a reasonable amount of time while preserving quantum properties is an active area of research [253–259].

A possible way to verify the proper functioning of quantum gates and quantum memories is to completely characterize their behavior via quantum process tomography [260, 261]. This, however, requires an effort exponentially increasing in the size of the

system. More importantly, it is desirable to determine the change of physical properties, such as entanglement and coherence, under the prescribed time evolution since these convey the quantumness of the underlying process. By contrast, a complete characterization does not distinguish between these characteristics and minor details, making it harder to identify the main features. Therefore, it is beneficial to describe devices directly by their effect on physical phenomena.

Several methods have been suggested to characterize quantum memories: The quantumness of channels has been assessed based on whether they preserve entanglement or not, focusing on reducing the number of measurements in bipartite optical systems [262, 263]. Furthermore, quantum steering has been considered as a way to evaluate the performance of quantum channels in the case of untrusted measurement devices, again distinguishing channels that do and do not preserve entanglement [264]. Finally, a resource theory of quantum memories has been developed [265]. The free resources are measure-and-prepare channels since they simply store classical information. Using arbitrary pre- and post-processing accompanied by unlimited classical memory as free operations, the authors establish a game-theoretic way to assess quantum memory performance based on the entanglement of the corresponding Choi state. Nonetheless, these attempts require either well characterized test states as inputs, many measurements on the output or an advanced scheme to be implemented.

First conditions on how to generally assess the performance of quantum memories were discussed in Ref. [250]. This work suggests to use the fidelity as a performance measure. In fact, instead of the fidelity, any distance measure between the input- and the output state would be suitable to measure the performance of such devices, e.g., a measure based on the coherence of the states [266]. As the authors of Ref. [250] note, however, the fidelity is sensitive to unitary transformations of the input, which may be compensated by the quantum computer controlling the interface. With this in mind, the authors propose to use the purity of the memory instead, which is indeed insensitive to unitary transformations. However, the purity of a channel yielding a fixed, pure state independent of the input is maximal, but such a channel would certainly not qualify as a proper memory.

With these considerations in mind, we introduce general criteria for quality measures of quantum memories. First, they should clearly distinguish schemes that require storing only classical information from perfect unitary transformations. Second, as we assume that unitary transformations can be corrected by the underlying quantum computer, the quality of a quantum memory should be invariant under such unitary transformations.

We then propose a measure that obeys these natural properties using the phenomenon of coherence. The key idea is that an ideal quantum memory preserves the coherence in any basis. The measure can be used to prove that a memory preserves entanglement and, moreover, it can be estimated with few measurements, without the need of well characterized input states. Our concept may be generalized to characterize also other quantum primitives such as teleportation schemes and, using generalized notions of coherence [267, 268], also to multi-particle quantum gates.

7.2 Memory quality measures

To start, let us study what physical properties a measure for the quality of a quantum memory should possess. As non-classical properties are essential for many quantum algorithms, the storage should preserve as many of these properties for as long as possible. However, the microscopic quantum system, such as a photon or an ion, storing the quantum state may get lost which is measured by the efficiency [250]. In this work, we measure the quality of a memory conditionally on the preserved systems, making the efficiency an additional, independent performance indicator. Additionally, the storage duration is essential but is also treated as an extra indicator.

Then, a quantum memory can be described as a quantum channel, with the optimal memory given by the identity channel. In practice, however, this is rather difficult to achieve. Contrary to that, measure-and-prepare (M&P) schemes (also known as entanglement-breaking channels) can be easily simulated using only classical storage. One just performs measurements on the input state and stores the result. Based on that, one then prepares a quantum state on demand as we described in Section 2.5.

These two examples show that a measure for the quality of a quantum memory should have two natural properties: First, it should be maximal for memories that preserve the input state perfectly. As we assume that we can perform unitary rotations, we also allow the memory to apply a known and fixed unitary rotation to the input. Second, the measure should have a non-maximal quality for the M&P schemes described above, certifying genuine quantum storage.

Definition 7.1. A map $Q(\mathcal{M}) \in [0,1]$ for a channel $\mathcal{M}$ is called *memory quality measure*, if it satisfies

M1: $Q(\mathcal{M}) = 1$ if $\mathcal{M}(\rho) = V\rho V^\dagger$ for some unitary V,

M2: $Q(\mathcal{M}) \leq c$ for some constant $c \in [0,1)$ if $\mathcal{M}$ is a M&P channel.

A memory quality measure is called *sharp*, if it additionally fulfills

M1′: $Q(\mathcal{M}) = 1 \Leftrightarrow \mathcal{M}(\rho) = V\rho V^\dagger$ for some unitary V.

Obviously, condition M1 implies that the identity channel has unit quality. Furthermore, for continuous sharp measures, M1′ implies M2 since M&P channels have a finite distance to the set of unitary channels due to the compactness of the set [20].

7.3 Definition of the measures

Recently, there has been growing interest in coherence in the context of resource theories [35]. This has led to the development of various coherence measures that quantify the amount of coherence present in a given D-dimensional state. For a fixed basis (defined by some unitary U such that $|b_i\rangle := U|i\rangle$), we use the normalized robustness of coherence [47]

$$C_U(\rho) := \frac{1}{D-1} \min_{\tau\in\mathcal{D}} \left\{ s \geq 0 \middle| \frac{\rho + s\tau}{1+s} \in \mathcal{I}_U \right\}, \tag{7.1}$$

where $\mathcal{D}$ is the set of all D-dimensional states and $\mathcal{I}_U$ is the set of incoherent (i.e., diagonal) D-dimensional states w.r.t. the basis $U|j\rangle$. However, our results are valid for any continuous and convex coherence measure with the property that the only states maximizing the measure for a fixed basis U are given by

$$|\Psi_U^{\vec{\alpha}}\rangle := \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{i\alpha_j} |b_j\rangle = UZ_{\vec{\alpha}} |+\rangle, \tag{7.2}$$

where $\vec{\alpha}$ is some D-dimensional vector of phases and $Z_{\vec{\alpha}}$ is a diagonal unitary matrix with entries $e^{i\alpha_j}$, acting on $|+\rangle := \frac{1}{\sqrt{D}} \sum_i |i\rangle$. Note that the states in Eq. (7.2) maximize any valid coherence monotone, and for many prominent coherence measures such as the robustness of coherence [44], the l_1-norm of coherence [45], and the relative entropy measure [46], they are the only states doing so. Furthermore, they are also maximally coherent in a resource theoretic sense as described in Section 2.4 [35, 45].

We define a physically motivated quality measure from the following considerations: Given a quantum channel $\mathcal{M}$, there is a "most classical" basis, in which even the most robust maximally coherent state with respect to that basis is mapped to a state with small coherence. This basis is identified by our proposed measure, and the conserved coherence in this basis defines the quality.

Definition 7.2. For a quantum channel $\mathcal{M}$, the quality Q_0 is given by

$$Q_0(\mathcal{M}) := \min_U \max_{\vec{\alpha}} C_U[\mathcal{M}(|\Psi_U^{\vec{\alpha}}\rangle)]. \tag{7.3}$$

Here, we write $\mathcal{M}(|\Psi_U^{\vec{\alpha}}\rangle)$ instead of $\mathcal{M}(|\Psi_U^{\vec{\alpha}}\rangle\langle\Psi_U^{\vec{\alpha}}|)$ for convenience. If $Q_0(\mathcal{M}) = 1$, then in any basis at least one maximally coherent state is preserved.

As the robustness of coherence has a clear operational interpretation, the same is also true for the quality measure Q_0. Indeed, it certifies the usefulness of the quantum memory $\mathcal{M}$ for the phase discrimination task described in Section 2.4 independent from the preferred basis.

Despite the clear physical interpretation of this measure, there are related quantities which turn out to be useful for the discussion. Therefore, we introduce two additional parameters, which provide an upper and lower bound on Q_0. First, we consider the minimal coherence left in any basis of the most robust maximally coherent states if one minimizes over their bases.

Definition 7.3. For a quantum channel $\mathcal{M}$, the quantity Q_- is defined by

$$Q_-(\mathcal{M}) := \min_{U,U'} \max_{\vec{\alpha}} C_{U'}[\mathcal{M}(|\Psi_U^{\vec{\alpha}}\rangle)]. \tag{7.4}$$

In contrast to Q_0, the basis of coherence is varied independently of the basis of the maximally coherent states. Thus, we have that $Q_-(\mathcal{M}) \leq Q_0(\mathcal{M})$. Second, as an upper bound to Q_0, we consider the minimal coherence in any basis maximized over all states in the range.

Definition 7.4. For a quantum channel $\mathcal{M}$, the quantity Q_+ is defined by

$$\begin{aligned} Q_+(\mathcal{M}) &:= \min_U \max_\rho C_U[\mathcal{M}(\rho)] \\ &= \min_U \max_{|\psi\rangle} C_U[\mathcal{M}(|\psi\rangle)], \end{aligned} \tag{7.5}$$

where the equality is due to the convexity of the coherence measure and linearity of $\mathcal{M}$.

Here, in contrast to Q_0, the maximization is not limited to maximally coherent states. Hence, it holds that

$$Q_-(\mathcal{M}) \leq Q_0(\mathcal{M}) \leq Q_+(\mathcal{M}). \tag{7.6}$$

Due to the minimization over all bases U (and U' for Q_-), for all channels $\mathcal{M}$ and unitary channels $\mathcal{V}$ with $\mathcal{V}(\rho) = V\rho V^\dagger$ where V is some unitary, we have the following identities:

$$\begin{aligned} Q_\pm(\mathcal{M}) &= Q_\pm(\mathcal{V} \circ \mathcal{M}) = Q_\pm(\mathcal{M} \circ \mathcal{V}), \\ Q_0(\mathcal{M}) &= Q_0(\mathcal{V} \circ \mathcal{M} \circ \mathcal{V}^{-1}). \end{aligned} \tag{7.7}$$

The quantities $Q_\pm$ are completely invariant under prior and subsequent rotations, whereas Q_0 is only invariant under joint rotations. As such, the quantities $Q_\pm$ are useful to obtain bounds on Q_0.

During our work, a similar approach has been investigated that also uses a coherence-based performance measure. Instead of considering the most robust or maximally coherent states, the authors are interested in the average coherence preserved over all states [269].

7.4 Properties of the measures

Using the Sinkhorn normal form of unitaries [270] together with the continuity of $Q_\pm$ and Q_0, one can show the following theorem.

Theorem 7.5. *The quantities $Q_\pm$ and Q_0 are sharp memory quality measures.*

Additionally, the quality measure Q_+ satisfies a useful pre-processing property:

Lemma 7.6. *The quality measure Q_+ cannot be increased by pre-processing the input, i.e., $Q_+(\mathcal{M} \circ \mathcal{N}) \leq Q_+(\mathcal{M})$ for all quantum channels $\mathcal{M}$ and $\mathcal{N}$.*

The proofs can be found in Refs. [18, 29]. For Q_-, we can prove a similar statement for the case of unital, i.e., channels that map the maximally mixed state to itself, single-qubit channels (see Lemma 7.10).

