

DISSERTATION | DOCTORAL THESIS

Titel | Title

Bound Entanglement and Distillability of Bell-Diagonal Qudits

verfasst von | submitted by

Christopher Popp B.Sc. B.Sc. M.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 796 605 411

Dissertationsgebiet lt. Studienblatt | Field of
study as it appears on the student record
sheet:

Physik

Betreut von | Supervisor:

Mag. Dr. Beatrix Hiesmayr Privatdoz.

Abstract

Quantum entanglement is a central resource in quantum information science. However, its structure and operational role remain only partially understood. Two open problems illustrate this: the separability problem, which asks whether a given state is separable or entangled, and the distillability problem, which concerns whether multiple copies of a state can be transformed into maximally entangled states using local operations and classical communication (LOCC). Bound entangled states, which are entangled but undistillable and difficult to distinguish from separable or distillable states, lie at the intersection of these challenges.

This thesis investigates the entanglement structure of bipartite Belldiagonal systems, i.e., mixtures of maximally entangled states, and its relation to distillability and bound entanglement. Employing the special properties of Belldiagonal systems, we develop a numerical ecosystem for state sampling and entanglement analysis, ensuring reproducibility of results. Detecting a significant portion of bound entangled states, we classify the majority of lowdimensional Belldiagonal states generated via the WeylHeisenberg construction. We demonstrate that the underlying group structure plays a decisive role in shaping entanglement properties and the prevalence and detectability of bound entanglement in such systems. These structural insights are harnessed to develop optimized stabilizerbased distillation protocols. Building on this framework, we design the FIMAX protocol, achieving superior performance compared to established protocols in efficiency and robustness to noise, even for lowfidelity or nonBelldiagonal inputs. Finally, we facilitate extensions for more general resource-state distillation, including secret-key distillation. Establishing local invariance under isometric channels and providing an algorithm for computing relevant information measures enables the use of stabilizerbased methods to distill states optimized for the corresponding operational tasks.

The results establish Belldiagonal systems as a fertile platform for exploring bound entanglement and distillability, as well as for operationally designing effective distillation protocols. They underscore the group structure as an exploitable principle that connects entanglement classification with distillation in Belldiagonal systems. While indicating promising directions for future investigations, the thesis advances both the theoretical understanding and the operational application of entanglement within quantum information science.

Kurzfassung

Verschränkung ist eine zentrale Ressource der Quanteninformationswissenschaft, deren Struktur und operative Rolle jedoch nur teilweise verstanden sind. Zwei offene Probleme verdeutlichen diese Lücken: das Separabilitätsproblem, das fragt, ob ein gegebener Zustand separabel oder verschränkt ist, und das Destillierbarkeitsproblem, das untersucht, ob mehrere Kopien eines Zustands mittels lokaler Operationen und klassischer Kommunikation (LOCC) in maximal verschränkte Zustände transformiert werden können. Gebundene verschränkte Zustände, die zwar verschränkt, aber nicht destillierbar sind und nicht zuverlässig von separablen oder destillierbaren Zuständen unterschieden werden können, liegen an der Schnittstelle beider Herausforderungen.

Diese Dissertation untersucht die Verschränkungsstruktur bipartiter Belldiagonaler Systeme, d.h. Mischungen maximal verschränkter Zustände, und deren Zusammenhang mit Destillierbarkeit und gebundener Verschränkung. Unter Ausnutzung spezieller Eigenschaften Belldiagonaler Systeme entwickeln wir ein numerisches Ökosystem zur Datenerzeugung und Verschränkungsanalyse, das Reproduzierbarkeit der Ergebnisse gewährleistet. Durch die Detektion eines bedeutenden Anteils gebundener verschränkter Zustände klassifizieren wir den Großteil der niedrigdimensionalen Bell-diagonalen Zustände, die mittels der WeylHeisenberg-Konstruktion erzeugt werden. Wir zeigen, dass die zugrunde liegende Gruppenstruktur eine entscheidende Rolle für die Verschränkungsstruktur als auch für das Vorkommen und die Nachweisbarkeit von gebundener Verschränkung in solchen Systemen spielt. Diese strukturellen Erkenntnisse werden genutzt, um optimierte stabilisatorbasierte Protokolle zur Verschränkungsdestillation zu entwickeln und präsentieren das FIMAXProtokoll, das sich durch eine hohe Effizienz und Robustheit gegenüber Rauschen auszeichnet, selbst für nichtBelldiagonale Zustände. Abschließend legen wir den Grundstein zur Erweiterung der Stabilisatormethoden auf andere Destillationsaufgaben, wie zum Beispiel die Destillation eines kryptografischen Schlüssels. Das Etablieren lokaler Invarianz unter isometrischen Transformationen relevanter Informationsmaße sowie die Entwicklung eines Algorithmus zur Berechnung dieser, erlauben den Einsatz stabilisatorbasierter Methoden zur Destillation von Zuständen, die für die jeweiligen operationellen Aufgaben optimiert sind.

Die Ergebnisse etablieren Belldiagonale Systeme als geeignete Plattform zur Untersuchung der gebundenen Verschränkung und Destillierbarkeit sowie zur Entwicklung wirksamer Destillationsprotokolle. Sie heben die Gruppenstruktur als ein nutzbares Prinzip hervor, das die Klassifikation und Destillation von Belldiagonalen Zuständen miteinander verbindet. Während sie vielversprechende Richtungen für zukünftige Untersuchungen aufzeigt, erweitert diese Arbeit sowohl das theoretische Verständnis als auch die operationelle Anwendung von Verschränkung in der Quanteninformationswissenschaft.

Acknowledgements

I thank Beatrix Hiesmayr for her confidence and trust in me to start this endeavor, her expertise and support throughout the project, always encouraging and enabling me to follow my scientific interests, her personal and professional guidance, and many enjoyable experiences throughout these years.

I thank Tobias Sutter for his motivating companionship and his valuable professional support.

I thank Thomas Juffmann for his advice and guidance.

I thank Mara Vogler for her unconditional support at all times.

I thank my mother Marliese Megner for encouraging me to follow this path.

Thank you all for making this journey possible.

The author acknowledges the Austrian Science Fund (FWF) [10.55776/P36102] for financial support.

The author also acknowledges the use of the AI-based software “LanguageTool”, “Microsoft Copilot”, and “xAI Grok” for linguistic refinement.

Inhaltsverzeichnis

Abstract	i
Kurzfassung	iii
Acknowledgements	v
1. Preamble	1
1.1. Background and motivation	1
1.2. Scope of the dissertation	5
2. Publication overview	9
2.1. Entanglement classification of Bell-diagonal states	11
2.1.1. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits [72]	11
2.1.2. Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits [73]	11
2.1.3. Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts [74]	12
2.2. Relation of group and entanglement structure in Bell-diagonal systems . .	13
2.2.1. Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states [75]	13
2.2.2. A Group-Theoretic Perspective on the PPT and Realignment Ent- anglement Criteria [76]	13
2.3. Entanglement distillation	14
2.3.1. A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits [77]	14
2.3.2. Low-Fidelity Entanglement Distillation with FIMAX [78]	15
2.4. Properties and computation of distillation-relevant information measures .	15
2.4.1. Local Invariance of Divergence-Based Quantum Information Mea- sures [79]	16
2.4.2. Computation of the Smooth Max-Mutual Information via Semide- finite Programming [80]	16
2.5. Review of bipartite bound entanglement	17
2.5.1. Bipartite Bound Entanglement [21]	17
2.6. Dissipative quantum neural networks	17
2.6.1. The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks [81]	17

3. Synopsis and conclusion	19
3.1. Synopsis of results	19
3.1.1. Developing numerical methods for entanglement classification and distillation of Bell-diagonal states	19
3.1.2. Validating Bell-diagonal systems for the analysis and characterization of bound entanglement	20
3.1.3. Identifying the underlying group structure of Bell states as a relevant factor for the entanglement structure of Bell-diagonal systems	21
3.1.4. Leveraging the group structure of standard Bell states for entanglement distillation via stabilizer-based methods	22
3.1.5. Laying the foundation for the generalization and extension of stabilizer-based distillation	23
3.1.6. Open questions	24
3.1.7. Review of bipartite bound entanglement	25
3.1.8. Further work	26
3.2. Synopsis of methods	26
3.2.1. Hybrid approach and software ecosystem	26
3.2.2. Methods for entanglement classification and validation of Bell-diagonal systems	27
3.2.3. Methods for identifying group structure as a relevant factor for the entanglement structure of Bell-diagonal systems	29
3.2.4. Methods for stabilizer-based entanglement distillation	30
3.2.5. Methods for the generalization and extension of stabilizer-based distillation	31
3.3. Concluding discussion	31
Literatur	35
A. Preliminaries	43
A.1. Notation	43
A.2. Definitions	43
B. Publications	49
B.1. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits	49
B.2. Almost complete solution for the NPhard separability problem of Bell diagonal qutrits	52
B.3. Comparing bound entanglement of Bell-diagonal pairs of qutrits and ququarts	65
B.4. Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states	81
B.5. A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria	96
B.6. A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits	109
B.7. Low-fidelity entanglement distillation with FIMAX	139

B.8. Local Invariance of Divergence-based Quantum Information Measures . .	156
B.9. Computation of the Smooth Max-Mutual Information via Semidefinite Programming	168
B.10. Bipartite bound entanglement	178
B.11. The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks	253

1. Preamble

This chapter introduces the background, motivation, and scope of the primary focus of the thesis: the entanglement structure of bipartite Bell-diagonal systems. We also introduce the main theoretical objects and concepts with direct relation to the results of the publications presented in Chapter 2 and discussed in Chapter 3. A collection of the theoretical essentials of quantum information theory and notation as presupposed throughout this thesis can be found in appendix A. All considered publications are presented in their published form in appendix B.

1.1. Background and motivation

Utilizing non-classical phenomena of quantum physics holds significant promise for the application to future technology [1]. Despite many experimental challenges, the performance of important tasks in fields like communication [2], computing [3, 4], cryptography [5–7] and metrology [8] can be improved beyond the capabilities of any system relying on classical methods only.

Quantum entanglement [9] is one of the defining non-classical correlations of quantum mechanics and a central resource for quantum information processing, underpinning protocols from teleportation [10] and dense coding [11, 12] to quantum key distribution [5, 7] and distributed sensing [8]. Resource theoretic and operational frameworks [13, 14] make these connections precise, providing measures and operational tasks that translate abstract correlations into concrete capabilities and costs.

Despite its relevance, a complete understanding of quantum entanglement has yet to be established. One symptom of this lack of understanding of this phenomenon manifests in its detection. To date, there is no sufficient and necessary criterion known that detects all entangled quantum states. Existing methods such as the PPT criterion [15], the realignment criterion [16, 17], approaches based on correlation tensors [18], or entanglement witnesses [19] exhibit limited detection power, which depends on the dimension and other structural properties of the quantum systems under consideration. This challenge is known as the separability problem [20]. It can be formulated as the decision problem of whether a given quantum state is in the convex set of separable, i.e., non-entangled, states or not. More precisely, following the notation as presented in appendix A and Ref. [21]:

Definition 1.1.1 (Separability and Entanglement)

A bipartite quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called separable if it can be written as a convex combination of product states, i.e., $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ with probabilities $p_i \geq 0$,

$\sum_i p_i = 1$, $\rho_A^i \in \mathcal{D}(\mathcal{H}_A)$, and $\rho_B^i \in \mathcal{D}(\mathcal{H}_B)$. If ρ_{AB} is not separable, it is called entangled. We denote the set of all separable (entangled) states by SEP (ENT).

Definition 1.1.2 (Separability Problem)

Given $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, decide whether or not $\rho_{AB} \in \text{SEP}$.

Entanglement measures [22] extend this problem by trying to quantify how strongly a given state is entangled. For pure states the entropy of entanglement [23] is the canonical measure. However, several generalizations to mixed states are possible. Proposed measures like the entanglement cost [24], the distillable secret key [25] or the distillable entanglement [23, 26] are relevant for certain operational tasks. Which single measure universally quantifies the resource entanglement, if such a canonical measure exists at all, is an open question.

Bell states are pure, maximally entangled states, which can be characterized by their reduced or marginal states. For a bipartite state ρ_{AB} , these are defined as $\rho_{A/B} := \text{Tr}_{B/A}[\rho_{AB}]$. Maximally entangled states are pure states whose marginal states are maximally mixed:

Definition 1.1.3 (Maximally Entangled States)

A pure bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$ is said to be maximally entangled if the partial trace on either system yields the maximally mixed state, i.e., $\text{Tr}_A(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_B$ or $\text{Tr}_B(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_A$.

Equivalently, maximally entangled pure states can also be defined by their Schmidt coefficients being all equal to $1/\sqrt{d}$ [27]. Note that this implies that if the reduced state of one party of a pure bipartite state is maximally mixed, the reduced state of the other party is also maximally mixed.

Bell states yield maximal entropy of entanglement and are considered the “gold standard” of the resource entanglement, lying at the heart of many applications in quantum information processing like quantum key distribution [7] or entanglement swapping [28]. There exist many bases of maximally entangled states for bipartite systems. A standard construction for a basis of Bell states for systems of arbitrary dimension d is based on the Weyl-Heisenberg operators (see, e.g., Ref.[1]), which we define as

Definition 1.1.4 (Weyl-Heisenberg Operators)

$$W_{k,l} := \sum_{j=0}^{d-1} \omega^{jk} |j\rangle\langle j+l|, \quad k, l \in \mathcal{Z}_d. \quad (1.1)$$

Here, we denote by $\mathcal{Z}_d := \mathbb{Z}/d\mathbb{Z}$ the ring of integers modulo d and $\omega := e^{2\pi i/d}$. These unitary operators form a group, the so-called Weyl-Heisenberg group $\mathcal{V}_d = \{\omega^i W_{k,l}\}_{i,k,l \in \mathcal{Z}_d}$ [29], which is induced by the Weyl relations, i.e.,

$$\begin{aligned} W_{k_1,l_1} W_{k_2,l_2} &= \omega^{l_1 k_2} W_{k_1+k_2, l_1+l_2}, \\ W_{k,l}^\dagger &= \omega^{kl} W_{-k,-l} = W_{k,l}^{-1}. \end{aligned} \quad (1.2)$$

Identifying two elements of the group if they differ only by a phase factor yields the group of principal interest for this work, the projective Weyl-Heisenberg group

$$\mathcal{W}_d = \{W_{k,l}\}_{k,l \in \mathbb{Z}_d}. \quad (1.3)$$

For the technical details to define this group formally, see Sec.1.1.4.

Applying these d^2 unitary operators locally to one side of a maximally entangled state defines d^2 linearly independent, maximally entangled qudit states [30]:

Definition 1.1.5 (Standard Bell Basis)

The standard Bell basis for a bipartite system of two qudits of dimension d is defined by the following basis states:

$$|\Omega_{k,l}\rangle := (W_{k,l} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle := (W_{k,l} \otimes \mathbb{1}_d) \frac{1}{\sqrt{d}} \sum_{i \in \mathbb{Z}_d} |i\rangle \otimes |i\rangle, \quad (1.4)$$

$$\Omega_{k,l} \equiv |\Omega_{k,l}\rangle \langle \Omega_{k,l}|, \quad k, l \in \mathbb{Z}_d. \quad (1.5)$$

For $d = 2$, $|\Omega_{0,0}\rangle$ is often denoted by $|\Phi^+\rangle$:

$$|\Phi^+\rangle := 1/\sqrt{2} (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle). \quad (1.6)$$

In practical implementations involving entanglement, interactions with the environment inevitably introduce noise, which degrades this resource and renders operations and outcomes prone to errors. To this date, correcting or mitigating these errors are crucial challenges to implementing quantum algorithms with contemporary hardware [31–33]. Entanglement distillation (also called entanglement purification) [34] is an alternative method to quantum error correction that addresses this problem directly by reverting errors [23, 35, 36]. Assuming two distant parties that share multiple copies of entangled but noisedegraded (mixed) states, an entanglement distillation protocol converts many weakly entangled copies into a generally smaller number of strongly entangled, ideally maximally entangled pairs. Following the distantlab paradigm, only local quantum operations and classical communication (LOCC) are allowed. Formally, following Ref.[21], we define:

Definition 1.1.6 (Entanglement Distillation, Distillability, Distillation Rate)

Entanglement distillation describes any LOCC protocol that transforms N copies of a resource state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ to $M \leq N$ copies of the maximally entangled qubit state $\Phi^+ \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with nonzero probability:

$$\rho^{\otimes N} \xrightarrow{\text{LOCC}} (\Phi^+)^{\otimes M}. \quad (1.7)$$

A state ρ is distillable if in the asymptotic limit $N \rightarrow \infty$, a distillation rate $R := \frac{M}{N}$ larger than zero can be achieved via entanglement distillation. If no such operations exist, ρ is called undistillable.

Note that any maximally entangled state of finite dimensions other than Φ^+ , e.g., any $\Omega_{k,l}$ (Def.1.1.5), can be used as target state for distillation, as all Bell states of finite dimension are interconvertible via LOCC in the asymptotic limit [1]. Originally introduced for two-level (qubit) systems [34], many distillation protocols have been found that enhance efficiency for certain noise regimes or generalize such methods to d -level (qudit) systems. Recurrence protocols [34, 37–43] act on a small set of copies to increase the fidelity with a Bell state in an iterative way. They allow for the distillation of highly noisy states but often exhibit poor efficiency in the asymptotic many-copy limit [44]. Hashing or breeding protocols [45–47] act on numerous input states collectively, yielding a positive rate of maximally entangled output states, but generally require low levels of noise [44].

As the set of separable states (SEP) is closed under LOCC (cf. Ref.[48]), any state that can be used for entanglement distillation must be entangled. Surprisingly, the converse is not true since there exist states that are entangled but are known to be undistillable [49–51]. These states are called bound entangled, reflecting the fact that the resource entanglement is “bound” to the states in contrast to the “free” entanglement of distillable states.

Definition 1.1.7 (Free and Bound Entanglement)

A bipartite quantum state ρ is called free entangled (denoted $\rho \in \text{FE}$) if it is distillable. If ρ is entangled and undistillable, it is called bound entangled ($\rho \in \text{BE}$).

To this date, all known bound entangled states are entangled states with a positive partial transposition ($\rho \in \text{PPT} \cap \text{ENT} \implies \rho \in \text{BE}$ [51]). It is not clear whether there exist bound entangled states with negative, i.e., non-positive partial transposition (NPT) [52, 53]. This problem is called the distillability problem [54]:

Definition 1.1.8 (Distillability Problem)

Does the set equivalence

$$\text{SEP} \cup \text{BE} \equiv \text{PPT} \tag{1.8}$$

hold generally?

Bound entanglement implies an irreversibility in the resource theory of entanglement, as entanglement is required to form these states via LOCC but cannot be extracted [14]. Consequently, bound entanglement lies at the heart of two major problems of quantum information theory: the distillability problem, by the potential existence of NPT bound entangled states, and the separability problem, because PPT bound entangled states are precisely those for which no general entanglement criterion is known. Beyond this foundational interest, bound entanglement has been shown to possess operational value in certain cryptographic and prepareandmeasure scenarios [55–57], and it therefore raises the practical question of when bound entanglement can be harnessed for information tasks.

These open questions are clearly an indication of our lack of understanding of quantum entanglement, both from an operational and a principle point of view. They motivate a focused, operationally oriented investigation into structured families of states that

are both analytically tractable and experimentally relevant. Belldiagonal systems, i.e., mixtures of Bell states, form a structured yet sufficiently rich family of bipartite states that lend themselves to systematic investigation. We refer to the subclass of mixed standard Bell basis states (Def.1.1.5), constructed via WeylHeisenberg operators, as the standard Belldiagonal system.

Definition 1.1.9 (Standard Bell-diagonal States, Magic Simplex)

$$\mathcal{M}_d := \left\{ \rho = \sum_{k,l \in \mathbb{Z}_d} c_{k,l} \Omega_{k,l} \mid \sum_{k,l \in \mathbb{Z}_d} c_{k,l} = 1, c_{k,l} \geq 0 \right\}. \quad (1.9)$$

Corresponding states can be viewed as maximally entangled states subject to local probabilistic Weyl errors, which makes the family operationally relevant, for example, if one part of the Bell state is sent through a noisy quantum channel [58]. The special algebraic properties of the Weyl operators forming the group $\mathcal{W}_d$ (1.3) are inherited by the standard Bell states and, in turn, by the standard Belldiagonal system. These properties induce distinctive entanglement structures and symmetries that can be exploited in both analytical and numerical studies [30]. Moreover, since the parametrization of mixed states by mixing probabilities naturally yields a geometric simplex, the standard Belldiagonal system is also known as the Magic Simplex [59].

Due to being challenging to detect, bound entanglement can be considered a rare phenomenon, especially in high dimensional systems [50]. However, initial investigations of low-dimensional Bell-diagonal states demonstrate that a significant share of Bell-diagonal states of bipartite qutrits is bound entangled [60] and that such states can be detected in experiments [61].

The foundational and operational relevance of bound entanglement and the combination of analytical and numerical accessibility of Bell-diagonal systems, including a significant amount of bound entanglement, motivate the scope of this thesis: Distillability and bound entanglement of Bell-diagonal states.

1.2. Scope of the dissertation

The overarching goal of this dissertation is to deepen our understanding of the entanglement structure of bipartite Bell-diagonal quantum states. Particular attention lies on bound entanglement and exploring how structural properties can be harnessed for entanglement distillation. More precisely, the central research question is:

How can the entanglement structure of bipartite Bell-diagonal systems be characterized, and in which way are their properties related to distillability and bound entanglement?

Answering this question requires overcoming three main challenges that we address

through a series of self-contained publications, each contributing analytical insights, reproducible numerical methods, or constructive protocol designs. In detail, these challenges are:

- **Establishing Bell-diagonal systems as a suitable platform for investigations.**

A prerequisite is to verify that Bell-diagonal systems contain a sufficiently rich set of bound entangled states to serve as a reliable foundation for systematic analysis. This involves developing and implementing analytical and numerical methods to classify large ensembles of states according to separability (SEP), entanglement (ENT), positive or negative partial transposition (PPT/NPT), and bound or free entanglement (BE/FE). Since only PPT-entangled states are known to be bound entangled, we focus on detecting PPT-bound entangled states. For local dimensions $d \geq 3$, no necessary and sufficient entanglement criterion is available, so a combination of methods must be employed to detect it. Comparing the performance of these methods for the detection of PPT entanglement strengthens classification capabilities and provides insights into the structure of PPT-entangled states. First, we focus on the standard Bell-diagonal system based on the Heisenberg-Weyl construction (Def.1.1.5). Building on prior investigations regarding analytical structure [30, 59] and numerical analyses [60, 62], we leverage the group properties of the system for enhancing the entanglement classification of its states. Additionally, we analyze other systems of non-standard maximally entangled bases and compare them to the magic simplex. In Sec.3.1.1 and Sec.3.1.2 we summarize and discuss the results of related publications. We show that Bell-diagonal systems are well suited for investigations within the scope of this thesis, providing us with a solid data foundation to analyze the entanglement structure and distillability of bipartite qudits.

- **Identifying characterizing properties of entanglement classes.**

Once a data foundation of entanglement classified states is established, the next step is to uncover structural features that correlate with separability, distillability and bound entanglement in Bell-diagonal systems. This requires comparing different Bell-diagonal systems and searching for numerical and analytical properties that distinguish separable, PPT-entangled, and NPT-entangled states. Of particular interest is the group structure of $\mathcal{W}_d$ (1.3) underlying the standard Bell basis constructed from WeylHeisenberg operators, which previous investigations [30] suggest plays a decisive role in shaping entanglement properties. In the publications summarized in Sec.3.1.3 we confirm this hypothesis, demonstrating that the group properties have significant influence on the entanglement structure of Bell-diagonal systems and that group-related properties of standard Bell-diagonal states contain significant information about their entanglement class.

Evaluating information-theoretic and entropic quantities that connect directly to distillation and related quantum information tasks presents a way towards additional characterizing factors. Generalized divergences offer a unifying framework to capture

various information or entropic measures [48]. Identifying such quantities with potential relevance for bound entanglement and evaluating them for Bell-diagonal states contributes to a characterization of distillable states in the magic simplex and other Bell-diagonal systems. Besides the asymptotic quantities related to the von Neumann entropy, like the coherent or the private information (cf. [1]), also the so-called one-shot regime [48, 63] corresponding to finite resources is considered. Quantities that can be related to the operational rates of distillation tasks, like quantum hypothesis-testing information [64] or the smooth max-mutual information [65], are expected to contain relevant information about the entanglement class. However, these measures are generally challenging to compute. Two publications summarized in Sec.3.1.5 address this challenge. One provides a method to prove local invariances for families of information quantities, and the other presents an explicit algorithm to compute the smooth max-mutual information by employing semidefinite programming (cf. [48]).

- **Leveraging structural properties for entanglement distillation.**

Finally, we investigate how the identified characterizing properties of the entanglement structure can be related to distillability. On the one hand, a better understanding of this relation contributes to our understanding of bound entanglement. On the other hand, such properties may be leveraged for the construction of effective entanglement distillation protocols.