In comparison, the measures introduced in Ref. [265] are monotonous under pre- and post-processing using unlimited classical memory and preexisting randomness. This is not true for Q_0 and $Q_\pm$. In the notation introduced in Section 2.3, a counterexample is given by the channel $\mathcal{N}$ defined by $\vec{\lambda} = (0, 0, 1)$ and vanishing $\vec{\kappa}$, and the M&P channel $\mathcal{M}$ maximizing Q_0, given by $\vec{\lambda} = (0, 0, \frac{1}{\sqrt{2}})$, $\vec{\kappa} = (\frac{1}{\sqrt{2}}, 0, 0)$. Then, $Q_0(\mathcal{M} \circ \mathcal{N}) = \frac{1}{\sqrt{2}} \not\leq Q_0(\mathcal{N}) = 0$. This counterexample also works for Q_+. For Q_-, choosing $\mathcal{N}$ as the planar channel with semi-axes $\vec{\lambda} = (0, \frac{1}{2}, \frac{1}{2})$ and zero displacement, and $\mathcal{M}$ as the channel maximizing Q_-, i.e., defined by $\vec{\lambda} = (0, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}})$ and displacement

$\vec{\kappa} = (\frac{1}{\sqrt{5}}, 0, 0)$, leads to $Q_-(\mathcal{M} \circ \mathcal{N}) = \frac{1}{2\sqrt{5}} \not\leq Q_-(\mathcal{N}) = 0$. The non-monotonicity is expected for measures based on coherence, because in contrast to entanglement, coherence can be created locally. Furthermore, if a measure is monotonous under the operations defined in Ref. [265], it would assign the same quality to all M&P channels. However, some M&P channels are more useful for the task of phase discrimination than others.

7.5 The single-qubit case

To find bounds on the quality of single-qubit M&P channels, we deploy a geometric approach using the action of qubit-qubit channels in the Bloch picture as described in Section 2.3. Any maximally coherent state is a pure state and, vice versa, any pure state is maximally coherent in some basis. Since any transformation of $\vec{v}$ can be decomposed into rotations, contractions and a translation, the set of maximally coherent states in a fixed basis, forming a great circle in the Bloch picture, is mapped onto the boundary of an ellipse given by a cut through the center of the ellipsoid. $Q_-(\mathcal{M})$ determines the axis in the Bloch sphere and the ellipse on the image's surface of $\mathcal{M}$ that minimize the maximal distance of any point on this ellipse to the axis. This is because, in the computational basis, $C_{\mathbb{1}}(\rho) = |v_x + iv_y| = \sqrt{v_x^2 + v_y^2}$ [47] which is the distance of a point at $\vec{v}$ from the $z-$axis that defines the computational basis. For any other basis, the Bloch sphere can simply be rotated leading to the same geometric result for any basis. For $Q_0(\mathcal{M})$, the ellipse is fixed by the axis depending on the channel $\mathcal{M}$. To find an upper bound on the measures Q_- and Q_+ (and from the latter for Q_0), it is sufficient to replace the minimization over all axes by a fixed set of directions in the Bloch sphere, which allows to obtain the following bounds.

Lemma 7.7. *Let $\mathcal{M}$ be a single-qubit channel defined by displacement vector $\vec{\kappa}$ and transformation matrix Λ with singular values $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Let $\vec{\kappa} = (\kappa_1, \kappa_2, \kappa_3)^T$ in the bases where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. Then, $Q_-(\mathcal{M}) \leq \min(\sqrt{\kappa_1^2 + \kappa_2^2} + \lambda_1, \lambda_2)$ and $Q_0(\mathcal{M}) \leq Q_+(\mathcal{M}) \leq \min(\sqrt{\kappa_1^2 + \kappa_2^2} + \lambda_2, \lambda_3)$.*

Proof. Instead of minimizing over all bases, we restrict the minimization to a discrete set to obtain an upper bound. For both $Q_-(\mathcal{M})$ and $Q_+(\mathcal{M})$, we consider the axes along $\vec{\kappa}$ and along the largest singular value of Λ.

To obtain an upper bound on $Q_+(\mathcal{M})$, we simply take into account all states on the surface of the ellipsoid. The largest possible distance to the axis along $\vec{\kappa}$ clearly is λ_3 since the axis goes through the center of the ellipsoid (see right in Fig. 7.1). Similarly, the distance from the axis along λ_3 is the distance to the center, which is given by

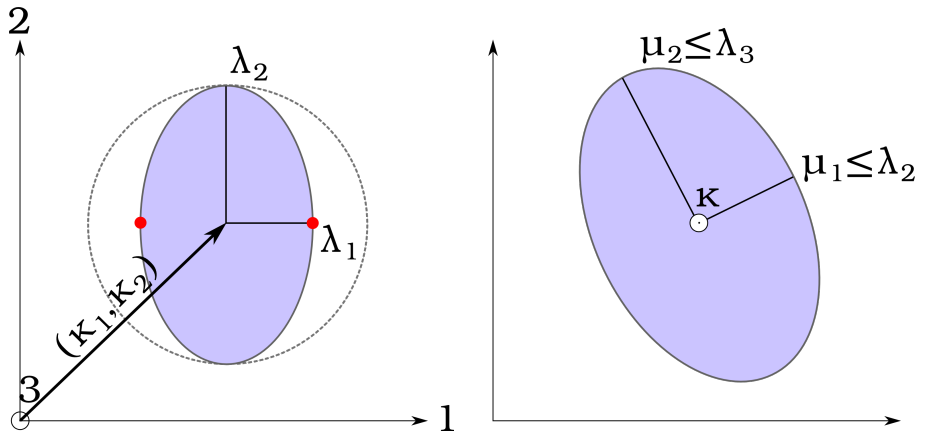


FIGURE 7.1: [29] Left: Projection of the ellipsoid in 1-2-direction to obtain upper bounds on the measures. The red dots indicate the points of the image of maximally coherent states in some basis which touch the boundary of the projected ellipse. Right: Projection of the ellipsoid in the direction of $\vec{\kappa}$. The semi-axes of the projection are bounded by the semi-axes of the ellipsoid.

$\sqrt{\kappa_1^2+\kappa_2^2}$, plus at most λ_2 since the axis is parallel to λ_3 (see left in Fig. 7.1). Because of the minimization over all bases, an upper bound is then given by $\min(\sqrt{\kappa_1^2+\kappa_2^2}+\lambda_2, \lambda_3)$.

In the case of Q_-, we can additionally choose the set of maximally coherent states. Since the channel $\mathcal{M}$ corresponds to an affine transformation of the Bloch vector, any ellipse on the surface of the ellipsoid with the same center as the ellipsoid is the image of a great circle on the surface of the Bloch sphere. Each of these circles is the set of maximally coherent states with respect to some basis. Hence, we can choose any ellipse on the surface of the ellipsoid and determine the maximal distance to the chosen axis to obtain an upper bound. For the axis along $\vec{\kappa}$, we choose the ellipse with semi-axes λ_1 and λ_2. Then, the maximal distance is at most λ_2 since the axis goes through the center of the ellipse. In the case of the axis along λ_3, the ellipse with semi-axes λ_1 and λ_2 limits the maximal distance to $\sqrt{\kappa_1^2+\kappa_2^2}+\lambda_1$ (see left in Fig. 7.1). Again, the minimum of the cases considered gives an upper bound on $Q_-(\mathcal{M})$. □

One can also find lower bounds on the quantities, which will later be useful for applications.

Lemma 7.8. *Let $\mathcal{M}$ be a single-qubit channel defined by displacement vector $\vec{\kappa}$ and transformation matrix Λ with singular values $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Then, $Q_0(\mathcal{M}) \geq Q_-(\mathcal{M}) \geq \lambda_1$ and $Q_+(\mathcal{M}) \geq \lambda_2$. If $\mathcal{M}$ is unital ($\vec{\kappa}=0$), equality holds for $Q_\pm$.*

Proof. In order to find lower bounds, we have to show the bound in all coherence bases.

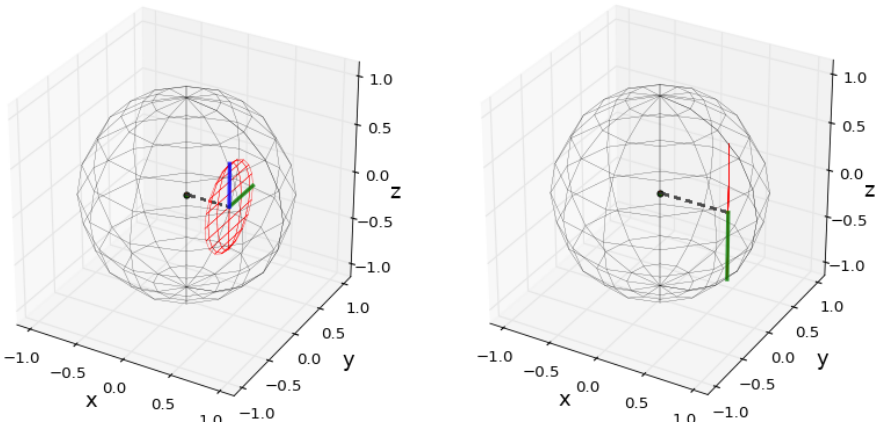


FIGURE 7.2: [29] Ellipsoid representations of the M&P channels that maximize the different quality measures. The displacement vector $\vec{\kappa}$ is depicted by a black dotted line, the semi-axes in blue and green. Left: The M&P channel maximizing Q_- maps to a disk of radius $\frac{1}{\sqrt{5}}$, displaced by $\frac{1}{\sqrt{5}}$. Right: The M&P channel maximizing Q_0 and Q_+, mapping to a straight line of length $\frac{2}{\sqrt{2}}$, displaced by $\frac{1}{\sqrt{2}}$.

For Q_+, we have to consider – for every coherence basis – the maximal distance to the center of the projection of the ellipsoid onto the plane perpendicular to the coherence direction. This projection is an ellipse with semi-axes $\mu_1 \geq \lambda_1$ and $\mu_2 \geq \lambda_2$, displaced by some vector from the center. If the displacement is zero, the maximal distance is given by μ_2 and therefore at least λ_2. For non-vanishing displacement, the maximal distance can only increase, yielding the lower bound for Q_+.

For $Q_-(\mathcal{M})$, we additionally have to minimize the maximal distance to the axis of two opposite points on this ellipse, due to the additional minimization over the input coherent states. This is in any case larger than μ_1 and therefore larger than λ_1.

Finally, if the channel is unital, note that the minimum over the coherence bases is attained in the direction of λ_3, where for $Q_-(\mathcal{M})$, we consider the states mapped to an ellipse along the λ_1-λ_3-axes, giving a maximum distance of λ_1. For $Q_+(\mathcal{M})$, the maximum distance of the non-displaced ellipsoid in this basis is given by λ_2. □

The upper bounds on the quality measures can be used to obtain tight bounds for M&P qubit channels.