Combining shift and phase errors, the Weyl operators can be interpreted as error operators forming a so-called nice error basis [66]. Considering the multi-copy setting of entanglement distillation (Def.1.1.6), the set of N -copy Weyl errors can be defined as follows:

Definition 1.2.1 (Weyl errors)

$$\mathcal{E}_N := \left\{ W(e) \mid e = (\vec{k}, \vec{l}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \right\} := \left\{ \bigotimes_{n=1}^N W_{k_n, l_n} \mid k_n, l_n \in \mathcal{Z}_d \right\}. \quad (1.10)$$

The group structure of the projective Weyl-Heisenberg group (1.3) naturally extends to the set of Weyl errors $\mathcal{E}_N$. A major result presented in the publications summarized in Sec.3.1.4 is that the common group structure of Bell states and Weyl errors can be utilized for highly effective and efficient entanglement distillation via stabilizer operations. A stabilizer is an abelian subgroup $S \subset \mathcal{E}_N$. Corresponding operations like stabilizer encodings or stabilizer measurements (cf. Ref.[41, 67]) are therefore inherently group-related and can be successfully applied in various fields, including measurement-based quantum computing [4] or quantum error correction [68, 69]. Building on a construction to use stabilizer operations for entanglement distillation [41], we leverage the shared group structure of states and operations to determine the action of stabilizer operations on Bell-diagonal states such that these operations allow the design of optimized distillation protocols. For two copies ($N = 2$) and prime dimensions d , we present and analyze a new entanglement distillation protocol

that exhibits superior performance compared to other protocols, especially for highly mixed states. A quantitative analysis of methods differentiating the necessarily mixed bound entangled states [70, 71] from similarly mixed but distillable states not only advances the theoretical understanding of entanglement but also holds practical relevance for quantum technologies operating in noisy or error-prone environments.

Overcoming these challenges, we demonstrate that Bell-diagonal systems provide both a structured setting for the study of bound entanglement and a fertile ground for stabilizer-based distillation protocols. Therefore, this work links the theoretical foundations of entanglement with operational progress in quantum communication and cryptographic protocols.

The remainder of this thesis is organized as follows. In Chapter 2, we present the publications and provide information about the author contributions and publication status. Chapter 3 summarizes and discusses the main reported results in Sec.3.1 and the related methodology in Sec.3.2. We conclude in Sec.3.3 with a discussion of the combined insights and their relevance for the goal of this thesis.

2. Publication overview

In this chapter, all publications created during the dissertation project are listed including the author's contribution, the publication status, contributing authors, publication information, and the abstract. We present the full publications in their published form in appendix B.

There are three types of contribution of the author to the publications, which are understood as follows:

- **First Author:** Contributions include the development of the research idea, the development of analytical and numerical methods, the execution of analytical and numerical analyses, the development of software and the generation of data, the verification and validation of results, designing, writing and reviewing the paper, submitting the paper to publishing journals, handling the review process, serving as corresponding author for the publication, and presenting the work at conferences and events.
- **Equally Contributing Author:** All authors contributed equally to the work. Contributions include the development of the research idea, the development of analytical and numerical methods, the execution of analytical and numerical analyses, the development of software and the generation of data, the verification and validation of results, designing, writing and reviewing the paper, submitting the paper to publishing journals, handling the review process, and presenting the work at conferences and events.
- **Contributing Author:** Contributions include the development of the research idea, supporting the development of analytical and numerical methods and the execution of analyses, the generation of data, the verification and validation of results, reviewing the paper, and presenting the results at conferences and events.

The status of the publications as of the thesis submission date is as follows:

- **Published:** The article has been accepted and published in a peer-reviewed journal.
- **Preprint, submitted to journal:** The article has been published on a preprint server and is currently in the review process of a peer-reviewed journal.

Funding and license: The research related to these publications was funded in whole, or in part, by the Austrian Science Fund (FWF) [10.55776/P36102]. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from these submissions.

The publications are ordered by the nature of their contribution as discussed in the synopsis of publications in Sec.3.1.

2.1. Entanglement classification of Bell-diagonal states

2.1.1. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits [72]

Contribution: First Author.

Status of the Publication: Published.

C. Popp, Journal of Open Source Software 8, 4924 (2023)

In the field of quantum information and technology, entanglement in quantum systems called qudits is regarded as resource for quantum-based or quantum-assisted information processing tasks. It allows new ways of information processing like quantum computation or quantum teleportation and provides possibilities for speedups in a variety of algorithms with applications like search or optimization. Despite its theoretical and practical relevance, there is no general method to determine whether a given quantum state is entangled or not due to a special and hard to detect form of entanglement called bound entanglement. Bipartite Bell states are maximally entangled states of two qudits and are of high relevance for application in quantum technologies. Mixtures of those states can be entangled, including its bound form, or separable. Leveraging their special properties, Bell states can be numerically generated and analyzed for their entanglement properties by various methods implemented in this Julia package.

2.1.2. Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits [73]

Contribution: First Author.

Status of the Publication: Published.

C. Popp and B. C. Hiesmayr, Scientific Reports 12, Number: 1 Publisher: Nature Publishing Group, 12472 (2022).

With a probability of success of 95% we solve the separability problem for Bell diagonal qutrit states with positive partial transposition (PPT). The separability problem, i.e. distinguishing separable and entangled states, generally lacks an efficient solution due to the existence of bound entangled states. In contrast to free entangled states that can be used for entanglement

distillation via local operations and classical communication, these states cannot be detected by the PeresHorodecki criterion or PPT criterion. We analyze a large family of bipartite qutrit states that can be separable, free entangled or bound entangled. Leveraging a geometrical representation of these states in Euclidean space, novel methods are presented that allow the classification of separable and bound entangled Bell diagonal states in an efficient way. Moreover, the classification allows the precise determination of relative volumes of the classes of separable, free and bound entangled states. In detail, out of all Bell diagonal PPT states $81.0\% \pm 0.1\%$ are determined to be separable while $13.9 \pm 0.1\%$ are bound entangled and only $5.1 \pm 0.1\%$ remain unclassified. Moreover, our applied criteria are compared for their effectiveness and relation as detectors of bound entanglement, which reveals that not a single criterion is capable to detect all bound entangled states.

2.1.3. Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts [74]

Contribution: First Author.

Status of the Publication: Published.

C. Popp and B. C. Hiesmayr, Scientific Reports 13, Number: 1 Publisher: Nature Publishing Group, 2037 (2023).

We compare the classification as entangled or separable of Bell diagonal bipartite qudits with positive partial transposition (PPT) and their properties for different dimensions. For dimension $d \geq 3$, a form of entanglement exists that is hard to detect and called bound entanglement due to the fact that such entangled states cannot be used for entanglement distillation. Up to this date, no efficient solution is known to differentiate bound entangled from separable states. We address and compare this problem named separability problem for a family of bipartite Bell diagonal qudits with special algebraic and geometric structures and applications in quantum information processing tasks in different dimensions. Extending analytical and numerical methods and results for Bell diagonal qutrits ($d = 3$), we successfully classify more than 75% of representative Bell diagonal PPT states for $d = 4$. Via those representative states we are able to estimate the volumes of separable and bound entangled states among PPT ququarts ($d = 4$). We find that at least 75.7% of all PPT states are separable, 1.7% bound entangled and for 22.6% it remains unclear whether they are separable or bound entangled. Comparing the structure of bound entangled states and their detectors, we find considerable differences in the detection capabilities for different dimensions

and relate those to differences of the Euclidean geometry for qutrits ($d = 3$) and ququarts ($d = 4$). Finally, using a detailed visual analysis of the set of separable and bound entangled Bell diagonal states in both dimensions, qualitative observations are made that allow to better distinguish bound entangled from separable states.

2.2. Relation of group and entanglement structure in Bell-diagonal systems

2.2.1. Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states [75]

Contribution: First Author.

Status of the Publication: Published.

C. Popp and B. C. Hiesmayr, New Journal of Physics 26, Publisher: IOP Publishing, 013039 (2024).

Bell states are of crucial importance for entanglement based methods in quantum information science. Typically, a standard construction of a complete orthonormal Bell-basis by WeylHeisenberg operators is considered. We show that the group structure of these operators has strong implication on error correction schemes and on the entanglement structure within Bell-diagonal states. In particular, it implies an equivalence between a Pauli channel and a twirl channel. Interestingly, other complete orthonormal Bell-bases do break the equivalence and lead to a completely different entanglement structure, for instance in the share of positive partial transposition (PPT)-entangled states. In detail, we find that the standard Bell basis has the highest observed share on PPT-states and PPT-entangled states compared to other Bell bases. In summary, our findings show that the standard Bell basis construction exploits a very special structure with strong implications to quantum information theoretic protocols if a deviation is considered.

2.2.2. A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria [76]

Contribution: Contributing Author.

Status of the Publication: Preprint, submitted for publishing.

T. C. Sutter, C. Popp and B. C. Hiesmayr, arXiv:2508.18393 [quant-ph], (2025).

Entanglement is a key feature in many quantum technologies, including secure communication protocols and quantum computing. However, detecting it in mixed quantum states remains a challenging task. While the PPT and realignment criteria are well-established tools for entanglement detection in general, and are especially effective in Bell-diagonal states, their connection to the underlying group structure of this state family has not been fully explored. In this work, we analyze the PPT and realignment criteria for Bell-diagonal states from a group-theoretic point of view. Our results demonstrate that the group structure of Bell-diagonal states provides a clear framework for analyzing and computing these two entanglement detection criteria. This unified perspective offers new insights into the mathematical and physical properties of entanglement in structured quantum systems, like the connection of the PPT criterion to experimentally accessible entanglement witnesses, thus opening a path for extending the group formalism to more complex quantum systems.

2.3. Entanglement distillation

2.3.1. A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits [77]

Contribution: First Author.

Status of the Publication: Published.

C. Popp, T. C. Sutter and B. C. Hiesmayr, Quantum 9, Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 1945 (2025).

Entanglement distillation, the process of converting weakly entangled states into maximally entangled ones using Local Operations and Classical Communication (LOCC), is pivotal for robust entanglement-assisted quantum information processing in error-prone environments. A construction based on stabilizer codes offers an effective method for designing such protocols. By analytically investigating the effective action of stabilizer protocols for systems of prime dimension d , we establish a standard form for the output states of recurrent stabilizer-based distillation. This links the properties of input states, stabilizers, and encodings to the properties of the protocol. Based

on those insights, we present a novel two-copy distillation protocol, applicable to all bipartite states in prime dimension, that maximizes the fidelity increase per iteration for Bell-diagonal states. The power of this framework and the protocol is demonstrated through numerical investigations, which provide evidence for superior performance in terms of efficiency and distillability of low-fidelity states compared to other well-established recurrence protocols. By elucidating the interplay between states, errors, and protocols, our contribution advances the systematic development of highly effective distillation protocols, enhancing our understanding of distillability.

2.3.2. Low-Fidelity Entanglement Distillation with FIMAX [78]

Contribution: First Author.

Status of the Publication: Published.

C. Popp, T. C. Sutter and B. C. Hiesmayr, International Journal of Quantum Information 23, Publisher: World Scientific Publishing Co., 2550017 (2025).

Uncontrolled interactions with the environment introduce errors that remain a significant challenge to the reliability of quantum technologies using entanglement. An essential method to overcome or mitigate these errors is entanglement distillation, the transformation of multiple copies of weakly entangled states into a smaller number of approximately maximally entangled states. We present a comparative analysis of the stabilizer-based two-copy entanglement distillation protocol, FIMAX, against other recurrent two-copy protocols, including ADGJ, DEJMPS, P1-P2, and the generalized BBPSSW protocol. We focus on low-fidelity bipartite quantum states in dimensions $d = 2$ and $d = 3$, which are particularly challenging to distill. Our findings demonstrate that FIMAX exhibits superior performance for these states. While other protocols struggle with highly noisy states, FIMAX successfully distills entanglement even when the initial state quality is severely compromised. These results highlight the protocol's capability to address the effects of environmental noise, advancing the robustness and scalability of quantum technologies leveraging entanglement distillation.

2.4. Properties and computation of distillation-relevant information measures

2.4.1. Local Invariance of Divergence-Based Quantum Information Measures [79]

Contribution: First Author.

Status of the Publication: Published.

C. Popp, T. C. Sutter and B. C. Hiesmayr, Entropy 27, Publisher: Multidisciplinary Digital Publishing Institute, 1051 (2025).

Quantum information quantities, such as mutual information and entropies, are essential for characterizing quantum systems and protocols in quantum information science. In this contribution, we identify types of information measures based on generalized divergences and prove their invariance under local isometric or unitary transformations. Leveraging the reversal channel for local isometries together with the data-processing inequality, we establish invariance for information quantities used in both asymptotic and one-shot regimes without relying on the specific functional form of the underlying divergence. These invariances can be applied to improve the computation of such information quantities or optimize protocols and their output states, whose performance is determined by some invariant measure. Our results improve the capability to characterize and compute many operationally relevant information measures with application across the field of quantum information processing.

2.4.2. Computation of the Smooth Max-Mutual Information via Semidefinite Programming [80]

Contribution: First Author.

Status of the Publication: Preprint, submitted for publishing.

C. Popp, T. C. Sutter and B. C. Hiesmayr, arXiv:2509.07743 [quant-ph], Oct. 2025.

We present an iterative algorithm based on semidefinite programming (SDP) for computing the quantum smooth max-mutual information $I_{\max}^{\epsilon}(\rho_{AB})$ of bipartite quantum states in any dimension. The algorithm is accurate if a rank condition for marginal states within the smoothing environment is satisfied and provides an upper bound otherwise. Central to our method is a novel SDP, for which we establish primal and dual formulations and prove strong duality. With the direct application of bounding the one-shot distillable key of a quantum state, this contribution extends SDP-based techniques in quantum information theory. Thereby it improves the capabilities

to compute or estimate information measures with application to various quantum information processing tasks.

2.5. Review of bipartite bound entanglement

2.5.1. Bipartite Bound Entanglement [21]

Contribution: Equally Contributing Author.

Status of the Publication: Published.

B. C. Hiesmayr, C. Popp and T. C. Sutter, International Journal of Quantum Information 23, Publisher: World Scientific Publishing Co., 2530003 (2025)

Bound entanglement is a special form of quantum entanglement that cannot be used for distillation, i.e. the local transformation of copies of arbitrarily entangled states into a smaller number of approximately maximally entangled states. Implying an inherent irreversibility of quantum resources, this phenomenon highlights the gaps in our current theory of entanglement. This review provides a comprehensive exploration of the key findings on bipartite bound entanglement. We focus on systems of finite dimensions, an area of high relevance for many quantum information processing tasks. We elucidate the properties of bound entanglement and its interconnections with various facets of quantum information theory and quantum information processing. The paper highlights areas where our understanding of bound entangled states, particularly in their detection and characterization, remains underdeveloped. By highlighting the need for further research into this phenomenon and underscoring relevant open questions, this paper invites researchers to unravel its relevance for our understanding of entanglement in Nature and how this resource can most effectively be used for applications in quantum technology.

2.6. Dissipative quantum neural networks

2.6.1. The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks [81]

Contribution: Contributing Author.

Status of the Publication: Preprint, submitted for publishing.

T. C. Sutter, C. Popp and B. C. Hiesmayr, arXiv:2502.09526 [quant-ph], Feb. 2025.

Combining machine learning and quantum computation is a potential path towards powerful applications on quantum devices. Regarding this, quantum neural networks are a prominent approach. In this work, we present a novel architecture for dissipative quantum neural networks (DQNNs) in which each building block can implement any quantum channel, thus introducing a clear notion of universality suitable for the quantum framework. To this end, we reformulate DQNNs using isometries instead of conventionally used unitaries, thereby reducing the number of parameters in these models. We furthermore derive a versatile one-to-one parametrization of isometries, allowing for an efficient implementation of the proposed structure. Focusing on the impact of different cost functions on the optimization process, we numerically investigate the trainability of extended DQNNs. This unveils significant training differences among the cost functions considered. Our findings facilitate both the theoretical understanding and the experimental implementability of quantum neural networks.

3. Synopsis and conclusion

In this chapter we summarize the results, methods, and implications of the scientific publications listed in the previous section. First, we discuss the results in Sec.3.1 before describing the methodology in Sec.3.2. Finally, we conclude and discuss the combined findings and their contribution to the central aim of this thesis in Sec.3.3.

3.1. Synopsis of results

The results generated as part of the project contribute to the fields of bound entanglement and distillation of bipartite qudits with special focus on Bell-diagonal systems. Here, we present a concise overview of the key findings and elucidate their alignment with the thesis’ central research questions, evaluate their contributions to the scientific community, and identify unresolved challenges alongside prospective directions for future research.

3.1.1. Developing numerical methods for entanglement classification and distillation of Bell-diagonal states

Due to the lack of a complete analytical characterization, a key requirement for the investigation of bound entanglement and distillability in Bell-diagonal systems is a suitable data foundation of entanglement-classified Bell-diagonal states. For this purpose, we create, release, and continue developing the open-source and peer-reviewed Julia software package “BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits” (cf. Sec.2.1.1 [72]).

The package provides methods to generate and analyze data that represents states of Bell-diagonal systems. Moreover, it implements general methods for entanglement classification of all bipartite qudits. The following features are used within the scope of the project:

- Efficient state sampling of Bell-diagonal and general bipartite states of arbitrary dimension.
- Entanglement classification of separable, PPT-entangled and NPT-entangled states via several entanglement criteria.
- Numerical creation of Bell-diagonal entanglement witnesses (Def.A.2.15).
- Implementation of several entanglement distillation protocols, including the newly developed stabilizer-based protocol “FIMAX” (Sec.2.3.1 [77]).

-
- Evaluation of quantum information and entropic quantities related to entanglement and secret-key distillation.

These results provide the numerical foundation for all following analyses. Releasing the software as an official peer-reviewed and open-source package enables the reproducibility of any reported results and makes the methods available for further scientific investigations in the field. More details about the implementation are described in Sec.3.2.1.

3.1.2. Validating Bell-diagonal systems for the analysis and characterization of bound entanglement

Combining known analytical properties of Bell-diagonal systems [30] and the numerical methods developed in the work of Sec.2.1.1 [72], the entanglement structure of Bell-diagonal systems is analyzed to validate the suitability of such systems for investigating bound entanglement.

The publication “Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits” (cf. Sec.2.1.2 [73]) provides an extensive analysis of the system of Bell-diagonal bipartite qutrits, classifying the significant majority ($> 95\%$) of states as separable, PPT-bound entangled or NPT-free-entangled. On the one hand, it shows that a significant share ($13.9\% \pm 0.1\%$) of bipartite Bell-diagonal qutrits with positive partial transposition is entangled and therefore bound entangled. On the other hand, the small number of unclassified PPT states verifies the effectiveness of methods leveraging symmetries of Bell-diagonal states to detect entanglement and separability in such systems.

The subsequent analysis is summarized in “Comparing bound entanglement of Bell-diagonal pairs of qutrits and ququarts” (cf. Sec.2.1.3 [74]). Comparing the results for $d = 3$ to $d \geq 4$, we analyze bipartite Bell-diagonal systems of dimension four and higher with similar methods. We observe that the relative amount of PPT states decreases quickly with higher dimensions. The restriction to systems with local dimension up to five for this project is therefore justified operationally since the potential amount of PPT-bound entanglement becomes insignificant with growing system size. The numerical analyses show that the entanglement classification methods are not as effective as in the less-dimensional case. Still, bound entangled states are detected among the PPT states (1.7%). While a decrease in performance is expected for the numerical methods due to the increase in computational complexity with the dimension, also the analytical criteria are considerably less effective for bipartite ququarts compared to qutrits. An analysis of classified states for dimension $d = 3$ and $d = 4$ shows that a large amount of relevant information for the classification of states as PPT or NPT is contained in special projections of rank d . These projections are related to subgroups of the projective Weyl-Heisenberg group (1.3), and the analysis indicates qualitatively that such projections are also relevant for differentiating separable and PPT bound entangled states.

Together, these results imply that Bell-diagonal systems are well suited to investigate bound entanglement. They contain a significant share of PPT bound entangled states, which can be detected by a combination of numerical and analytical methods. The

reproducible foundation of data representing entanglement-classified states can be further processed to find characterizing properties of separable, PPT-entangled, and NPT-entangled Bell-diagonal states. The methods used to generate the results reported in this section are described in detail in Sec.3.2.2.

3.1.3. Identifying the underlying group structure of Bell states as a relevant factor for the entanglement structure of Bell-diagonal systems

The data of entanglement-classified Bell-diagonal states shows that the group structure of the Weyl-Heisenberg operators (1.3) is related to the entanglement structure and classification of Bell-diagonal states as separable, PPT-entangled, or NPT-entangled. We take up this observation in the publication “Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states” (cf. Sec.2.2.1 [75]), which analyzes implications of the group structure underlying the standard Bell basis (cf. Def.1.1.5). We show that this group structure of standard Bell states implies a stabilizing property of so-called Weyl twirl operators for standard Bell states. This special property has several implications and applications. First, it implies that the projection of any bipartite state to the set of standard Bell-diagonal states is equivalent to twirling, i.e., randomly applying the local Weyl twirl operators. As a consequence, any state can be projected to Bell-diagonal form by LOCC operations. Hence, twirling conserves separability, allowing us to parameterize and sample separable Bell-diagonal states and generate optimal Bell-diagonal entanglement witnesses more efficiently. Second, the stabilizing property enables a simple error-detecting and correcting scheme for the maximally entangled qudit state.

We demonstrate the significance of the Weyl-Heisenberg group for these properties and implications by comparison with so-called generalized Bell bases that do not exhibit a similar group structure. Introducing additional parameterizable relative phases between the standard Bell states, an orthonormal basis of maximally entangled bipartite qudit states is created for which the group relation of the standard basis is broken. Both applications regarding the Weyl twirl and the stabilizing property depend on the algebraic structure of the standard Bell states and do not hold if that structure is lost. We also show that the entanglement structure of Bell-diagonal systems strongly depends on the algebraic structure of its Bell states. If the group structure of standard Bell states is diminished, the corresponding system of Bell-diagonal states generally has fewer PPT-entangled states. The publication “A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria” (cf. Sec.2.2.2 [76]) provides further insights about the relation of group and entanglement structure. In this work, the PPT and realignment entanglement criteria [15–17] are analyzed for Bell-diagonal states from a group-theoretic perspective. Deriving expressions for both criteria in terms of group-related objects (e.g., subgroups, cosets, etc.) underscores analytically the observations regarding the relation of group and entanglement structure made in the publications of Sec.2.1.2 [73], Sec.2.1.3 [74] and Sec.2.2.1 [75].

In summary, the following findings are especially significant for the analysis of bound entanglement in Bell-diagonal systems: First, PPT bound entanglement occurs more

frequently for mixtures of Bell states that are related by operators of the Weyl-Heisenberg group compared to other Bell-diagonal systems. Second, the entanglement criteria that allow for the effective detection of PPT entangled states can be expressed solely by group-related objects if they are applied to standard Bell-diagonal states. Finally, these findings identify the group structure of the Weyl-Heisenberg operators as a relevant factor for the entanglement structure of Bell-diagonal systems. Sec.3.2.3 discusses the related methodology.

3.1.4. Leveraging the group structure of standard Bell states for entanglement distillation via stabilizer-based methods

Having identified the algebraic structure of Bell states as a significant influencing factor for the entanglement structure of bipartite Bell-diagonal systems, the question arises whether this relation also holds for the entanglement distillability.

Using Bell-diagonal input states and the target state to be a standard Bell state (cf. Def.1.1.5, Def.1.1.6 and the comment below Def.1.1.6), both the target states and the operators to be interpreted as error operators (Def.1.2.1), i.e., the Weyl operators, have the same underlying group structure. The intuition is that this common structure can be leveraged for entanglement distillation.

The publication “A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits” (cf. Sec.2.3.1 [77]) builds on and extends a known construction [41] to use stabilizer codes for entanglement distillation. Stabilizers are commutative subgroups of the group of error operators and can be used to define stabilizer codes for multi-qudit systems [66, 69] (cf. A.2.17). Leveraging the connecting group structure of Bell states and stabilizer operations, we derive a general output form of multi-copy standard Bell-diagonal states under the action of stabilizer-based LOCC operations. In this form, the roles of the input state, stabilizer, encoding, and measurement outcomes are explicitly visible, allowing the optimization of stabilizer-based operations for the task of entanglement distillation. We present a new recurrent entanglement distillation protocol called FIMAX that acts on two copies of prime-dimensional bipartite qudits. The protocol is proven to maximally increase the fidelity of Bell-diagonal states in each iteration among all protocols of this stabilizer-based construction. FIMAX demonstrates superior performance regarding the distillation efficiency and distillability compared to other recurrent protocols presented in Refs.[34, 37–40]. Curiously, this advantage persists even if non-Bell-diagonal states requiring a prior twirl to Bell-diagonal form are used for distillation.