Theorem 7.9. *Let $\mathcal{M}$ be a single-qubit M&P channel. Then, it holds that*

$$Q_0(\mathcal{M}) \leq Q_+(\mathcal{M}) \leq \frac{1}{\sqrt{2}} \tag{7.8}$$

and $Q_-(\mathcal{M}) \leq \frac{1}{\sqrt{5}}$. *Additionally, if* $\mathcal{M}$ *is unital* ($\vec{\kappa} = 0$),

$$Q_0(\mathcal{M}) \leq Q_+(\mathcal{M}) \leq \frac{1}{2} \tag{7.9}$$

and $Q_-(\mathcal{M}) \leq \frac{1}{3}$. *All of these bounds are tight.*

Proof. Let $\mathcal{M}$ be defined by displacement vector $\vec{\kappa}$ and transformation matrix Λ with singular values $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Since we only consider Q_+ and Q_-, we can assume w.l.o.g. that $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, i.e., we ignore possible signs in the canonical form of Λ as they do not change the shape of the resulting ellipse-shaped image of the channel. Let $\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$. Complete positivity of a single-qubit channel $\mathcal{M}$ is equivalent to $\rho_\mathcal{M} \geq 0$ where $\rho_\mathcal{M}$ is the Choi matrix of $\mathcal{M}$ [30–32, 271]. Using Descartes's rule of signs [152] on the characteristic polynomial of the Choi matrix $\rho_\mathcal{M}$, complete positivity of the channel is equivalent to the following set of inequalities

$$|\vec{\kappa}|^2 + |\vec{\lambda}|^2 \leq 3, \tag{7.10}$$

$$|\vec{\kappa}|^2 + |\vec{\lambda}|^2 - 2\lambda_1\lambda_2\lambda_3 \leq 1, \tag{7.11}$$

$$(1 - |\vec{\kappa}|^2)^2 - 2(1 - |\vec{\kappa}|^2)|\vec{\lambda}|^2 - \frac{1}{2}|\vec{\lambda}|^4 + 8\lambda_1\lambda_2\lambda_3 + \frac{1}{2}\sum_i D_i^2 - 4\vec{K}\cdot\vec{L} \geq 0, \tag{7.12}$$

where $D_i = \sum_{j=1}^3 (-1)^{\delta_{ij}}\lambda_i^2$, $\vec{K} = (\kappa_1^2, \kappa_2^2, \kappa_3^2)^T$ and $\vec{L} = (\lambda_1^2, \lambda_2^2, \lambda_3^2)^T$ [272, 273]. Similarly, single-qubit channels are M&P channels if and only if $\frac{1}{2}\mathbb{1} - \rho_\mathcal{M}$ is positive semi-definite [271]. This yields the same set of equations with $\lambda_i \leftrightarrow -\lambda_i$. In the following, we apply these restrictions to Lemma 7.7. Clearly, the bounds from Lemma 7.7 only become worse if $\vec{\kappa}$ is rotated such that $\vec{\kappa} = (|\vec{\kappa}|, 0, 0)^T$. However, rotating a M&P channel in such a way always leads to another M&P channel as can be seen from Eqs. (7.10) to (7.12). Thus, we can restrict ourselves to this type of channels. For these channels, the eigenvalues can be evaluated analytically and maximization of the bounds over these channels for Q_- results in the channel

$$\mathcal{M}_-(\rho) = \frac{1}{2}\left[\mathbb{1} + \frac{1}{\sqrt{5}}\left(\sigma_x + \mathrm{Tr}(\rho\sigma_y)\sigma_y + \mathrm{Tr}(\rho\sigma_z)\sigma_z\right)\right]. \tag{7.13}$$

It is visualized in the Bloch picture in Fig. 7.2 and has a quality of $Q_-(\mathcal{M}_-) = \frac{1}{\sqrt{5}}$. For Q_+, the optimization of the bounds over the channels yields

$$\mathcal{M}_+(\rho) = \frac{1}{2}\left[\mathbb{1} + \frac{1}{\sqrt{2}}(\sigma_x + \mathrm{Tr}(\rho\sigma_z)\sigma_z)\right], \tag{7.14}$$

with $Q_0(\mathcal{M}_+) = Q_+(\mathcal{M}_+) = \frac{1}{\sqrt{2}}$. The channel is visualized in Fig. 7.2.

For unital channels, i.e. $\vec{\kappa} = 0$, the condition for separability reads $\sum_i \lambda_i \leq 1$ [69]. Maximizing under this constraint yields for Q_- the depolarizing channel

$$\mathcal{M}'_-(\rho) = \frac{1}{3}\rho + \frac{1}{3}\mathbb{1} \tag{7.15}$$

with $Q_-(\mathcal{M}'_-) = \frac{1}{3}$. For Q_0 and Q_+, we obtain the planar channel

$$\mathcal{M}'_+(\rho) = \frac{1}{2}\left[\mathbb{1} + \frac{1}{2}(\mathrm{Tr}(\rho\sigma_y)\sigma_y + \mathrm{Tr}(\rho\sigma_z)\sigma_z)\right] \tag{7.16}$$

with $Q_0(\mathcal{M}'_+) = Q_+(\mathcal{M}'_+) = \frac{1}{2}$. □

Finally, we have a statement similar to Lemma 7.6 for Q_- if the channel is unital:

Lemma 7.10. *Let $\mathcal{M}$ and $\mathcal{N}$ be unital channels acting on single qubits ($D = 2$). Then, it holds that $Q_-(\mathcal{M} \circ \mathcal{N}) \leq Q_-(\mathcal{M})$.*

Proof. First, note that the composition of unital channels is again a unital channel. As shown in Lemma 7.8, the quality measure $Q_-(\mathcal{M})$ for a unital channel $\mathcal{M}$ is given by the minimal singular value of the matrix $\Lambda_\mathcal{M}$, i.e. $\lambda_1(\Lambda_\mathcal{M})$. With this, we have that

$$\begin{aligned} Q_-(\mathcal{M} \circ \mathcal{N}) &= \lambda_1(\Lambda_{\mathcal{M}\circ\mathcal{N}}) \\ &\leq \lambda_1(\Lambda_\mathcal{M})\lambda_3(\Lambda_\mathcal{N}) \\ &\leq \lambda_1(\Lambda_\mathcal{M}) = Q_-(\mathcal{M}). \end{aligned} \tag{7.17}$$

For the first inequality, we used that $\Lambda_{\mathcal{M}\circ\mathcal{N}} = \Lambda_\mathcal{M}\Lambda_\mathcal{N}$ and Theorem 3.3.16 from Ref. [274]. The second inequality follows from the fact that for channels, all the singular values of the matrix Λ have to be smaller or equal to 1. □

To illustrate how the measures can be determined for specific single-qubit channels, we examine several well-known channels.

7.6 Examples of single-qubit channels

In the following, we will consider the phase-flip, the amplitude-damping and the depolarizing channel and derive their quality in terms of Q_0 and $Q_\pm$.

– *The phase-flip channel $\mathcal{P}$:* The matrix Λ for the unital (i.e., $\vec{\kappa} = 0$) phase-flip channel $\mathcal{P}$, is given by $\mathrm{diag}(1-p, 1-p, 1)$ with $0 \leq p \leq 1$. It can be realized by a M&P scheme for $p = 1$ only. Using the result from Lemma 7.8 for unital channels, we have that

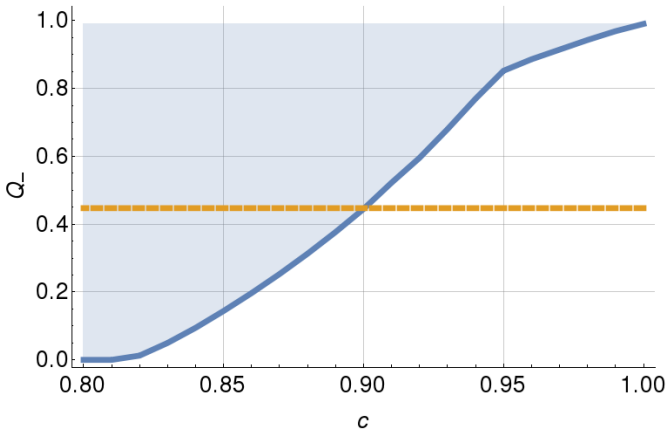


FIGURE 7.3: [29] Lower bound (solid blue line) and allowed values above this bound (in blue) for the quality measure Q_-, given that in certain directions a coherence of at least c is measured. The upper bound for M&P channel of $\frac{1}{\sqrt{5}}$ is displayed by a dashed orange line.

$Q_-(\mathcal{P}) = Q_0(\mathcal{P}) = Q_+(\mathcal{P}) = 1-p$. It should be noted that any bit-flip or bit-phase-flip channel is related to a phase-flip channel with the same error probability p via a transformation of the form $\mathcal{V} \circ \mathcal{P} \circ \mathcal{V}^{-1}$ where $\mathcal{V}(\rho) = V\rho V^\dagger$ is a unitary channel. Hence, the quality measures $Q_\pm$ and Q_0 for these channels with same error probability coincide. Note that Q_- excludes unital M&P schemes for $p < \frac{2}{3}$, while Q_+ and Q_0 exclude them for $p < \frac{1}{2}$.

– *The amplitude-damping channel $\mathcal{A}$:* The matrix Λ for the amplitude-damping channel $\mathcal{A}$ is given by $\mathrm{diag}(\sqrt{1-p}, \sqrt{1-p}, 1-p)$ and $\vec{\kappa} = (0,0,p)^T$, where $0 \leq p \leq 1$. This channel can again be implemented by M&P schemes only if $p = 1$. Considering the maximal coherence of the states in the image of this channel with respect to the computational basis shows that $Q_+(\mathcal{A}) \leq \sqrt{1-p}$. Using that $\lambda_1 \leq Q_- \leq Q_0 \leq Q_+$ leads to $Q_- = Q_0 = Q_+ = \sqrt{1-p}$. Thus, Q_- excludes M&P schemes for $p < \frac{4}{5}$, whereas Q_+ and Q_0 exclude them for $p < \frac{1}{2}$.

– *The depolarizing channel $\mathcal{D}$:* The matrix Λ for the unital depolarizing channel $\mathcal{D}$ is given by $\mathrm{diag}(p,p,p)$, where $0 \leq p \leq 1$. This channel is a M&P channel only if $p \leq \frac{1}{3}$. Because of symmetry, it is clear that $Q_- = Q_0 = Q_+ = p$. Thus, Q_- certifies the full range of non-M&P channels if it is known that the channel is unital, while Q_0 and Q_+ exclude M&P schemes in the case of $p > \frac{1}{2}$.

7.7 Experimental estimation of the quality of a quantum memory

In this section, we explain how to determine a lower bound on the quality measures from experimental data for qubit systems for channels close to the identity channel. This situation is of major interest as a perfect quantum memory corresponds to the identity channel.

Obviously, it is possible to obtain (lower bounds on) the quality measures by performing process tomography of the channel and then using the obtained characterization. However, process tomography requires the ability to prepare a set of input states with high precision as well as many well characterized measurements [260, 261]. Here, we only assume that one can prepare three different states $\{\rho_i\}_{i=1}^3$ such that for the output states one can certify a lower bound $c_i \in [0,1]$ on the following coherences

$$\begin{aligned} C_{U_x}[\mathcal{M}(\rho_1)] \geq c_1, \quad C_{U_y}[\mathcal{M}(\rho_1)] \geq c_1, \\ C_{U_x}[\mathcal{M}(\rho_2)] \geq c_2, \quad C_{U_z}[\mathcal{M}(\rho_2)] \geq c_2, \\ C_{U_y}[\mathcal{M}(\rho_3)] \geq c_3, \quad C_{U_z}[\mathcal{M}(\rho_3)] \geq c_3, \end{aligned} \tag{7.18}$$

where the U_j correspond to the usual x, y and z direction on the Bloch sphere (i.e., $U_j = e^{i\sigma_j \pi/4}$ for $j = x, y, z$). This can for instance be achieved using the method from Ref. [275]. If the input states are chosen carefully and the channel is close enough to a unitary transformation, it suffices to only conduct three measurements in total. These measurements certify that there are states close to the eigenstates of the Pauli matrices in the image of $\mathcal{M}$. Furthermore, we only assume a bound on the coherence of the output of the quantum memory, nothing additional is assumed on the input- or output states.

For simplicity, we consider the case where $c := c_1 = c_2 = c_3$. As the smallest semi-axis is a lower bound on Q_-, one can determine the channel that shows the smallest possible λ_1 compatible with the observed data. In particular, it is required that the image of the channel contains states for which the bounds given in Eqs. (7.18) are fulfilled. For $c > \sqrt{\frac{2}{3}} \approx 0.82$, there must be at least three different states close to the boundary of the Bloch sphere. Numerically optimizing over all compatible channels leads to the lower bounds depicted in Fig. 7.3. Hence, for values of $c \gtrsim 0.82$ it is possible to obtain non-trivial lower bounds on the quality measure Q_- (and hence, also on Q_0 and Q_+) by having access only to a few lower bounds on the coherences of three different states. M&P channels can be excluded with certainty if $Q_- > \frac{1}{\sqrt{5}} \approx 0.45$, which is given for $c \gtrsim 0.9$. As an example, consider the amplitude-damping channel $\mathcal{A}$ from

above. One can find states for which $c = \sqrt{1-p}$, and thus exclude M&P channels for $p \lesssim 0.19$.