The subsequent publication “Low-Fidelity Entanglement Distillation with FIMAX” (cf. Sec.2.3.2 [78]) investigates the distillation capability of FIMAX extensively and compares its performance to other protocols (see Refs.[34, 37–40]). Focusing on bipartite qubits and qutrits, the new protocol demonstrates best-in-class performance for the capability to distill highly noisy states. The distillation effectivity exceeds those of other protocols, especially for the standard Bell-diagonal states, but also for randomly generated pure states and generalized Bell-diagonal states lacking the special group properties of the standard Bell-diagonal system. Notably, FIMAX is the only one among the investigated protocols that can distill any of the considered bipartite qutrit states that have less than

$1/d$ overlap with all standard Bell-states.

Combining the positive results regarding the distillation efficiency and effectivity, these works demonstrate that utilizing the algebraic structure of Bell states with stabilizer-based methods like FIMAX provides both a highly effective and resource-efficient approach to entanglement distillation. Moreover, the general analytical results regarding the action of stabilizer-based operations on Bell states are the foundation for enabling the application of the stabilizer formalism to other distillation tasks like, e.g., secret-key distillation (cf. [48]). Methodological aspects related to these results are discussed in Sec.3.2.4.

3.1.5. Laying the foundation for the generalization and extension of stabilizer-based distillation

With the established power of stabilizer-based methods for entanglement distillation by increasing the fidelity with the maximally entangled state, a natural question arising is whether these methods can also be applied to other tasks. Secret-key distillation is a task of practical relevance for cryptographic applications [82]. It is also relevant for quantifying entanglement, as it defines the entanglement measure “distillable secret-key” [22]. This measure is of special interest in the context of bound entanglement, as bound entangled states have zero distillable entanglement by definition but may have a non-zero distillable key [55, 83].

Guided by the operational principle of FIMAX to maximize the iterative increase of the fidelity, one way to extend stabilizer-based methods to other distillation tasks like secret-key distillation is to consider other quantities that are iteratively increased. For example, instead of the fidelity with the Bell state, one might consider the private information (cf. Ref.[84]) of the state to maximize the increase of a state’s capability to distill secret-key states in the asymptotic limit of infinite copies. Many other information quantities exist that measure a state’s capability for some information processing task (see Ref.[48] for a comprehensive overview). To optimize stabilizer operations for the LOCC transformation of input states into target states that are suitable for a certain task, the corresponding information measure needs to be evaluated. This, however, can be analytically and numerically difficult.

The publication “Local Invariance of Divergence-Based Quantum Information Measures” (cf. Sec.2.4.1 [79]) addresses this problem. In this work, we investigate several information and entropic quantities that are operationally relevant for different distillation tasks. Based on generalized divergences (cf. Ref.[48]), we define general families of information measures and prove local invariance under isometric transformation, independent of the specific form of the divergence. This simplifies the computation of various such quantities and allows for protocol optimization for operational tasks, whose performance is measured by an invariant quantity. In addition, the methods presented in this work are formulated as generally as possible, allowing them to be applied in various fields of quantum information science.

Nonetheless, the computation of some operationally relevant information measures can be very challenging. The so-called smooth max-mutual information can be used to determine a lower bound on the one-shot distillable key of a quantum state [85]. However, its

definition involves a numerically costly optimization, making it difficult to calculate. Semidefinite programming (SDP) has been proven to be an effective tool to calculate similar quantities. For the smooth max-mutual information, however, this method has not yet been successfully applied. In “Computation of the Smooth Max-Mutual Information via Semidefinite Programming” (cf. Sec.2.4.2 [80]) an algorithm is presented that allows the computation of the smooth max-mutual information for any quantum states if a weak restriction corresponding to the rank of its marginal states is fulfilled. It involves a particular SDP, which we fully characterize and prove strong duality, ensuring the numerical stability of the algorithm.

Both publications contribute to extending the methods of optimizing stabilizer-based operations for other or more general distillation tasks. The corresponding methods are summarized in Sec.3.2.5.

Based on these results, at the date of writing this dissertation, a publication regarding the extension of stabilizer-based methods to the task of secret-key distillation and more general resource-state distillation is in preparation. Preliminary results include the formalization of the stabilizer distillation protocol as a quantum channel. This makes the stabilizer distillation methods applicable to any bipartite quantum state. It can be shown how output states for different choices of stabilizer encodings can be related by a local unitary transformation. Moreover, the bipartite channel is extended to the tripartite setting of secret-key distillation (cf. Ref.[48]). In this setting, the output states relevant for drawing a secret key, potentially influenced by the third malicious party, are again related by local isometries. Together with the results of Sec.2.4.1 [79] and Sec.2.4.2 [80] this enables a more efficient evaluation of relevant information measures in order to optimize the generalized stabilizer distillation routine for the distillation of resource states with strong entanglement or a high capability for secret-key drawing.

3.1.6. Open questions

We conclude this section by collecting open questions with a close relation to the presented publications.

1. **How does the entanglement structure of Bell-diagonal systems of higher dimensions d compare to $d \leq 4$?**

Notwithstanding the quickly decreasing relative amount of Bell-diagonal PPT states for local dimensions $d \geq 5$, an analysis of Bell-diagonal systems as in Sec.2.1.3 [74] of higher dimension may reveal novel features of bound entangled states.

2. **What is the information-theoretic relation between bound entanglement of Bell-diagonal systems and the group structure of the Bell basis?**

Comparing the structure of standard Bell-diagonal bound entangled states to generalized Bell-diagonal bound entangled states as discussed in Sec.2.2.1 [75] may lead to a better understanding of the relation of group structure of Bell states and bound entanglement. Moreover, an answer to the question of why there exists more bound entanglement for systems of mixed group related Bell states may be found.

3. **How can FIMAX be generalized for entanglement distillation of multiple copies of general bipartite states in non-prime dimension?**

Regarding stabilizer-based entanglement distillation (cf. Sec.2.1.2 [73]), analyzing the possibility of extending the two-copy FIMAX protocol to non-prime dimensions and multi-copy regimes are natural next steps. Furthermore, investigating its action on general bipartite qudits, without prior twirl to the set of Bell-diagonal states, may further improve its distillation performance.

4. **Is the class of stabilizer-based distillation protocols strictly larger than the class of permutation-based distillation protocols?**

Focusing on the connecting properties of stabilizer-based and permutation-based protocols (see discussion in Sec.2.1.2 [73] and Ref.[86]), answering whether for each permutation-based protocol there exists an equivalent stabilizer-based protocol contributes to finding optimal entanglement distillation protocols.

5. **Which properties characterize NPT states that cannot be distilled by stabilizer-based methods?**

Given the remarkable ability of FIMAX to distill highly noisy states (cf. Sec.2.3.2 [78]), determining these properties may contribute to the distillability problem (Def.1.1.8) by identifying candidates for NPT bound entangled states.

6. **Can invariance under other local channels for divergence-based information quantities be proven by similar methods as those successful for local isometric channels?**

This open question corresponds to the method of combining a reversal channel with the data processing inequality of generalized divergences to prove invariance under local isometric transformations (cf. Sec.2.4.1 [79]). Noting that mainly properties regarding the local reversal of the isometric channel are used, this method may be extended to other channels and reversals, e.g., the Petz recovery map (cf. Ref.[48]), which is relevant for the general question of recoverability of quantum information [87].

7. **How can the algorithm for computing the smooth max-mutual information be generalized and applied to other information and entropic quantities?**

Given the collection of several information and entropic quantities in families as presented in the publication of Sec.2.4.1 [79], the SDP-based algorithm of Sec.2.4.2 [80] may be generalized for other quantities with similar properties to the smooth max-mutual information. Additionally, a natural next step would be to investigate the possibility of lifting the limiting rank requirement of the algorithm.

3.1.7. Review of bipartite bound entanglement

The review paper “Bipartite Bound Entanglement” (cf. Sec.2.5.1 [21]) provides a comprehensive review of existing results about bound entanglement in bipartite systems of finite

dimension. On the one hand, the review focuses on characterizing properties of PPT bound entangled states and how such states can be detected and used in applications. On the other hand, it discusses bound entanglement in the broader context of quantum information theory and highlights open questions.

Providing relevant information about existing results and methods to analyze bound entanglement, the review paper enables the scientific community to quickly identify open questions and applicable methods for novel investigations regarding bound entanglement.

3.1.8. Further work

The publication “The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks” (cf. Sec.2.6.1 [81]) contributes to the field of quantum machine learning. It introduces an architecture for Dissipative Quantum Neural Networks (DQNNs) based on building blocks that satisfy a notion of universality. Furthermore, the training process of such networks in dependence on the applied cost function is investigated.

3.2. Synopsis of methods

In this section, we discuss the methods used to generate the results as reported in the synopsis above. Before we review the techniques for specific results, we consider the general approach and the numerical environment.

3.2.1. Hybrid approach and software ecosystem

We use a hybrid approach of numerical and analytical methods with varying relevance for the different publications. This approach is well suited to the task of characterizing the entanglement structure of Bell-diagonal states. Analytical properties of the system enable or improve certain numerical methods, which, in turn, provide crucial observations and reveal new analytical structures regarding the entanglement structure and distillability of Bell-diagonal states.

The entanglement-conserving symmetries are a crucial example of this synergy. They allow for a significant increase in entanglement classification and detection capability of implemented methods. Conversely, analysis of the resulting classification data reveals certain structures to be investigated analytically. Of particular importance for the project, the observation regarding the relation of group structure of Bell states and entanglement structure of Bell-diagonal states shifts the focus to stabilizer-based methods for entanglement distillation. Another example is the optimization of extreme points of the set of separable Bell-diagonal states and the construction of Bell-diagonal entanglement witnesses. For that, a reduction of computational complexity can be achieved with the analytical results regarding the projection to Bell-diagonal systems by the Weyl twirl, as discussed in the publication of Sec.2.2.1 [75].

The numerical methods for entanglement classification and distillation have been consistently implemented in the dynamic and general-purpose programming language “Julia”

[88]. Its extensive ecosystem for software development, data science, and scientific computing offers a well-suited platform for numerical investigations in the scope of this project. Additionally, in comparison with other programming languages like C or Python, Julia often demonstrates the best performance [89]. Collecting all newly developed methods in an official Julia package (cf. Sec.2.1.1 [72]) certifies reproducibility of all numerical analyses. Furthermore, publishing the software as an open-source package enables the scientific community to contribute to the package. The package furthermore enables verification and benchmarking of analytical results (see, e.g., the publication of Sec.2.3.2 [78] regarding the performance of FIMAX).

3.2.2. Methods for entanglement classification and validation of Bell-diagonal systems

Analytical:

The combination of the group properties of Bell states constructed via the Weyl-Heisenberg operators (cf. Def.1.1.5) with existing entanglement classification methods is a main concept behind the results reported in Sec.3.1.2. One method related to the entanglement classification of Bell-diagonal systems (cf. Sec.2.1.2 [73] and Sec.2.1.3 [74]) is the combination of the projection to the set of Bell-diagonal states (cf. Sec.2.2.1 [75]) with an efficient parametrization of unitaries and quantum states [90, 91]. This significantly reduces the number of parameters and therefore improves the performance and stability of numerical classification methods. Utilizing the group properties of Bell states inherited from the Weyl-Heisenberg group (1.3) also improves the performance of analytical entanglement criteria for bipartite qutrits and provides a new perspective on these criteria (cf. Sec.2.2.2 [76]).

Numerical:

Adapting and optimizing existing methods for entanglement classification of quantum states for Bell-diagonal states enables analyses related to the entanglement structure in Bell-diagonal systems of local dimension three and four. The generated data sets of classified Bell-diagonal qutrit and ququart states provide a statistically solid and reproducible data foundation for the reported results of Sec.2.1.2 [73], Sec.2.1.3 [74] and Sec.2.2.1 [75].