For higher dimensional channels, the estimation is more involved, however, notable results can still be obtained [18, 29].

7.8 Conclusion

We introduced a physically motivated measure Q_0 that characterizes quantum memories by their ability to preserve coherence and fulfills all the desirable properties for such a quantifier. In particular, entanglement-breaking channels offer only restricted quality as quantified by the coherence-based performance measure Q_0, revealing an insightful connection between the quantum phenomena coherence and entanglement. For a single-qubit quantum memory, the measure can be evaluated for many scenarios, even if only restricted experimental data is available. In contrast to full process tomography, our scheme does not require the precise preparation of states but only demands the certification of (sufficiently high) lower bounds on certain coherences of three unknown states.

For future work, it is desirable to extend the method to characterize and verify other basic elements of quantum information processing. A simple extension is the case of quantum teleportation, where the results can be applied directly. More interesting is an application to two-qubit gates. The fact that a two-qubit gate generates entanglement, can be seen as the property that a certain two-level coherence increases [267, 268]. In this sense, our method may be extended to characterize the entangling capability of multi-qubit quantum gates.

"Die Quanten sind doch eine hoffnungslose Schweinerei!"
Max Born

Concluding remarks

This thesis is devoted to deepening the understanding of the interplay between convex optimization, coherence, and quantum entanglement. During this venture we touched upon various areas of research in quantum information and optimization theory such as semidefinite programming, symmetries in optimization problems, the marginal problem, absolutely maximally entangled states, quantum error-correcting codes, entanglement and its detection in experiments, entropic uncertainty relations, the joint numerical range of observables, quantum memories, and last but not least, quantum coherence. We found new connections between seemingly independent problems. For instance, we rephrased rank-constrained optimization problems as hierarchies of semidefinite programs using entanglement theory, transformed the existence problem of AME states to a separability problem, contributed to the theory of entanglement detection in different practical scenarios, and quantified the performance of quantum memories via their ability to preserve coherence.

After having introduced the mathematical fundamentals, we first described a quantum-inspired hierarchy for rank-constrained optimization. We used the purification of mixed quantum states and the well-known fact that the minimal dimension of the added ancilla system is exactly the rank of the purified state to rephrase the optimization over rank-constrained matrices as an optimization over pure quantum states of an enlarged system. The constraint that the considered state is pure is still difficult to ensure, however, we characterized it via separable two-copy states in the symmetric subspace. Then, inspired by the DPS hierarchy, we transformed the optimization over the separable cone to a hierarchy of, in principle efficiently solvable, semidefinite programs implementing N-copy quantum states. Furthermore, we applied the method to several problem instances from quantum information and computer science, namely, the optimization over pure quantum states and unitary channels, orthonormal representations and the maximum-cut problem of graphs, as well as pseudo-boolean optimization. Finally, we provide some insights into the quantum de Finetti theorem, especially, considering the uniqueness question of multi-copy decompositions.

Second, special attention is given to the use of our method in the context of the pure-state marginal problem in quantum mechanics, in particular, considering the existence of absolutely maximally entangled states and quantum error-correcting codes. We utilized unitary and permutation symmetries to prove that the existence of n-partite locally d-dimensional AME states is equivalent to the bipartite separability of a certain quantum state of $2n$ particles. Moreover, we derived an explicit expression for that state and recover many known existence results via the positivity and PPT constraints which become easily computable in our framework. The symmetries also allowed us to compute high orders of our hierarchy, which will be useful when investigating the existence of quantum codes which we transformed to a marginal problem as well making our method applicable.

Third, we investigated an entanglement detection scenario which we termed scrambled data. Here, the mapping of measurement outcomes to the corresponding outcome probabilites is lost. Challenges of this type might occur in experimental setups where different outcomes cannot be distinguished. It turned out that the state space of in principle detectable states is nonconvex implying that their detection is involved. In particular, ordinary entanglement witnesses are insufficient to detect all such states, however, we found scrambling-invariant families of entanglement witnesses as well as entropic uncertainty relations that prove useful for entanglement detection in this scenario.

Fourth, we examined the volume ratio of the joint separable compared to the general numerical range of various observables. A small volume ratio generically allows for entanglement detection even in the case of relatively large experimental confidence regions due to statistical and systematic errors. Thus, we found general bounds and investigated the case of qubits in more detail. In particular, we explicitly computed both the separable and general numerical range for two locally traceless two-qubit observables. Furthermore, we described six locally noncommuting qutrit-qutrit observables whose measurements do not allow to detect any entanglement, and evidence that this might be the maximal number of such observables indicates a new connecting between entanglement and commutativity.

Finally, we introduced a quality measure that quantifies the performance of quantum memories based on their ability to preserve the coherence of the stored quantum state. Since quantum memories are just as important for future universal quantum computers as classical memories for their ordinary counterparts, the certification of desirable properties is essential in practical implementations. We identified indispensible features and showed that our performance measure satisfies these requirements. For instance, the measure distinguishes between entanglement-breaking channels that

merely store classical information and unitary transformations that can be corrected or accounted for in a quantum computer. Moreover, we discussed single-qubit channels in more detail, finding several bounds and a simple measurement scheme to approximate our quality measure.

We hope that our findings not only advance the knowledge in the various fields of research considered, but also facilitate the application of tools from one area to another via the newly established connections. Specifically, intriguing open questions include applying our hierarchy for rank-constrained optimization to specific instances of the existence problem of quantum error-correcting codes, finding further applications in and beyond quantum information theory, and investigating further the connection between entanglement detection and commutativity.

Acknowledgments

First and foremost, I want to thank my Doktorvater, Otfried Gühne, who not only gave me the chance to study in his group as a PhD student, but also made an extraordinary job at supervising me over the years. In particular, I admire his patience to sit down with his students and simply calculate, almost always coming up with new ideas worth investigating when I was stuck with a seemingly insurmountable problem. Thank you.

Also, former and current members of our working group contributed to the best research environment possible. Special thanks go to Xiao-Dong Yu and Nikolai Wyderka with whom I worked the most during my quantum information venture. Without their invaluable ideas and fruitful collaboration, this thesis would not have been possible. Thanks are also due to all my coauthors, office mates, and all other colleagues that became friends during my studies. Although I don't want to fill this page solely with names, I have to mention Mariami Gachechiladze and our uncountable jogging rounds around the Häusling. I really miss having the lunch breaks all together due to the corona virus, and I even miss everyone asking if I also want a cup of coffee even though, up to this day, I still haven't acquired a taste for it.

Deepest gratitude are due to Daniela Lehmann for always being happy and cheerful while organizing everything and finding unusual ways around the university's madness of bureaucracy.

The quantum information group at the Jagiellonian university in Cracow deserves my sincerest thanks for their hospitality during my visit in the middle of the pandemic. In particular, I want to thank Karol Życzkowski, Jakub Czartowski and Konrad Szymański for our rewarding collaboration, and Adam Burchardt for his bike and interesting discussions.

I acknowledge the financial and non-financial support by the House of Young Talents Siegen.

Besides, learning Chinese would not have been so much fun without the support of my chinese colleagues Zhen-Peng Xu, Yuanyuan Mao, and Xiao-Dong Yu. I really appreciate you for not being frustrated because of my terrible pronunciation.

Last but not least, I want to thank my family, my love, and my friends for their constant moral support and that they never left me by myself when a sense of despair was about to overwhelm me.

List of publications

(A) Timo Simnacher, Nikolai Wyderka, René Schwonnek, and Otfried Gühne
Entanglement detection with scrambled data
Phys. Rev. A **99**, 062339 (2019)

(B) Timo Simnacher, Nikolai Wyderka, Cornelia Spee, Xiao-Dong Yu, and Otfried Gühne
Certifying quantum memories with coherence
Phys. Rev. A **99**, 062319 (2019)

(C) Xiao-Dong Yu, Timo Simnacher, Nikolai Wyderka, H. Chau Nguyen, and Otfried Gühne
A complete hierarchy for the pure state marginal problem in quantum mechanics
Nature Communications **12**, 1012 (2021)

(D) Xiao-Dong Yu, Timo Simnacher, H. Chau Nguyen, and Otfried Gühne
Quantum-inspired hierarchy for rank-constrained optimization
arXiv:2012.00554

(E) Timo Simnacher, Jakub Czartowski, Konrad Szymański, and Karol Życzkowski
Confident entanglement detection via (separable) numerical range
arXiv:2107.04365

List of Figures and Tables

List of Figures

List of Tables

A Projectors onto invariant subspaces of partially transposed fourpartite permutation matrices

Let $\boldsymbol{V}^{T_P}$ be the vector of partially transposed fourpartite permutation matrices with respect to the subsystems P in the canonical order

$$\boldsymbol{V}^{T_P} = \left(V_{1234}^{T_P}, V_{1243}^{T_P}, V_{1324}^{T_P}, V_{1342}^{T_P}, V_{1423}^{T_P}, \dots\right)^T, \tag{A.1}$$

where $V_{xyzw} = \sum_{i_1 i_2 i_3 i_4} |i_x i_y i_z i_w\rangle \langle i_1 i_2 i_3 i_4|$.

First, we consider the partial transposition w.r.t. the first subsystem. For local dimension $d = 2$, the projectors onto the invariant subspaces are given by

$$\begin{aligned} S_0^A = \Big(&\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, \\ &-\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}, -\frac{1}{24}\Big)^T \cdot \boldsymbol{V}^{T_A}, \end{aligned} \tag{A.2}$$

$$\begin{aligned} S_1^A = \Big(&\frac{1}{6}, -\frac{1}{6}, -\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{6}, \frac{1}{9}, -\frac{1}{9}, \frac{1}{18}, -\frac{1}{18}, -\frac{1}{18}, \frac{1}{18}, \frac{1}{18}, -\frac{1}{18}, \frac{1}{9}, \frac{1}{18}, -\frac{1}{9}, \\ &-\frac{1}{18}, -\frac{1}{18}, \frac{1}{18}, \frac{1}{18}, \frac{1}{9}, -\frac{1}{18}, -\frac{1}{9}\Big)^T \cdot \boldsymbol{V}^{T_A}, \end{aligned} \tag{A.3}$$

$$\begin{aligned} S_2^A = \Big(&0, 0, 0, 0, 0, \frac{3}{8}, \frac{3}{8}, -\frac{1}{8}, -\frac{1}{8}, -\frac{1}{8}, -\frac{1}{8}, -\frac{1}{8}, -\frac{1}{8}, -\frac{1}{8}, \frac{3}{8}, -\frac{1}{8}, \frac{3}{8}, -\frac{1}{8}, -\frac{1}{8}, \\ &-\frac{1}{8}, -\frac{1}{8}, \frac{3}{8}, -\frac{1}{8}, \frac{3}{8}\Big)^T \cdot \boldsymbol{V}^{T_A}, \end{aligned} \tag{A.4}$$

$$\begin{aligned} S_3^A = \Big(&\frac{2}{3}, 0, 0, -\frac{1}{3}, -\frac{1}{3}, 0, -\frac{4}{9}, -\frac{2}{9}, \frac{1}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}, \frac{2}{9}, -\frac{4}{9}, \frac{1}{9}, -\frac{2}{9}, \frac{2}{9}, \frac{2}{9}, \\ &\frac{1}{9}, \frac{1}{9}, -\frac{4}{9}, \frac{2}{9}, -\frac{2}{9}\Big)^T \cdot \boldsymbol{V}^{T_A}. \end{aligned} \tag{A.5}$$