The main methods for creating this data representing entanglement-classified Bell-diagonal states are as follows: First, Bell-diagonal states are represented by their mixing probabilities $c_{k,l}$ ($c_{k,l} \geq 0$, $\sum_{k,l} c_{k,l} = 1$), forming a $(d^2 - 1)$ -simplex in Euclidean space (cf. Def.1.1.9). A uniform distribution of points in this simplex then corresponds to a uniform distribution of Bell-diagonal quantum states induced by the Lebesgue measure in the simplex (see Sec.5.2 in [21] for more details) and is used to sample Bell-diagonal states. Second, analytical entanglement and separability criteria are implemented (cf. Sec.2.1.2 [73]). Third, leveraging an efficient parameterization [90] of separable states, random extreme points of the set of separable Bell-diagonal states $\text{SEP} \cap \mathcal{M}_d$ are determined. On the one hand, these are used for a polyhedral approximation of the convex set $\text{SEP} \cap \mathcal{M}_d$ allowing to classify separable states. On the other hand, the extreme

points can be used to define Bell-diagonal entanglement witnesses [19] (cf. Def.A.2.15), capable of detecting PPT- and NPT-entangled states. Finding these extreme points requires a multivariate nonlinear optimization over the set of parameterized separable Bell-diagonal states. There is no general exact solution to this optimization problem. Therefore, a heuristic combination of gradient-free and gradient-based methods providing the most consistent results is used. The gradient-free method "Nelder-Mead" [92] and the gradient-based method "Newton-Trust-Region" [93] are applied in their implementation of the package "Optim.jl" [94]. This is done for many iterations for different random starting points. The number of iterations for this multistart optimization is chosen such that the total minimum/maximum is the global minimum/maximum of the function with high probability. Additional cross-checks with all other applied entanglement criteria or witnesses provide further confidence. Notwithstanding, the solution may not be the true global optimum. For a separability classifier based on the polyhedral approximation of the set of separable states, this would imply that some separable states are not classified. For the entanglement witnesses, some separable states might get classified as entangled. To reduce the probability for this false-positive classification, we only classify states as entangled if the violation of a corresponding criterion exceeds an empirically selected value.

Finally, entanglement-conserving symmetries of Bell-diagonal systems [30, 59] implied by the group structure of Bell states are leveraged. Bell-diagonal states that are related by such a symmetry transformation are necessarily of the same entanglement class. Hence, if (bound) entanglement or separability is detected for any state contained in the set of states related by these symmetries, it is implied for the whole set. This significantly improves the inner polyhedral approximation of the set of separable Bell-diagonal states and the detection capability of entanglement criteria and witnesses.

Similar methods can be applied for higher dimensions as well. In this case, the computational complexity generally increases, mainly due to the optimization over $\text{SEP} \cap \mathcal{M}_d$. However, also the number of entanglement-conserving symmetries rises with the dimension, partially mitigating the decreased effectivity of entanglement criteria and more complex optimization. With the availability of training data of sufficient quantity and quality regarding Bell-diagonal qutrits and ququarts, methods from machine learning may reveal additional relevant structures for the characterization of bound entangled states in future investigations. Initial analyses show that convolutional neural network (CNN) classifiers are capable of identifying features related to the group structure of Bell states for classifying a state as separable, bound entangled, or free entangled with high success probability ¹.

¹A related investigation has been presented by the author during the 2nd International Workshop on Machine Learning and Quantum Computing Applications in Medicine and Physics, 04-07 June 2024 in Warsaw, Poland. A related publication is in progress.

3.2.3. Methods for identifying group structure as a relevant factor for the entanglement structure of Bell-diagonal systems

Analytical:

The common ground of analyses providing the insights described in Sec.3.1.3 is the relation of two-copy Weyl errors in $\mathcal{E}_2$ (cf. Def.1.2.1) and Bell states. The Weyl twirl operators $T_{i,j} := W_{i,j} \otimes W_{i,j}^*$ form a commutative subgroup of $\mathcal{E}_2$. They take diagonal form in the standard Bell basis, implying the equivalence of Bell-diagonal projection and Weyl twirl (cf. Sec.2.2.1 [75]). Moreover, standard Bell-diagonal states can be represented via these Weyl twirl operators. In this representation, the effect of applying Weyl errors or the PPT and realignment criteria can be easily determined. Leveraging this, we evaluate the PPT and realignment entanglement criteria for Bell-diagonal qutrits in terms of quantities that are related to the group properties of the Weyl operators in Ref.[76].

We introduce generalized Bell bases (cf. Sec.2.2.1 [75]) as bases of maximally entangled states that may not have a similar group structure as the standard Bell basis. They enable the systematic analysis of the influence of a group structure on the entanglement structure of Bell-diagonal systems. This is achieved by the possibility to effectively sample generalized Bell bases and therefore various Bell-diagonal systems for numerical comparison. One benefit of this family of bases is that its elements can be interpreted as disturbed Weyl-Heisenberg bases. The disturbance can be easily parameterized, allowing us to consider undisturbed and heavily disturbed bases in a continuous parameterization while ensuring that the basis states are maximally entangled and orthonormal.

Numerical:

Except for the symmetry-based techniques, which rely on the group properties of the Weyl-Heisenberg operators, we use the methods for the classification of standard Bell-diagonal states to classify numerous generalized Bell-diagonal systems (cf. Sec.3.2.2). We evaluate these data sets statistically via standard methods such as computing the mean, the median, the variance, the standard deviation, and correlation coefficients of the relative number of states for each entanglement class. Thereby, we can estimate the relative volume of entanglement classes in general Bell-diagonal systems and are able to compare these on equal footing to the standard Bell-diagonal system as described in Sec.3.1.3.

Two main numerical methods contribute to qualitatively identifying the group structure as a relevant factor for entanglement properties of Bell-diagonal states. The first method is to consider projections of the classified data to a subspace spanned by d Bell operators. For the standard Bell-diagonal system, these projections can reveal significant information about the entanglement class of Bell-diagonal states (cf. Ref.[74] [74]). For example, if a Bell-diagonal state has most of its support on the subspace spanned by d Bell operators that are related to the same coset of a subgroup $S \subset \mathcal{W}_d$, that state is very likely separable. This property is lost for generalized Bell bases. The second method is to analyze and statistically compare the classification data for many Bell-diagonal systems. We find that the volume of PPT states is typically significantly smaller for generalized Bell-diagonal systems than for the standard Bell-diagonal system. Likewise, the applied

entanglement criteria typically detect significantly less PPT-bound entanglement for mixed generalized Bell states than for mixed standard Bell states. Furthermore, the relative amount of PPT-bound entanglement among the PPT states correlates strongly with the total amount of PPT states. This indicates that for systems with a large share of PPT states, as is the case for systems close to the standard system exhibiting the group structure, the systems also contain an above-proportional amount of PPT-bound entanglement.

3.2.4. Methods for stabilizer-based entanglement distillation

Analytical:

The combination of group theoretical results and established methods of quantum information theory facilitates the development of the generalized stabilizer-based distillation theory and the entanglement distillation protocol FIMAX (cf. Sec.3.1.4). Stabilizer codes (Def.A.2.17) are inherently related to the Weyl-Heisenberg group and allow the construction of likewise related operations. We utilize the shared properties of operations and Bell states and derive the analytical action of such stabilizer operations on Bell-diagonal states in terms of group-theoretic objects. For prime dimensions d and two copies of input states $N = 2$, the group of Weyl errors $\mathcal{E}_2$ (Def.1.2.1) and the spectra of the Weyl operators (1.1) have special properties that are used in Sec.2.3.1 [77] to define a specific “canonical” stabilizer encoding. The developed theory describing the action of stabilizer-based operations on Bell-diagonal states and the properties of the canonical encoding together imply the reported maximization of the output fidelity of FIMAX. In the form presented in Sec.2.3.1 [77], the FIMAX protocol is only directly applicable to Bell-diagonal states. A general input state needs to be projected to Bell-diagonal form first, e.g., by Weyl-twirling (cf. Sec.2.2.1 [75]), potentially introducing new noise and impeding distillation.

Numerical:

The performance comparisons of FIMAX and other distillation protocols as reported in Sec.2.3.1 [77] and Sec.2.3.2 [78] are carried out by applying the protocols to the same sample of states and evaluating performance indicators such as the distillation efficiency and distillable share. Comparing the performance as well for different state families contributes to a wide applicability of reported results. Using the methods implemented in the BellDiagonalQudits package, we sample uniformly distributed standard and generalized Bell-diagonal states of certain families according to their distribution in the simplex (cf. Sec.3.2.2). Pure states are sampled according to the Haar measure. We then post-select generated states by their fidelity, creating uniformly distributed sample sets for analyses regarding the distillation performance. The reported performance indicators represent average values for these sample sets. The sample sizes are chosen so that the standard deviation of the mean value is as specified in the corresponding publications, assuming a uniform distribution. Additionally, the reported analytical results are generally verified by comparing numerical outputs to analytically expected results for random inputs.

3.2.5. Methods for the generalization and extension of stabilizer-based distillation

Analytical:

The publications of Sec.2.4.1 [79] and Sec.2.4.2 [80] extend the focus to methods that are not based on constructions related to the Weyl-Heisenberg operators and Bell states. They provide a way to extend the methods described above to more general systems and distillation tasks. Of central importance are generalized divergence-based information quantities (cf. Ref.[48] for a comprehensive overview) that characterize the suitability of a quantum state for certain operational tasks. The defining property of generalized divergences is that they obey a data-processing inequality (cf. Ref.[48]), implying that the application of a quantum channel to two quantum states cannot render them more distinguishable as measured by the divergence. Combining the data-processing inequality with suitable reversal channels (cf. Ref.[48]) is the main method of proving invariance under local isometric channels for general families of divergence-based information and entropic measures (cf. Sec.2.4.1 [79]). A major benefit of this technique is its generality, depending only on the definition of the information measure via the divergence, but not on the specific divergence itself. Thereby, local isometric invariance is implied for several information measures based on different divergences and is therefore applicable in various operational settings.

The algorithm presented in the publication of Sec.2.4.2 [80] to compute the smooth max-mutual information relies on the iterative solution of two semidefinite programs. One of them has not been characterized before, and this is done by determining its primal and dual formulation and proving strong duality, ensuring equal solution of the primal and dual problem and numerical stability. The presented algorithm is shown to converge to the smooth max-mutual information if the state satisfies a rank condition for its marginals in the smoothing environment.

Numerical:

Again, the Julia ecosystem provides effective methods such as the package “Quantuminformation.jl” [95] to verify the analytical findings. To implement and verify the results regarding the semidefinite programs presented in the publication of Sec.2.4.2 [80], the “splitting conic solver (SCS)” [96] together with the optimization framework and Julia package “JuMP” [97] is used.

3.3. Concluding discussion

The overarching objective of this thesis is to address an operationally meaningful question: How can we characterize the entanglement structure of bipartite Bell-diagonal systems, and how are their properties related to distillability and bound entanglement? The work combines group-related analysis, reproducible numerics, and constructive protocol

design. It shows that Belldiagonal systems reveal a structured entanglement landscape, including manifestations of bound entanglement. Moreover, Bell-diagonal systems admit stabilizerbased methods that exploit that structure to distill entanglement effectively. In the following, we synthesize the main results, place them in a broader context, and outline promising directions for future research.

A major practical outcome of the thesis is a reproducible numerical ecosystem. It generates large, statistically grounded data sets and provides numerical methods for entanglement analyses that support the analytic observations. The “BellDiagonalQudits.jl” package and its sampling and classification methods, applied to systems of local dimensions $d \leq 5$ in the scope of reported results, provide a transparent foundation for the distribution and prevalence of PPT-bound entanglement we observe in the standard Belldiagonal system. The data also reveals the limits of current criteria as the dimension grows beyond $d = 4$, motivating both algorithmic and theoretical extensions.

These numerical results demonstrate that a significant fraction of PPT Belldiagonal qutrits and ququarts are bound entangled. In addition, much of the discriminative information between separable, PPTentangled, and NPT states is reflected in the group structure of the Weyl-Heisenberg operators. We identify the algebraic relations of Bell states as a core characterizing property for the entanglement structure of corresponding Bell-diagonal states: standard Bell-diagonal states, governed by group-induced symmetries, exhibit enhanced bound entanglement compared to generalized variants, where group structures are disrupted. This is analytically underpinned by reformulations of PPT and realignment criteria in terms of subgroups and their cosets when applied to standard Bell-diagonal states.

Interpreting these structural findings operationally, we demonstrate that stabilizerbased constructions aligned with the Weyl-Heisenberg group yield powerful distillation tools. By deriving and utilizing a general form for the action of LOCC stabilizer operations on multicopy Belldiagonal states, we introduce the FIMAX protocol. Acting on two copies of prime dimensions, we prove the eponymous property of the protocol: it maximizes the fidelity increase per iteration among stabilizerbased constructions. Empirically, FIMAX outperforms competing recurrent protocols in distillation efficiency and robustness to noise. Furthermore, it remains effective even for low-fidelity and nonBelldiagonal inputs that require prior twirling to Bell-diagonal form. These results illustrate a broader principle: when the error model, target states, and applicable operations share an algebraic backbone, one can design resourceefficient protocols that exploit this common structure to achieve high performance.