For $d \geq 3$, they are given by

$$\begin{aligned} S_0^A = \Big(&\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, \ -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, \\ &-\alpha_+, -\alpha_+, \ -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+, -\alpha_+\big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_A}, \end{aligned} \tag{A.6}$$

$$\begin{aligned} S_1^A = \Big(&\frac{1}{6}, -\frac{1}{6}, -\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{6}, -\alpha_-, \alpha_-, \alpha_-, -\alpha_-, \ -\alpha_-, \alpha_-, \alpha_-, -\alpha_-, -\alpha_-, \\ &\alpha_-, \alpha_-, -\alpha_-, -\alpha_-, \ \alpha_-, \alpha_-, -\alpha_-, -\alpha_-, \alpha_-\big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_A}, \end{aligned} \tag{A.7}$$

$$\begin{aligned} S_{2a}^A = (&0, 0, 0, 0, 0, 0, \beta_+, \beta_+, -\gamma_+, -\gamma_+, -\gamma_+, -\gamma_+, \ -\gamma_+, -\gamma_+, \beta_+, -\gamma_+, \\ &\beta_+, -\gamma_+, -\gamma_+, -\gamma_+, \ -\gamma_+, \beta_+, -\gamma_+, \beta_+)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_A}, \end{aligned} \tag{A.8}$$

$$\begin{aligned} S_{2b}^A = (&0, 0, 0, 0, 0, 0, \beta_-, -\beta_-, -\gamma_-, \gamma_-, \gamma_-, -\gamma_-, \ -\gamma_-, \gamma_-, \beta_-, -\gamma_-, -\beta_-, \\ &\gamma_-, \gamma_-, -\gamma_-, -\gamma_-, \ \beta_-, \gamma_-, -\beta_-)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_A}, \end{aligned} \tag{A.9}$$

$$\begin{aligned} S_3^A = \Big(&\frac{2}{3}, 0, 0, -\frac{1}{3}, -\frac{1}{3}, 0, -2d\delta, -2\delta, \delta, d\delta, d\delta, \ \delta, \delta, d\delta, -2d\delta, \delta, -2\delta, \\ &d\delta, d\delta, \delta, \delta, \ -2d\delta, d\delta, -2\delta)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_A}, \end{aligned} \tag{A.10}$$

where

$$\alpha_\pm = \frac{1}{6(d \pm 2)}, \tag{A.11}$$

$$\beta_\pm = \frac{d \pm 1}{2(d \pm 2)(d \mp 1)} = (d \pm 1)\gamma_\pm, \tag{A.12}$$

$$\delta = \frac{1}{3(d^2 - 1)}. \tag{A.13}$$

Second, we compute the projectors onto the invariant subspaces for the partial transposition w.r.t. the first two subsystems. In this case, there is no distinction for different local dimensions. With $\eta = 1/(d^4 - 1)$, the projectors are given by

$$\begin{aligned} S_0^{AB} = \Big(&\frac{1}{4}, \frac{1}{4}, -\frac{1}{4d}, -\frac{1}{4d}, -\frac{1}{4d}, -\frac{1}{4d}, -\frac{1}{4}, -\frac{1}{4}, \frac{1}{4d}, \ \frac{1}{4d}, \frac{1}{4d}, \frac{1}{4d}, \frac{1}{4d}, \frac{1}{4d}, \\ &-\frac{1}{4d}, -\frac{1}{4d}, 0, 0, \frac{1}{4d}, \frac{1}{4d}, \ -\frac{1}{4d}, -\frac{1}{4d}, 0, 0\Big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_{AB}}, \end{aligned} \tag{A.14}$$

$$\begin{aligned} S_1^{AB} = \Big(&\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4d}, \frac{1}{4d}, \frac{1}{4d}, -\frac{1}{4d}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4d}, \ \frac{1}{4d}, \frac{1}{4d}, -\frac{1}{4d}, -\frac{1}{4d}, \\ &\frac{1}{4d}, -\frac{1}{4d}, \frac{1}{4d}, 0, 0, \frac{1}{4d}, \ -\frac{1}{4d}, \frac{1}{4d}, -\frac{1}{4d}, 0, 0\Big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_{AB}}, \end{aligned} \tag{A.15}$$

$$S_2^{AB} = \big(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, d^2\eta, -\eta, \ 0, 0, 0, 0, -\eta, d^2\eta\big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_{AB}}, \tag{A.16}$$

$$\begin{aligned} S_3^{AB} = \Big(&\frac{1}{2}, 0, \frac{1}{2d}, 0, 0, \frac{1}{2d}, 0, \frac{1}{2}, 0, -\frac{1}{2d}, -\frac{1}{2d}, 0, 0, -\frac{1}{2d}, \ \frac{1}{2d}, \\ &0, 0, 0, -\frac{1}{2d}, 0, 0, \frac{1}{2d}, 0, 0\Big)^T \cdot \boldsymbol{V}^{\boldsymbol{T}_{AB}}. \end{aligned} \tag{A.17}$$

Bibliography

[1] M. Planck, "Über das Gesetz der Energieverteilung im Normalspectrum", Ann. Phys. (Berl.) **309**, 553 (1901).

[2] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", Ann. Phys. (Berl.) **322**, 132 (1905).

[3] C. Marletto and V. Vedral, "Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity", Phys. Rev. Lett. **119**, 240402 (2017).

[4] R. S. Ingarden, "Quantum information theory", Rep. Math. Phys. **10**, 43 (1976).

[5] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?", Phys. Rev. **47**, 777 (1935).

[6] J. S. Bell, "On the einstein podolsky rosen paradox", Physics Physique Fizika **1**, 195 (1964).

[7] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, C. Abellán, *et al.*, "Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres", Nature **526**, 682 (2015).

[8] M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, *et al.*, "Significant-loophole-free test of bell's theorem with entangled photons", Phys. Rev. Lett. **115**, 250401 (2015).

[9] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, *et al.*, "Strong loophole-free test of local realism", Phys. Rev. Lett. **115**, 250402 (2015).

[10] R. P. Feynman, "Simulating physics with computers", Int. J. Theor. Phys **21** (1982).

[11] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", SIAM review **41**, 303 (1999).

[12] S. Wiesner, "Conjugate coding", ACM Sigact News **15**, 78 (1983).

[13] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing", Proc. of IEEE International Conference on Computers, Systems and Signal Processing **175**, 8 (1984).

[14] A. K. Ekert, "Quantum cryptography based on Bell's theorem", Phys. Rev. Lett. **67**, 661 (1991).

[15] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, "MIP*=RE" (2020), arXiv:2001.04383 [quant-ph] .

[16] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement", Rev. Mod. Phys. **81**, 865 (2009).

[17] F. Huber and N. Wyderka, "Table of AME states", `http://www.tp.nt.uni-siegen.de/+fhuber/ame.html`.

[18] N. Wyderka, *Learning from correlations: What parts of quantum states tell about the whole*, Ph.D. thesis, Universität Siegen (2020).

[19] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (Cambridge University Press, 2002).

[20] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: An introduction to quantum entanglement* (Cambridge University Press, 2017).

[21] L. Landau, "Das Dämpfungsproblem in der Wellenmechanik", Zeitschrift für Physik **45**, 430 (1927).

[22] J. von Neumann, "Thermodynamik quantenmechanischer Gesamtheiten", Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1927**, 273 (1927).

[23] I. Gelfand and M. Neumark, "On the imbedding of normed rings into the ring of operators in Hilbert space", Matematicheskii Sbornik **12**, 197 (1943).

[24] W. Gerlach and O. Stern, "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld", Zeitschrift für Physik **9**, 349 (1922).

[25] R. Shankar, *Principles of quantum mechanics* (Springer Science & Business Media, 2012).

[26] J. J. Sakurai and E. D. Commins, *Modern quantum mechanics* (American Association of Physics Teachers, 1995).

[27] M. M. Wolf, "Quantum channels & operations, guided tour", Lecture notes, Technische Universität München, `https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf` (2012).

[28] W. F. Stinespring, "Positive functions on C*-algebras", Proc. of the American Mathematical Society **6**, 211 (1955).

[29] T. Simnacher, N. Wyderka, C. Spee, X.-D. Yu, and O. Gühne, "Certifying quantum memories with coherence", Phys. Rev. A **99**, 062319 (2019).

[30] J. de Pillis, "Linear transformations which preserve Hermitian and positive semidefinite operators", Pac. J. Math. **23**, 129 (1967).

[31] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators", Rep. Math. Phys. **3**, 275 (1972).

[32] M.-D. Choi, "Completely positive linear maps on complex matrices", Linear Algebra Its Appl. **10**, 285 (1975).

[33] K. Kraus, "General state changes in quantum theory", Annals of Physics **64**, 311 (1971).

[34] J. Aberg, "Quantifying superposition" (2006), arXiv:0612146 [quant-ph] .

[35] T. Baumgratz, M. Cramer, and M. B. Plenio, "Quantifying coherence", Phys. Rev. Lett. **113**, 140401 (2014).

[36] A. Winter and D. Yang, "Operational resource theory of coherence", Phys. Rev. Lett. **116**, 120404 (2016).

[37] F. G. Brandao and G. Gour, "Reversible framework for quantum resource theories", Physical review letters **115**, 070503 (2015).

[38] E. Chitambar and G. Gour, "Quantum resource theories", Rev. Mod. Phys. **91**, 025001 (2019).

[39] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying entanglement", Phys. Rev. Lett. **78**, 2275 (1997).

[40] C. Eltschka and J. Siewert, "Quantifying entanglement resources", J. Phys. A: Math. Theor. **47**, 424005 (2014).

[41] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality", Rev. Mod. Phys. **86**, 419 (2014).

[42] J. I. de Vicente, "On nonlocality as a resource theory and nonlocality measures", J. Phys. A: Math. Theor. **47**, 424017 (2014).

[43] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, "Quantifying Bell: The resource theory of nonclassicality of common-cause boxes", Quantum **4**, 280 (2020).

[44] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, "Robustness of asymmetry and coherence of quantum states", Phys. Rev. A **93**, 042107 (2016).

[45] Y. Peng, Y. Jiang, and H. Fan, "Maximally coherent states and coherence-preserving operations", Phys. Rev. A **93**, 032326 (2016).

[46] Z. Bai and S. Du, "Maximally coherent states", Quantum Inf. Comput. **15**, 1355 (2015).

[47] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, "Robustness of coherence: An operational and observable measure of quantum coherence", Phys. Rev. Lett. **116**, 150502 (2016).

[48] T. S. Cubitt, F. Verstraete, W. Dür, and J. I. Cirac, "Separable states can be used to distribute entanglement", Phys. Rev. Lett. **91**, 037902 (2003).

[49] E. Schrödinger, "Probability relations between separated systems", Math. Proc. of the Cambridge Philosophical Society **32**, 446–452 (1936).

[50] L. P. Hughston, R. Jozsa, and W. K. Wootters, "A complete classification of quantum ensembles having a given density matrix", Phys. Lett. A **183**, 14 (1993).

[51] N. Hadjisavvas, "Properties of mixtures on non-orthogonal states", Lett. Math. Phys. **5**, 327 (1981).

[52] O. Gühne and G. Tóth, "Entanglement detection", Phys. Rep. **474**, 1 (2009).

[53] G. Svetlichny, "Distinguishing three-body from two-body nonseparability by a Bell-type inequality", Phys. Rev. D **35**, 3066 (1987).

[54] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going beyond Bell's theorem", in *Bell's theorem, quantum theory and conceptions of the universe* (Springer, 1989) pp. 69–72.

[55] A. Zeilinger, M. A. Horne, and D. M. Greenberger, "Higher-order quantum entanglement", in *NASA Conf. Publ*, Vol. 3135 (1992) pp. 73–81.

[56] W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways", Phys. Rev. A **62**, 062314 (2000).

[57] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: Necessary and sufficient conditions", Phys. Lett. A **223**, 1 (1996).

[58] A. Peres, "Separability criterion for density matrices", Phys. Rev. Lett. **77**, 1413 (1996).

[59] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition", Phys. Lett. A **232**, 333 (1997).

[60] L. Gurvits, "Classical deterministic complexity of Edmonds' Problem and quantum entanglement", in *Proc. of the thirty-fifth annual ACM symposium on theory of computing* (ACM, 2003) p. 10.

[61] S. Gharibian, "Strong NP-hardness of the quantum separability problem", Quantum Inf. Comput. **10**, 343 (2010).

[62] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, "Distinguishing separable and entangled states", Phys. Rev. Lett. **88**, 187904 (2002).

[63] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, "Complete family of separability criteria", Phys. Rev. A **69**, 022308 (2004).

[64] E. Chitambar, "Local quantum transformations requiring infinite rounds of classical communication", Phys. Rev. Lett. **107**, 190502 (2011).

[65] F. Mintert, A. R. Carvalho, M. Kuś, and A. Buchleitner, "Measures and dynamics of entangled states", Phys. Rep. **415**, 207 (2005).

[66] M. B. Plenio and S. S. Virmani, "Going beyond Bell's theorem", in *Quantum information and coherence* (Springer, 2014) pp. 173–209.

[67] M. A. Nielsen, "Conditions for a class of entanglement transformations", Phys. Rev. Lett. **83**, 436 (1999).

[68] A. S. Holevo, "Coding theorems for quantum channels", Russian Math. Surveys **53**, 1295 (1998).

[69] M. Horodecki, P. W. Shor, and M. B. Ruskai, "Entanglement breaking channels", Rev. Math. Phys. **15**, 629 (2003).

[70] M. Christandl, A. Müller-Hermes, and M. M. Wolf, "When do composed maps become entanglement breaking?", in *Annales Henri Poincaré*, Vol. 20 (Springer, 2019) pp. 2295–2322.

[71] L. Chen, Y. Yang, and W.-S. Tang, "Positive-partial-transpose square conjecture for $n = 3$", Phys. Rev. A **99**, 012337 (2019).

[72] S. Singh and I. Nechita, "The PPT^2 conjecture holds for all Choi-type maps" (2006), arXiv:2011.03809 [quant-ph] .

[73] F. Hausdorff, "Der Wertvorrat einer Bilinearform", Mathematische Zeitschrift **3**, 314 (1919).

[74] O. Toeplitz, "Das algebraische Analogon zu einem Satze von Fejér", Mathematische Zeitschrift **2**, 187 (1918).

[75] P. Gawron, Z. Puchała, J. A. Miszczak, Ł. Skowronek, and K. Życzkowski, "Restricted numerical range: A versatile tool in the theory of quantum information", J. Math. Phys. **51**, 102204 (2010).

[76] F. D. Murnaghan, "On the field of values of a square matrix", Proc. of the National Academy of Sciences of the United States of America **18**, 246 (1932).

[77] R. Kippenhahn, "Über den Wertevorrat einer Matrix", Mathematische Nachrichten **6**, 193 (1951).

[78] M. Fiedler, "Geometry of the numerical range of matrices", Linear Algebra Its Appl. **37**, 81 (1981).

[79] A. Klyachko, "Quantum marginal problem and representations of the symmetric group" (2004), arXiv:0409113 [quant-ph] .

[80] W. Helwig, W. Cui, J. I. Latorre, A. Riera, and H.-K. Lo, "Absolute maximal entanglement and quantum secret sharing", Phys. Rev. A **86**, 052335 (2012).

[81] W. Helwig and W. Cui, "Absolutely maximally entangled states: Existence and applications" (2013), arXiv:1306.2536 [quant-ph] .

[82] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned", Nature **299**, 802 (1982).

[83] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory", Phys. Rev. A **52**, R2493 (1995).

[84] A. Steane, "Multiple-particle interference and quantum error correction", Proc. of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **452**, 2551 (1996).

[85] E. Knill and R. Laflamme, "Theory of quantum error-correcting codes", Phys. Rev. A **55**, 900 (1997).

[86] A. J. Scott, "Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions", Phys. Rev. A **69**, 052330 (2004).

[87] E. M. Rains, "Nonbinary quantum codes", IEEE Transactions on Information Theory **45**, 1827 (1999).

[88] L. G. Khachiyan, "A polynomial algorithm in linear programming", in *Doklady Akademii Nauk*, Vol. 244 (Russian Academy of Sciences, 1979) pp. 1093–1096.

[89] S. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).

[90] L. Vandenberghe and S. Boyd, "Semidefinite programming", SIAM Rev. **38**, 49 (1996).

[91] M. Slater, "Lagrange multipliers revisited", in *Traces and emergence of nonlinear programming* (Springer, 2014) p. 293.

[92] R. M. Freund, "Introduction to semidefinite programming", Lecture notes, Massachusetts Institute of Technology, `https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-251j-introduction-to-mathematical-programming-fall-2009/readings/MIT6_251JF09_SDP.pdf` (2009).

[93] C. E. Shannon, "A mathematical theory of communication", Bell Syst. tech. j. **27**, 379 (1948).

[94] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", Z. Phys. **43**, 172 (1927).

[95] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, "Entropic uncertainty relations and their applications", Rev. Mod. Phys. **89**, 015002 (2017).

[96] H. Maassen and J. B. Uffink, "Generalized entropic uncertainty relations", Phys. Rev. Lett. **60**, 1103 (1988).

[97] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics", J. Stat. Phys. **52**, 479 (1988).

[98] A. Rényi, "On measures of information and entropy", Proc. of the fourth Berkeley Symposium on Mathematics, Statistics and Probability , 547– (1960).

[99] X.-D. Yu, T. Simnacher, H. C. Nguyen, and O. Gühne, "Quantum-inspired hierarchy for rank-constrained optimization" (2020), arXiv:2012.00554 [quant-ph] .

[100] M. Navascués, S. Pironio, and A. Acín, "Bounding the set of quantum correlations", Phys. Rev. Lett. **98**, 010401 (2007).

[101] T. Barthel and R. Hübener, "Solving condensed-matter ground-state problems by semidefinite relaxations", Phys. Rev. Lett. **108**, 200404 (2012).

[102] D. Simmons-Duffin, "A semidefinite program solver for the conformal bootstrap", J. High Energy Phys. **2015**, 174.

[103] L. Lovász, "On the Shannon capacity of a graph", IEEE Trans. Inf. Theory **25**, 1 (1979).

[104] J. B. Lasserre, "Global optimization with polynomials and the problem of moments", SIAM J. Optim. **11**, 796 (2001).

[105] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, Ph.D. thesis, California Institute of Technology (2000).

[106] M. Navascués, S. Pironio, and A. Acín, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations", New J. Phys. **10**, 073013 (2008).

[107] M. Navascués and T. Vértesi, "Bounding the set of finite dimensional quantum correlations", Phys. Rev. Lett. **115**, 020501 (2015).

[108] O. Gühne, Y. Mao, and X.-D. Yu, "Geometry of faithful entanglement", Phys. Rev. Lett. **126**, 140503 (2021).

[109] F. Barahona, M. Grötschel, M. Jünger, and G. Reinelt, "An application of combinatorial optimization to statistical physics and circuit layout design", Oper. Res. **36**, 493 (1988).

[110] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum state tomography via compressed sensing", Phys. Rev. Lett. **105**, 150401 (2010).

[111] I. Markovsky, *Low-rank approximation* (Springer, 2019).

[112] R. Orsi, U. Helmke, and J. B. Moore, "A Newton-like method for solving rank constrained linear matrix inequalities", Automatica **42**, 1875 (2006).

[113] C. Sun and R. Dai, "Rank-constrained optimization and its applications", Automatica **82**, 128 (2017).

[114] M. Weilenmann, B. Dive, D. Trillo, E. A. Aguilar, and M. Navascués, "Entanglement detection beyond measuring fidelities", Phys. Rev. Lett. **124**, 200502 (2020), Erratum: Phys. Rev. Lett. **125**, 159903(E) (2020).

[115] P. M. Alberti and A. Uhlmann, *Stochasticity and partial order* (Deutscher Verlag der Wissenschaften, Berlin, 1982).

[116] C. D.-Y. Lee and J. Watrous, "Detecting mixed-unitary quantum channels is NP-hard", Quantum **4**, 253 (2020).

[117] Ł. Rudnicki, Z. Puchała, and K. Życzkowski, "Strong majorization entropic uncertainty relations", Phys. Rev. A **89**, 052115 (2014).

[118] G. Gour, A. Grudka, M. Horodecki, W. Kłobus, J. Łodyga, and V. Narasimhachar, "Conditional uncertainty principle", Phys. Rev. A **97**, 042130 (2018).

[119] L. Lovász, *Graphs and geometry* (American Mathematical Society, 2019).

[120] A. Cabello, S. Severini, and A. Winter, "Graph-theoretic approach to quantum correlations", Phys. Rev. Lett. **112**, 040401 (2014).

[121] R. Ramanathan and P. Horodecki, "Necessary and sufficient condition for state-independent contextual measurement scenarios", Phys. Rev. Lett. **112**, 040404 (2014).

[122] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming", JACM **42**, 1115 (1995).

[123] E. Boros and P. L. Hammer, "Pseudo-boolean optimization", Discrete Appl. Math. **123**, 155 (2002).

[124] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", Science **220**, 671 (1983).

[125] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks", Rev. Mod. Phys. **80**, 1275 (2008).

[126] A. Lucas, "Ising formulations of many NP problems", Front. Phys. **2**, 5 (2014).

[127] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, "Evidence for quantum annealing with more than one hundred qubits", Nat. Phys. **10**, 218 (2014).

[128] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm" (2014), arXiv:1411.4028 [quant-ph] .

[129] E. Farhi and A. W. Harrow, "Quantum supremacy through the quantum approximate optimization algorithm" (2019), arXiv:1602.07674 [quant-ph] .

[130] F. Vallentin, "Symmetry in semidefinite programs", Linear Algebra Its Appl. **430**, 360 (2009).

[131] D. Rosset, "SymDPoly: Symmetry-adapted moment relaxations for noncommutative polynomial optimization" (2018), arXiv:1808.09598 [quant-ph] .

[132] E. A. Aguilar, J. J. Borkała, P. Mironowicz, and M. Pawłowski, "Connections between mutually unbiased bases and quantum random access codes", Phys. Rev. Lett. **121**, 050501 (2018).

[133] G. Tóth and O. Gühne, "Entanglement and permutational symmetry", Phys. Rev. Lett. **102**, 170503 (2009).

[134] X.-D. Yu, T. Simnacher, N. Wyderka, H. C. Nguyen, and O. Gühne, "A complete hierarchy for the pure state marginal problem in quantum mechanics", Nat. Commun. **12**, 1012 (2021).

[135] M. Christandl, R. König, G. Mitchison, and R. Renner, "One-and-a-half quantum de Finetti theorems", Commun. Math. Phys. **273**, 473 (2007).

[136] M. Navascués, M. Owari, and M. B. Plenio, "Power of symmetric extensions for entanglement detection", Phys. Rev. A **80**, 052306 (2009).

[137] M. Berta, F. Borderi, O. Fawzi, and V. Scholz, "Semidefinite programming hierarchies for quantum error correction" (2018), arXiv:1810.12197 [quant-ph] .