The results presented in this thesis also extend the stabilizer perspective beyond entanglement distillation by developing tools to optimize other operational quantities besides fidelity. Proving local invariance of a broad class of divergencebased information measures under isometric transformations simplifies the evaluation of taskrelevant quantities. Additionally, it supports the generalization of stabilizer methods to secretkey distillation and related resource tasks. Complementing this, we develop the SDPbased algorithm for computing the smooth maxmutual information. This algorithm provides a practical route to optimizing protocols for cryptographic objectives. Together these theoretical

and computational advances lay the groundwork for treating stabilizer distillation as a flexible channel design problem rather than a single fidelity-centric routine.

At the same time, the thesis acknowledges important limitations and open problems. The relative abundance of PPT states and the detectability of bound entanglement decline with increasing local dimension, so conclusions drawn for $d = 3$ and $d = 4$ must be confirmed for higher dimensions. Many of the strongest structural and operational statements depend on the Weyl-Heisenberg group underlying the standard Bell basis. Generalized Bell bases that break this group relation exhibit markedly different entanglement landscapes and reduced effectiveness of group-based criteria and stabilizer methods. Crucially, the precise information-theoretic relation between the underlying group structure and bound entanglement, as well as distillability, remains to be established. Also, the entanglement and distillability landscape is not yet fully charted: whether there exist NPT-bound entangled states or whether all Bell-diagonal NPT states are distillable remain open questions.

The broader implications of the work reach into resource theory, cryptography, and experimental quantum information. By connecting structural features of Bell-diagonal systems to operational resources such as entanglement and distillable key, the thesis bridges foundational questions about bound entanglement, including the separability and the distillability problem, with practical communication and cryptography tasks. Stabilizer-based distillation protocols that align with natural laboratory noise models are promising candidates for near-term experimental validation in qubit and qutrit settings. Methodologically, the combination of open-source software and symmetry-aware analysis for methods that are applied in different fields of (quantum) information processing offers a transferable template. It can be applied to other information processing problems where symmetry can reduce complexity and guide protocol design.

Looking forward, several concrete research directions emerge. Extending the numerical and analytic studies of Bell-diagonal systems to higher local dimensions will clarify whether new structural phenomena regarding bound entanglement appear beyond $d = 4$. Generalizing FIMAX and the stabilizer framework to nonprime dimensions and multicopy regimes could broaden their operational reach. The same holds for extending them to the distillation of non-Bell-diagonal inputs without prior twirling. A systematic analysis of general LOCC stabilizer operations and their action on group-related Bell-diagonal states can reveal which subgroup structures relate to distillability via stabilizer operations. This analysis can be combined with group-theoretic formulations of entanglement criteria. Together, they may yield a general characterization of how group structure relates to distillability in Bell-diagonal systems. Moreover, finding NPT states that resist stabilizer-based distillation identifies candidates for NPT bound entanglement. Finally, experimental tests on small quantum processors could validate and refine the protocols under realistic noise.

These contributions collectively elevate the study of Bell-diagonal systems from descriptive analyses to actionable frameworks. This benefits quantum information science in several ways. It provides reproducible tools, analytical characterizations that advance our understanding of bound entanglement, and protocols for quantum information processing

in noisy environments. The emphasis on group structure as a unifying theme bridges entanglement classification and distillation, offering a blueprint for tackling similar phenomena in other systems. The open questions and technical challenges identified here form a clear and actionable agenda for future work. They chart a path toward deepening the theory, expanding the computational toolkit, and designing protocols and experiments that advance quantum communication and cryptography.

Literatur

- ¹M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
- ²D. Cozzolino, B. Da Lio, D. Bacco und L. K. Oxenløwe, “HighDimensional Quantum Communication: Benefits, Progress, and Future Challenges”, en, *Advanced Quantum Technologies* **2**, 1900038 (2019).
- ³Y. Wang, Z. Hu, B. C. Sanders und S. Kais, “Qudits and High-Dimensional Quantum Computing”, en, *Frontiers in Physics* **8**, 589504 (2020).
- ⁴H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf und M. Van den Nest, “Measurement-based quantum computation”, en, *Nature Physics* **5**, Publisher: Nature Publishing Group, 19–26 (2009).
- ⁵S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi und P. Wallden, “Advances in quantum cryptography”, EN, *Advances in Optics and Photonics* **12**, Publisher: Optica Publishing Group, 1012–1236 (2020).
- ⁶C. H. Bennett und G. Brassard, “Quantum cryptography: Public key distribution and coin tossing”, *Theoretical Computer Science, Theoretical Aspects of Quantum Cryptography celebrating 30 years of BB84* **560**, 7–11 (2014).
- ⁷A. K. Ekert, “Quantum cryptography based on Bell’s theorem”, *Physical Review Letters* **67**, Publisher: American Physical Society, 661–663 (1991).
- ⁸C. L. Degen, F. Reinhard und P. Cappellaro, “Quantum sensing”, *Reviews of Modern Physics* **89**, Publisher: American Physical Society, 035002 (2017).
- ⁹R. Horodecki, P. Horodecki, M. Horodecki und K. Horodecki, “Quantum entanglement”, en, *Reviews of Modern Physics* **81**, arXiv:quant-ph/0702225, 865–942 (2009).
- ¹⁰C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres und W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”, *Physical Review Letters* **70**, Publisher: American Physical Society, 1895–1899 (1993).
- ¹¹C. H. Bennett und S. J. Wiesner, “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states”, *Physical Review Letters* **69**, Publisher: American Physical Society, 2881–2884 (1992).
- ¹²R. F. Werner, “All teleportation and dense coding schemes”, en, *Journal of Physics A: Mathematical and General* **34**, 7081 (2001).

-
- ¹³S. Bäuml, S. Das, X. Wang und M. M. Wilde, “Resource theory of entanglement for bipartite quantum channels”, arXiv:1907.04181 [quant-ph], 10.48550/arXiv.1907.04181 (2019).
- ¹⁴E. Chitambar und G. Gour, “Quantum resource theories”, Reviews of Modern Physics **91**, Publisher: American Physical Society, 025001 (2019).
- ¹⁵A. Peres, “Separability Criterion for Density Matrices”, Physical Review Letters **77**, Publisher: American Physical Society, 1413–1415 (1996).
- ¹⁶K. Chen und L.-A. Wu, “A matrix realignment method for recognizing entanglement”, Quantum Information and Computation **3**, 10.26421/QIC3.3-1 (2002).
- ¹⁷O. Rudolph, “Further Results on the Cross Norm Criterion for Separability”, en, Quantum Information Processing **4**, 219–239 (2005).
- ¹⁸G. Sarbicki, G. Scala und D. Chruściński, “Family of multipartite separability criteria based on a correlation tensor”, en, Physical Review A **101**, 012341 (2020).
- ¹⁹B. M. Terhal, “Bell Inequalities and the Separability Criterion”, en, Physics Letters A **271**, arXiv:quant-ph/9911057, 319–326 (2000).
- ²⁰L. Gurvits, “Classical deterministic complexity of Edmonds’ Problem and quantum entanglement”, en, in Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (Juni 2003), S. 10–19.
- ²¹B. C. Hiesmayr, C. Popp und T. C. Sutter, “Bipartite bound entanglement”, International Journal of Quantum Information **23**, Publisher: World Scientific Publishing Co., 2530003 (2025).
- ²²M. Plenio und S. Virmani, “An introduction to entanglement measures”, Quantum Information and Computation **7**, 1–51 (2007).
- ²³C. H. Bennett, D. P. DiVincenzo, J. A. Smolin und W. K. Wootters, “Mixed State Entanglement and Quantum Error Correction”, en, Physical Review A **54**, arXiv:quant-ph/9604024, 3824–3851 (1996).
- ²⁴P. M. Hayden, M. Horodecki und B. M. Terhal, “The asymptotic entanglement cost of preparing a quantum state”, en, Journal of Physics A: Mathematical and General **34**, 6891 (2001).
- ²⁵K. Horodecki, M. Horodecki, P. Horodecki, D. Leung und J. Oppenheim, “Quantum Key Distribution Based on Private States: Unconditional Security Over Untrusted Channels With Zero Quantum Capacity”, IEEE Transactions on Information Theory **54**, Conference Name: IEEE Transactions on Information Theory, 2604–2620 (2008).
- ²⁶C. H. Bennett, H. J. Bernstein, S. Popescu und B. Schumacher, “Concentrating partial entanglement by local operations”, en, Physical Review A **53**, 2046–2052 (1996).
- ²⁷M. A. Nielsen und I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, en, ISBN: 9780511976667 Publisher: Cambridge University Press, Dez. 2010.

-
- ²⁸M. ukowski, A. Zeilinger, M. A. Horne und A. K. Ekert, ““Event-ready-detectors” Bell experiment via entanglement swapping”, *Physical Review Letters* **71**, Publisher: American Physical Society, 4287–4290 (1993).
- ²⁹I. Bengtsson und K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
- ³⁰B. Baumgartner, B. Hiesmayr und H. Narnhofer, “A special simplex in the state space for entangled qudits”, en, *Journal of Physics A: Mathematical and Theoretical* **40**, arXiv:quant-ph/0610100, 7919–7938 (2007).
- ³¹J. Preskill, “Quantum Computing in the NISQ era and beyond”, en-GB, *Quantum* **2**, Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 79 (2018).
- ³²S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall und T. J. Yoder, “High-threshold and low-overhead fault-tolerant quantum memory”, en, *Nature* **627**, Publisher: Nature Publishing Group, 778–782 (2024).
- ³³R. Acharya, D. A. Abanin, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, N. Astrakhantsev, J. Atalaya, R. Babbush, D. Bacon, B. Ballard, J. C. Bardin, J. Bausch, A. Bengtsson, A. Bilmes, S. Blackwell, S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, D. A. Browne, B. Buchea, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, A. Cabrera, J. Campero, H.-S. Chang, Y. Chen, Z. Chen, B. Chiaro, D. Chik, C. Chou, J. Claes, A. Y. Cleland, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, S. Das, A. Davies, L. De Lorenzo, D. M. Debroy, S. Demura, M. Devoret, A. Di Paolo, P. Donohoe, I. Drozdov, A. Dunsworth, C. Earle, T. Edlich, A. Eickbusch, A. M. Elbag, M. Elzouka, C. Erickson, L. Faoro, E. Farhi, V. S. Ferreira, L. F. Burgos, E. Forati, A. G. Fowler, B. Foxen, S. Ganjam, G. Garcia, R. Gasca, É. Genois, W. Jiang, C. Gidney, D. Gilboa, R. Gosula, A. G. Dau, D. Graumann, A. Greene, J. A. Gross, S. Habegger, J. Hall, M. C. Hamilton, M. Hansen, M. P. Harrigan, S. D. Harrington, F. J. H. Heras, S. Heslin, P. Heu, O. Higgott, G. Hill, J. Hilton, G. Holland, S. Hong, H.-Y. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, S. Jordan, C. Joshi, P. Juhas, D. Kafri, H. Kang, A. H. Karamlou, K. Kechedzhi, J. Kelly, T. Khair, T. Khatkar, M. Khezri, S. Kim, P. V. Klimov, A. R. Klots, B. Kobrin, P. Kohli, A. N. Korotkov, F. Kostritsa, R. Kothari, B. Kozlovskii, J. M. Kreikebaum, V. D. Kurilovich, N. Lacroix, D. Landhuis, T. Lange-Dei, B. W. Langley, P. Laptev, K.-M. Lau, L. Le Guevel, J. Ledford, J. Lee, K. Lee, Y. D. Lensky, S. Leon, B. J. Lester, W. Y. Li, Y. Li, A. T. Lill, W. Liu, W. P. Livingston, A. Locharla, E. Lucero, D. Lundahl, A. Lunt, S. Madhuk, F. D. Malone, A. Maloney, S. Mandrà, J. Manyika, L. S. Martin, O. Martin, S. Martin, C. Maxfield, J. R. McClean, M. McEwen, S. Meeks, A. Megrant, X. Mi, K. C. Miao, A. Mieszala, R. Molavi, S. Molina, S. Montazeri, A. Morvan, R. Movassagh, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, H. Neven, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, C.-H. Ni, M. Y. Niu, T. E. O’Brien, W. D. Oliver, A. Opremcak, K. Ottosson, A. Petukhov, A. Pizzuto, J. Platt, R. Potter, O. Pritchard, L. P. Pryadko, C. Quintana, G. Ramachandran,