[138] C. M. Caves, C. A. Fuchs, and R. Schack, "Unknown quantum states: The quantum de Finetti representation", J. Math. Phys. **43**, 4537 (2002).

[139] P. Seymour and T. Zaslavsky, "Averaging sets: A generalization of mean values and spherical designs", Adv. Math. **52**, 213 (1984).

[140] D. Gross, K. Audenaert, and J. Eisert, "Evenly distributed unitaries: On the structure of unitary designs", J. Math. Phys. **48**, 052104 (2007).

[141] L. Chen, D. Chu, L. Qian, and Y. Shen, "Separability of completely symmetric states in a multipartite system", Phys. Rev. A **99**, 032312 (2019).

[142] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, Vol. 255 (Springer, 2009).

[143] M. Ray, N. G. Boddu, K. Bharti, L.-C. Kwek, and A. Cabello, "Graph-theoretic approach to dimension witnessing", New J. Phys. **23**, 033006 (2021).

[144] R. Hamerly *et al.*, "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer", Sci. Adv. **5**, eaau0823 (2019).

[145] G. E. Crooks, "Performance of the quantum approximate optimization algorithm on the maximum cut problem" (2018), arXiv:1811.08419 [quant-ph] .

[146] Y. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization", Optim. Method. Softw. **9**, 141 (1998).

[147] M. Nakata, "A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP,-QD and -DD", in *2010 IEEE international symposium on computer-aided control system design* (IEEE, 2010) pp. 29–34.

[148] M. Nakata, B. J. Braams, K. Fujisawa, M. Fukuda, J. K. Percus, M. Yamashita, and Z. Zhao, "Variational calculation of second-order reduced density matrices by strong n-representability conditions and an accurate semidefinite programming solver", J. Chem. Phys. **128**, 164113 (2008).

[149] M. Yamashita, K. Fujisawa, M. Fukuda, K. Kobayashi, K. Nakata, and M. Nakata, "Latest developments in the SDPA family for solving large-scale SDPs", in *Handbook on semidefinite, conic and polynomial optimization* (Springer, 2012) pp. 687–713.

[150] U. Leverrier, "Sur les variations séculaire des élements des orbites pour les sept planétes principales", J. de Math , 5 (1840).

[151] D. Faddeev and I. Sominskii, "Collection of problems on higher algebra", Gostekhizdat, Moscow, (1949).

[152] R. Descartes, *La géométrie* (De l'Imprimerie de Ian Maire, 1637).

[153] O. Gamel, "Entangled Bloch spheres: Bloch matrix and two-qubit state space", Phys. Rev. A **93**, 062320 (2016).

[154] J. Watrous, *The theory of quantum information* (Cambridge University Press, 2018).

[155] V. I. Bogachev, *Measure theory*, Vol. 1 & 2 (Springer Science & Business Media, 2007).

[156] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik", Naturwissenschaften **23**, 807 (1935).

[157] A. J. Coleman, "Structure of fermion density matrices", Rev. Mod. Phys. **35**, 668 (1963).

[158] A. Klyachko, "Quantum marginal problem and N-representability", J. Phys. Conf. Ser. **36**, 72 (2006).

[159] C. Schilling, *Quantum Marginal Problem and its Physical Relevance*, Ph.D. thesis, ETH Zürich (2015).

[160] N. Linden, S. Popescu, and W. Wootters, "Almost every pure state of three qubits is completely determined by its two-particle reduced density matrices", Phys. Rev. Lett. **89**, 207901 (2002).

[161] A. Sawicki, M. Walter, and M. Kuś, "When is a pure state of three qubits determined by its single-particle reduced density matrices?", J. Phys. A: Math. Theor. **46**, 055304 (2013).

[162] N. Wyderka, F. Huber, and O. Gühne, "Almost all four-particle pure states are determined by their two-body marginals", Phys. Rev. A **96**, 010102 (2017).

[163] F. Huber and O. Gühne, "Characterizing ground and thermal states of few-body Hamiltonians", Phys. Rev. Lett. **117**, 010403 (2016).

[164] S. Karuvade, P. D. Johnson, F. Ticozzi, and L. Viola, "Uniquely determined pure quantum states need not be unique ground states of quasi-local Hamiltonians", Phys. Rev. A **99**, 062104 (2019).

[165] J. Eisert, T. Tyc, T. Rudolph, and B. C. Sanders, "Gaussian quantum marginal problem", Commun. Math. Phys. **280**, 263 (2008).

[166] A. Aloy, M. Fadel, and J. Tura, "The quantum marginal problem for symmetric states: Applications to variational optimization, nonlocality and self-testing", New J. Phys. **23**, 033026 (2021).

[167] M. Walter, B. Doran, D. Gross, and M. Christandl, "Entanglement polytopes: Multiparticle entanglement from single-particle information", Science **340**, 1205 (2013).

[168] R. Chaves, C. Majenz, and D. Gross, "Information-theoretic implications of quantum causal structures", Nat. Commun. **6**, 5766 (2015).

[169] C. Schilling, C. L. Benavides-Riveros, A. Lopes, T. Maciążek, and A. Sawicki, "Implications of pinned occupation numbers for natural orbital expansions: I. Generalizing the concept of active spaces", New J. Phys. **22**, 023001 (2020).

[170] T. Maciążek, A. Sawicki, D. Gross, A. Lopes, and C. Schilling, "Implications of pinned occupation numbers for natural orbital expansions. II: Rigorous derivation and extension to non-fermionic systems", New J. Phys. **22**, 023002 (2020).

[171] D. Goyeneche, D. Alsina, J. I. Latorre, A. Riera, and K. Życzkowski, "Absolutely maximally entangled states, combinatorial designs, and multiunitary matrices", Phys. Rev. A **92**, 032316 (2015).

[172] F. Huber, O. Gühne, and J. Siewert, "Absolutely maximally entangled states of seven qubits do not exist", Phys. Rev. Lett. **118**, 200502 (2017).

[173] F. Huber, C. Eltschka, J. Siewert, and O. Gühne, "Bounds on absolutely maximally entangled states from shadow inequalities, and the quantum MacWilliams identity", J. Phys. A: Math. Theor. **51**, 175301 (2018).

[174] J. Bryan, S. Leutheusser, Z. Reichstein, and M. V. Raamsdonk, "Locally maximally entangled states of multipart quantum systems", Quantum **3**, 115 (2019).

[175] Z. Raissi, A. Teixidó, C. Gogolin, and A. Acín, "Constructions of k-uniform and absolutely maximally entangled states beyond maximum distance codes", Phys. Rev. Research **2**, 033411 (2020).

[176] M. Grassl, "Bounds on the minimum distance of linear codes and quantum codes", `http://www.codetables.de/`.

[177] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, "Quantum computers", Nature **464**, 45 (2010).

[178] J. Preskill, "Quantum computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[179] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505 (2019).

[180] P. Horodecki, Ł. Rudnicki, and K. Życzkowski, "Five open problems in quantum information" (2020), arXiv:2002.03233 [quant-ph] .

[181] S. A. Rather, A. Burchardt, W. Bruzda, G. Rajchel-Mieldzioć, A. Lakshminarayan, and K. Życzkowski, "Thirty-six entangled officers of Euler" (2021), arXiv:2104.05122 [quant-ph] .

[182] Y.-K. Liu, "Consistency of local density matrices is QMA-complete", in *Approximation, randomization, and combinatorial optimization. Algorithms and techniques* (2006) pp. 438–449.

[183] Y.-K. Liu, M. Christandl, and F. Verstraete, "Quantum computational complexity of the N-representability problem: QMA complete", Phys. Rev. Lett. **98**, 110503 (2007).

[184] P. Bürgisser, M. Christandl, K. D. Mulmuley, and M. Walter, "Membership in moment polytopes is in NP and coNP", SIAM J. Comput. **46**, 972 (2017).

[185] P. Bürgisser, A. Garg, R. Oliveira, M. Walter, and A. Wigderson, "Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory", in *9th innovations in theoretical computer science conference (ITCS 2018)* (2018) pp. 24:1–24:20.

[186] P. Bürgisser, C. Franks, A. Garg, R. Oliveira, M. Walter, and A. Wigderson, "Efficient algorithms for tensor scaling, quantum marginals, and moment polytopes", in *2018 IEEE 59th annual symposium on foundations of computer science (FOCS)* (2018) pp. 883–897.

[187] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model", Phys. Rev. A **40**, 4277 (1989).

[188] L. P. Lebedev, M. J. Cloud, and V. A. Eremeyev, *Tensor analysis with applications in mechanics* (World Scientific, Singapore, 2010).

[189] A. Burchardt and Z. Raissi, "Stochastic local operations with classical communication of absolutely maximally entangled states", Phys. Rev. A **102**, 022413 (2020).

[190] W. Fulton and J. Harris, *Representation theory: A first course*, Vol. 129 (Springer-Verlag, Berlin, 1991).

[191] H. Boerner, *Representations of groups* (North-Holland, Amsterdam, 1963).

[192] *GAP — groups, algorithms, and programming, version 4.11.0*, The GAP Group (2020), `https://www.gap-system.org`.

[193] Y. Zhang, L. H. Kauffman, and R. F. Werner, "Permutation and its partial transpose", Int. J. Quantum Inf. **5**, 469 (2007).

[194] M. Studziński, M. Horodecki, and M. Mozrzymas, "Commutant structure of $U^{\otimes(n-1)} \otimes U^*$ transformations", J. Phys. A: Math. Theor. **46**, 395303 (2013).

[195] M. Mozrzymas, M. Horodecki, and M. Studziński, "Structure and properties of the algebra of partially transposed permutation operators", J. Math. Phys. **55**, 032202 (2014).

[196] M. Mozrzymas, M. Studziński, and M. Horodecki, "A simplified formalism of the algebra of partially transposed permutation operators with applications", J. Phys. A: Math. Theor. **51**, 125202 (2018).

[197] T. Eggeling and R. F. Werner, "Separability properties of tripartite states with $U \otimes U \otimes U$ symmetry", Phys. Rev. A **63**, 042111 (2001).

[198] B. Buchberger, "Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems", Aequationes mathematicae **4**, 374 (1970).

[199] T. Simnacher, N. Wyderka, R. Schwonnek, and O. Gühne, "Entanglement detection with scrambled data", Phys. Rev. A **99**, 062339 (2019).

[200] M. Seevinck and J. Uffink, "Local commutativity versus Bell inequality violation for entangled states and versus non-violation for separable states", Phys. Rev. A **76**, 042105 (2007).

[201] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, "Squashing models for optical measurements in quantum communication", Phys. Rev. Lett. **101**, 093601 (2008).

[202] T. Moroder and O. Gittsovich, "Calibration-robust entanglement detection beyond Bell inequalities", Phys. Rev. A **85**, 032301 (2012).

[203] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, "Device-independent witnesses of genuine multipartite entanglement", Phys. Rev. Lett. **106**, 250404 (2011).

[204] J. T. Barreiro, J.-D. Bancal, P. Schindler, D. Nigg, M. Hennrich, T. Monz, N. Gisin, and R. Blatt, "Demonstration of genuine multipartite entanglement with device-independent witnesses", Nat. Phys. **9**, 559 (2013).

[205] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, "Experimental violation of a Bell's inequality with efficient detection", Nature **409**, 791 (2001).

[206] Z. Wang, S. Singh, M. Navascués, *et al.*, "Entanglement and nonlocality in infinite 1D systems", Phys. Rev. Lett. **118**, 230401 (2017).

[207] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, "Permutationally invariant state reconstruction", New J. Phys. **14**, 105001 (2012).

[208] T. Bastin, P. Mathonet, and E. Solano, "Operational entanglement families of symmetric mixed N-qubit states", Phys. Rev. A **91**, 022310 (2015).

[209] P. Migdał, J. Rodriguez-Laguna, and M. Lewenstein, "Entanglement classes of permutation-symmetric qudit states: Symmetric operations suffice", Phys. Rev. A **88**, 012335 (2013).

[210] S. Brierley, S. Weigert, and I. Bengtsson, "All mutually unbiased bases in dimensions two to five", Quantum Inf. Comput. **10**, 803 (2010).

[211] R. Schwonnek, "Additivity of entropic uncertainty relations", Quantum **2**, 59 (2018).

[212] K. Abdelkhalek, R. Schwonnek, H. Maassen, F. Furrer, J. Duhme, P. Raynal, B.-G. Englert, and R. F. Werner, "Optimality of entropic uncertainty relations", Int. J. Quantum Inf. **13**, 1550045 (2015).

[213] D. W. Berry and B. C. Sanders, "Bounds on general entropy measures", J. Phys. A: Math. Gen. **36**, 12255 (2003).

[214] S. Wehner and A. Winter, "Entropic uncertainty relations — A survey", New J. Phys. **12**, 025009 (2010).

[215] A. Sanpera, R. Tarrach, and G. Vidal, "Local description of quantum inseparability", Phys. Rev. A **58**, 826 (1998).

[216] P. Horodecki and A. Ekert, "Method for direct detection of quantum entanglement", Phys. Rev. Lett. **89**, 127902 (2002).

[217] S. Johri, D. S. Steiger, and M. Troyer, "Entanglement spectroscopy on a quantum computer", Phys. Rev. B **96**, 195136 (2017).

[218] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, "Measuring the Rényi entropy of a two-site Fermi-Hubbard model on a trapped ion quantum computer", Phys. Rev. A **98**, 052334 (2018).

[219] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, "Probing Rényi entanglement entropy via randomized measurements", Science **364**, 260 (2019).

[220] M. Curty, O. Gühne, M. Lewenstein, and N. Lütkenhaus, "Detecting two-party quantum correlations in quantum-key-distribution protocols", Phys. Rev. A **71**, 022306 (2005).

[221] T. Simnacher, J. Czartowski, K. Szymański, and K. Życzkowski, "Confident entanglement detection via separable numerical range" (2021), arXiv:2107.04365 [quant-ph] .

[222] F. Shahandeh, M. Ringbauer, J. C. Loredo, and T. C. Ralph, "Ultrafine entanglement witnessing", Phys. Rev. Lett. **118**, 110502 (2017).

[223] M. Gachechiladze, N. Wyderka, and O. Gühne, "The structure of ultrafine entanglement witnesses", J. Phys. A: Math. Theor. **51**, 365307 (2018).

[224] P. Horodecki, "From limits of quantum operations to multicopy entanglement witnesses and state-spectrum estimation", Phys. Rev. A **68**, 052101 (2003).

[225] O. Gühne and N. Lütkenhaus, "Nonlinear entanglement witnesses", Phys. Rev. Lett. **96**, 170502 (2006).

[226] T. Moroder, O. Gühne, and N. Lütkenhaus, "Iterations of nonlinear entanglement witnesses", Phys. Rev. A **78**, 032326 (2008).

[227] J.-Y. Wu, H. Kampermann, D. Bruß, C. Klöckl, and M. Huber, "Determining lower bounds on a measure of multipartite entanglement from few local observables", Phys. Rev. A **86**, 022319 (2012).

[228] Ł. Pawela, P. Gawron, J. A. Miszczak, Z. Puchała, K. Życzkowski, P. Lewandowska, and R. Kukulski, "Numerical shadow and numerical range", `https://www.numericalshadow.org/`.

[229] P. Wu and R. Tang, "Joint separable numerical range and bipartite ultrafine entanglement witnessing", J. Phys. A: Math. Theor. **53**, 445302 (2020).

[230] J. Czartowski, K. Szymański, B. Gardas, Y. V. Fyodorov, and K. Życzkowski, "Separability gap and large-deviation entanglement criterion", Phys. Rev. A **100**, 042326 (2019).

[231] W. Hoeffding, "Probability inequalities for sums of bounded random variables", J. Am. Stat. Assoc. **58**, 13 (1963).

[232] C. C. Cowen and E. Harel, "An effective algorithm for computing the numerical range", Unpublished manuscript. `https://www.math.iupui.edu/~ccowen/Downloads/33NumRange.pdf` (1995).

[233] P. Gawron and P. Sadowski, "Approximation of separable numerical range using simulated annealing", Theoretical and Applied Informatics , 149 (2014).

[234] L. Gurvits and H. Barnum, "Better bound on the exponent of the radius of the multipartite separable ball", Phys. Rev. A **72**, 032322 (2005).

[235] R. Hildebrand, "Entangled states close to the maximally mixed state", Phys. Rev. A **75**, 062330 (2007).

[236] L. Gurvits and H. Barnum, "Largest separable balls around the maximally mixed bipartite quantum state", Phys. Rev. A **66**, 062311 (2002).

[237] F. T. Hioe and J. H. Eberly, "N-level coherence vector and higher conservation laws in quantum optics and quantum mechanics", Phys. Rev. Lett. **47**, 838 (1981).

[238] P. B. Slater, "A concise formula for generalized two-qubit Hilbert-Schmidt separability probabilities", J. Phys. A: Math. Theor. **46**, 445302 (2013).

[239] P. B. Slater and C. F. Dunkl, "Formulas for rational-valued separability probabilities of random induced generalized two-qubit states", Adv. Theor. Math. Phys. , 621353 (2015).

[240] J. Fei and R. Joynt, "Numerical computations of separability probabilities", Rep. Math. Phys. **78**, 177 (2016).

[241] A. Lovas and A. Andai, "Invariance of separability probability over reduced states in 4×4 bipartite systems", J. Phys. A: Math. Theor. **50**, 295303 (2017).

[242] K. Życzkowski and H.-J. Sommers, "Hilbert-Schmidt volume of the set of mixed quantum states", J. Phys. A: Math. Gen. **36**, 10115 (2003).

[243] M. Kuś and K. Życzkowski, "Geometry of entangled states", Phys. Rev. A **63**, 032307 (2001).

[244] S. Ishizaka and T. Hiroshima, "Maximally entangled mixed states under nonlocal unitary operations in two qubits", Phys. Rev. A **62**, 022310 (2000).

[245] F. Verstraete, K. Audenaert, and B. De Moor, "Maximally entangled mixed states of two qubits", Phys. Rev. A **64**, 012316 (2001).

[246] K. R. Parthasarathy, "On the maximal dimension of a completely entangled subspace for finite level quantum systems", Proc.: Math. Sci. **114**, 365 (2004).

[247] J. Walgate and A. J. Scott, "Generic local distinguishability and completely entangled subspaces", J. Phys. A: Math. Theor. **41**, 375305 (2008).

[248] T. Heinosaari, D. Reitzner, and P. Stano, "Notes on joint measurability of quantum observables", Found. Phys. **38**, 1133 (2008).

[249] D. P. DiVincenzo, "The physical implementation of quantum computation", Fortschr. Phys. **48**, 771 (2000).

[250] C. Simon, M. Afzelius, J. Appel, A. B. de La Giroday, S. Dewhurst, N. Gisin, C. Hu, F. Jelezko, S. Kröll, J. Müller, *et al.*, "Quantum memories", Eur. Phys. J. D **58**, 1 (2010).

[251] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum repeaters: The role of imperfect local operations in quantum communication", Phys. Rev. Lett. **81**, 5932 (1998).

[252] L.-M. Duan, M. Lukin, J. I. Cirac, and P. Zoller, "Long-distance quantum communication with atomic ensembles and linear optics", Nature **414**, 413 (2001).

[253] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurášek, and E. S. Polzik, "Experimental demonstration of quantum memory for light", Nature **432**, 482 (2004).

[254] K. S. Choi, H. Deng, J. Laurat, and H. J. Kimble, "Mapping photonic entanglement into and out of a quantum memory", Nature **452**, 67 (2008).

[255] R. Zhao, Y. Dudin, S. Jenkins, C. Campbell, D. Matsukevich, T. Kennedy, and A. Kuzmich, "Long-lived quantum memory", Nat. Phys. **5**, 100 (2009).

[256] M. P. Hedges, J. J. Longdell, Y. Li, and M. J. Sellars, "Efficient quantum memory for light", Nature **465**, 1052 (2010).

[257] K. Jensen, W. Wasilewski, H. Krauter, T. Fernholz, B. M. Nielsen, M. Owari, M. B. Plenio, A. Serafini, M. Wolf, and E. Polzik, "Quantum memory for entangled continuous-variable states", Nat. Phys. **7**, 13 (2011).

[258] Y. Pu, N. Jiang, W. Chang, H. Yang, C. Li, and L. Duan, "Experimental realization of a multiplexed quantum memory with 225 individually accessible memory cells", Nat. Commun. **8**, 15359 (2017).

[259] W. Zhang, D.-S. Ding, Y.-B. Sheng, L. Zhou, B.-S. Shi, and G.-C. Guo, "Quantum secure direct communication with quantum memory", Phys. Rev. Lett. **118**, 220501 (2017).

[260] I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box", J. Mod. Opt. **44**, 2455 (1997).

[261] M. Mohseni, A. Rezakhani, and D. Lidar, "Quantum-process tomography: Resource analysis of different strategies", Phys. Rev. A **77**, 032322 (2008).

[262] H. Häseler, T. Moroder, and N. Lütkenhaus, "Testing quantum devices: Practical entanglement verification in bipartite optical systems", Phys. Rev. A **77**, 032303 (2008).

[263] H. Häseler and N. Lütkenhaus, "Probing the quantumness of channels with mixed states", Phys. Rev. A **80**, 042304 (2009).

[264] M. F. Pusey, "Verifying the quantumness of a channel with an untrusted device", J. Opt. Soc. Am. B **32**, A56 (2015).

[265] D. Rosset, F. Buscemi, and Y.-C. Liang, "Resource theory of quantum memories and their faithful verification with minimal assumptions", Phys. Rev. X **8**, 021033 (2018).

[266] M.-L. Hu and H. Fan, "Relative quantum coherence, incompatibility, and quantum correlations of states", Phys. Rev. A **95**, 052106 (2017).

[267] M. Ringbauer, T. R. Bromley, M. Cianciaruso, L. Lami, W. S. Lau, G. Adesso, A. G. White, A. Fedrizzi, and M. Piani, "Certification and quantification of multilevel quantum coherence", Phys. Rev. X **8**, 041007 (2018).

[268] T. Kraft and M. Piani, "Genuine correlated coherence", J. Phys. A: Math. Theor. **51**, 414013 (2018).

[269] F. Shahbeigi and S. J. Akhtarshenas, "Quantumness of quantum channels", Phys. Rev. A **98**, 042313 (2018).

[270] M. Idel and M. M. Wolf, "Sinkhorn normal form for unitary matrices", Linear Algebra Its Appl. **471**, 76 (2015).

[271] M. B. Ruskai, "Qubit entanglement breaking channels", Rev. Math. Phys. **15**, 643 (2003).

[272] M. B. Ruskai, S. Szarek, and E. Werner, "An analysis of completely-positive trace-preserving maps on $\mathcal{M}_2$", Linear Algebra Its Appl. **347**, 159 (2002).

[273] D. Braun, O. Giraud, I. Nechita, C. Pellegrini, and M. Žnidarič, "A universal set of qubit quantum channels", J. Phys. A: Math. Theor. **47**, 135302 (2014).

[274] R. A. Horn and C. R. Johnson, *Topics in matrix analysis* (Cambridge University Press, 1991).

[275] X.-D. Yu and O. Gühne, "Detecting coherence via spectrum estimation", Phys. Rev. A **99**, 062310 (2019).