-
- M. J. Reagor, J. Redding, D. M. Rhodes, G. Roberts, E. Rosenberg, E. Rosenfeld, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, A. W. Senior, M. J. Shearn, A. Shorter, N. Shutty, V. Shvarts, S. Singh, V. Sivak, J. Skrzynny, S. Small, V. Smelyanskiy, W. C. Smith, R. D. Somma, S. Springer, G. Sterling, D. Strain, J. Suchard, A. Szasz, A. Sztein, D. Thor, A. Torres, M. M. Torunbalci, A. Vaishnav, J. Vargas, S. Vdovichev, G. Vidal, B. Villalonga, C. V. Heidweiller, S. Waltman, S. X. Wang, B. Ware, K. Weber, T. Weidel, T. White, K. Wong, B. W. K. Woo, C. Xing, Z. J. Yao, P. Yeh, B. Ying, J. Yoo, N. Yosri, G. Young, A. Zalcman, Y. Zhang, N. Zhu, N. Zobrist und Google Quantum AI and Collaborators, “Quantum error correction below the surface code threshold”, en, *Nature*, Publisher: Nature Publishing Group, 1–7 (2024).
- ³⁴C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin und W. K. Wootters, “Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels”, en, *Physical Review Letters* **76**, 722–725 (1996).
- ³⁵A. M. Steane, “Error Correcting Codes in Quantum Theory”, *Physical Review Letters* **77**, Publisher: American Physical Society, 793–797 (1996).
- ³⁶D. Gottesman, “Stabilizer Codes and Quantum Error Correction”, en, *arXiv:quant-ph/9705052*, <https://doi.org/10.48550/arXiv.quant-ph/9705052> (1997).
- ³⁷D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu und A. Sanpera, “Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels”, *Physical Review Letters* **77**, Publisher: American Physical Society, 2818–2821 (1996).
- ³⁸G. Alber, A. Delgado, N. Gisin und I. Jex, “Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces”, en, *Journal of Physics A: Mathematical and General* **34**, 8821 (2001).
- ³⁹M. Horodecki und P. Horodecki, “Reduction criterion of separability and limits for a class of distillation protocols”, *Physical Review A* **59**, Publisher: American Physical Society, 4206–4216 (1999).
- ⁴⁰J. Miguel-Ramiro und W. Dür, “Efficient entanglement purification protocols for d-level systems”, en, *Physical Review A* **98**, 042309 (2018).
- ⁴¹R. Matsumoto, “Conversion of a general quantum stabilizer code to an entanglement distillation protocol”, en, *Journal of Physics A: Mathematical and General* **36**, *arXiv:quant-ph/0209091*, 8113–8127 (2003).
- ⁴²J. Dehaene, M. Van Den Nest, B. De Moor und F. Verstraete, “Local permutations of products of Bell states and entanglement distillation”, en, *Physical Review A* **67**, 022310 (2003).
- ⁴³H. Bombin und M. A. Martin-Delgado, “Entanglement distillation protocols and number theory”, en, *Physical Review A* **72**, 032313 (2005).
- ⁴⁴K. G. H. Vollbrecht und F. Verstraete, “Interpolation of recurrence and hashing entanglement distillation protocols”, *Physical Review A* **71**, Publisher: American Physical Society, 062325 (2005).

-
- ⁴⁵K. G. H. Vollbrecht und M. M. Wolf, “Efficient distillation beyond qubits”, en, *Physical Review A* **67**, 012303 (2003).
- ⁴⁶I. Devetak und A. Winter, “Distillation of secret key and entanglement from quantum states”, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **461**, arXiv:quant-ph/0306078, 207–235 (2005).
- ⁴⁷R. Matsumoto, *Breeding protocols are advantageous for finite-length entanglement distillation*, arXiv:2401.02265 [quant-ph], Feb. 2024.
- ⁴⁸S. Khatri und M. M. Wilde, *Principles of Quantum Communication Theory: A Modern Approach*, arXiv:2011.04672 [cond-mat, physics:hep-th, physics:math-ph, physics:quant-ph], Feb. 2024.
- ⁴⁹P. Horodecki, “Separability criterion and inseparable mixed states with positive partial transposition”, *Physics Letters A* **232**, 333–339 (1997).
- ⁵⁰P. Horodecki, J. I. Cirac und M. Lewenstein, “Bound Entanglement for Continuous Variables is a Rare Phenomenon”, en, in *Quantum Information with Continuous Variables*, hrsg. von S. L. Braunstein und A. K. Pati (Springer Netherlands, Dordrecht, 2003), S. 211–228.
- ⁵¹M. Horodecki, P. Horodecki und R. Horodecki, “Mixed-State Entanglement and Distillation: Is there a Bound Entanglement in Nature?”, en, *Physical Review Letters* **80**, 5239–5242 (1998).
- ⁵²P. Pankowski, M. Piani, M. Horodecki und P. Horodecki, “A few steps more towards NPT bound entanglement”, *IEEE Transactions on Information Theory* **56**, arXiv:0711.2613 [quant-ph], 4085–4100 (2010).
- ⁵³S.-Y. Qi, G. Gupur, Y.-C. Wu und G.-P. Guo, “Nonpositive-transpose entanglement and bound entanglement: From distillability sets to inequalities and multivariable insights”, *Physical Review A* **110**, Publisher: American Physical Society, 012406 (2024).
- ⁵⁴P. Horodecki, . Rudnicki und K. yczkowski, “Five Open Problems in Quantum Information Theory”, en, *PRX Quantum* **3**, 010101 (2022).
- ⁵⁵K. Horodecki, M. Horodecki, P. Horodecki und J. Oppenheim, “Secure Key from Bound Entanglement”, *Physical Review Letters* **94**, Publisher: American Physical Society, 160502 (2005).
- ⁵⁶Y. Qin, J. Cheng, J. Ma, D. Zhao, Z. Yan, X. Jia, C. Xie und K. Peng, “Efficient and secure quantum secret sharing for eight users”, *Physical Review Research* **6**, Publisher: American Physical Society, 033036 (2024).
- ⁵⁷C. Roch i Carceller und A. Tavakoli, “Bound Entangled States Are Useful in Prepare-and-Measure Scenarios”, *Physical Review Letters* **134**, Publisher: American Physical Society, 120203 (2025).
- ⁵⁸J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs und A. Zeilinger, “Experimental entanglement purification of arbitrary unknown states”, en, *Nature* **423**, Publisher: Nature Publishing Group, 417–422 (2003).

-
- ⁵⁹B. Baumgartner, B. C. Hiesmayr und H. Narnhofer, “State space for two qutrits has a phase space structure in its core”, en, *Physical Review A* **74**, 032327 (2006).
- ⁶⁰B. C. Hiesmayr, “Free versus bound entanglement, a NP-hard problem tackled by machine learning”, en, *Scientific Reports* **11**, Publisher: Nature Publishing Group, 19739 (2021).
- ⁶¹B. C. Hiesmayr und W. Löffler, “Complementarity reveals bound entanglement of two twisted photons”, en, *New Journal of Physics* **15**, 083036 (2013).
- ⁶²B. Baumgartner, B. C. Hiesmayr und H. Narnhofer, “The geometry of bipartite qutrits including bound entanglement”, en, *Physics Letters A* **372**, arXiv:0705.1403 [quant-ph], 2190–2195 (2008).
- ⁶³M. Tomamichel, *Quantum Information Processing with Finite Resources – Mathematical Foundations*, Bd. 5, arXiv:1504.00233 [quant-ph] (2016).
- ⁶⁴L. Wang und R. Renner, “One-Shot Classical-Quantum Capacity and Hypothesis Testing”, *Physical Review Letters* **108**, Publisher: American Physical Society, 200501 (2012).
- ⁶⁵F. Buscemi und N. Datta, “The Quantum Capacity of Channels With Arbitrarily Correlated Noise”, *IEEE Transactions on Information Theory* **56**, 1447–1460 (2010).
- ⁶⁶E. Knill, *Non-binary unitary error bases and quantum codes*, en, *Techn. Ber. LA-UR–96-2717*, 373768 (Juni 1996), LA-UR–96–2717, 373768.
- ⁶⁷C. Popp, T. C. Sutter und B. C. Hiesmayr, “A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits”, arXiv:2408.02383 [quant-ph], 10.48550/arXiv.2408.02383 (2024).
- ⁶⁸E. Rains, “Nonbinary quantum codes”, *IEEE Transactions on Information Theory* **45**, Conference Name: IEEE Transactions on Information Theory, 1827–1832 (1999).
- ⁶⁹A. Ashikhmin und E. Knill, “Nonbinary quantum stabilizer codes”, *IEEE Transactions on Information Theory* **47**, 3065–3072 (2001).
- ⁷⁰L. Chen und D. . okovi, “Distillability and PPT entanglement of low-rank quantum states”, en, *Journal of Physics A: Mathematical and Theoretical* **44**, 285303 (2011).
- ⁷¹P. Horodecki, J. A. Smolin, B. M. Terhal und A. V. Thapliyal, “Rank two bipartite bound entangled states do not exist”, *Theoretical Computer Science, Algorithms in Quantum Information Processing* **292**, 589–596 (2003).
- ⁷²C. Popp, “BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits”, en, *Journal of Open Source Software* **8**, 4924 (2023).
- ⁷³C. Popp und B. C. Hiesmayr, “Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits”, en, *Scientific Reports* **12**, Number: 1 Publisher: Nature Publishing Group, 12472 (2022).
- ⁷⁴C. Popp und B. C. Hiesmayr, “Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts”, en, *Scientific Reports* **13**, Number: 1 Publisher: Nature Publishing Group, 2037 (2023).

-
- ⁷⁵C. Popp und B. C. Hiesmayr, “Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states”, en, *New Journal of Physics* **26**, Publisher: IOP Publishing, 013039 (2024).
- ⁷⁶T. C. Sutter, C. Popp und B. C. Hiesmayr, *A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria*, arXiv:2508.18393 [quant-ph], Aug. 2025.
- ⁷⁷C. Popp, T. C. Sutter und B. C. Hiesmayr, “A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits”, en-GB, *Quantum* **9**, Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 1945 (2025).
- ⁷⁸C. Popp, T. C. Sutter und B. C. Hiesmayr, “Low-fidelity entanglement distillation with FIMAX”, *International Journal of Quantum Information* **23**, Publisher: World Scientific Publishing Co., 2550017 (2025).
- ⁷⁹C. Popp, T. C. Sutter und B. C. Hiesmayr, “Local Invariance of Divergence-Based Quantum Information Measures”, en, *Entropy* **27**, Publisher: Multidisciplinary Digital Publishing Institute, 1051 (2025).
- ⁸⁰C. Popp, T. C. Sutter und B. C. Hiesmayr, *Computation of the Smooth Max-Mutual Information via Semidefinite Programming*, arXiv:2509.07743 [quant-ph], Okt. 2025.
- ⁸¹T. C. Sutter, C. Popp und B. C. Hiesmayr, *The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks*, arXiv:2502.09526 [quant-ph], Feb. 2025.
- ⁸²C. H. Bennett und G. Brassard, “Quantum cryptography: Public key distribution and coin tossing”, *Theoretical Computer Science, Theoretical Aspects of Quantum Cryptography celebrating 30 years of BB84* **560**, 7–11 (2014).
- ⁸³P. Horodecki und R. Augusiak, *On quantum cryptography with bipartite bound entangled states*, en, arXiv:0712.3999 [quant-ph], Dez. 2007.
- ⁸⁴I. Devetak, “The private classical capacity and quantum capacity of a quantum channel”, *IEEE Transactions on Information Theory* **51**, 44–55 (2005).
- ⁸⁵S. Khatri, E. Kaur, S. Guha und M. M. Wilde, *Second-order coding rates for key distillation in quantum key distribution*, arXiv:1910.03883 [quant-ph], Juli 2021.
- ⁸⁶E. Hostens, J. Dehaene und B. De Moor, *The equivalence of two approaches to the design of entanglement distillation protocols*, arXiv:quant-ph/0406017, Juni 2004.
- ⁸⁷M. Junge, R. Renner, D. Sutter, M. M. Wilde und A. Winter, “Universal recoverability in quantum information”, in *2016 IEEE International Symposium on Information Theory (ISIT)*, ISSN: 2157-8117 (Juli 2016), S. 2494–2498.
- ⁸⁸J. Bezanson, A. Edelman, S. Karpinski und V. B. Shah, “Julia: A Fresh Approach to Numerical Computing”, *SIAM Review* **59**, Publisher: Society for Industrial and Applied Mathematics, 65–98 (2017).

-
- ⁸⁹S. V. Krishna, O. Jadhav, P. Ardhapurkar, M. Valmiki, S. Agrawal, R. Bulusu, P. Dinde und S. Wandhekar, “Performance Comparison of Julia with C and Python for Solving Computational Problems”, en, in *Advances in Computing and Data Sciences*, hrsg. von M. Singh, V. Tyagi, P. K. Gupta, J. Flusser, T. Ören, A. R. Cherif und R. Tomar (2025), S. 35–45.
- ⁹⁰C. Spengler, M. Huber und B. C. Hiesmayr, “A composite parameterization of unitary groups, density matrices and subspaces”, en, *Journal of Physics A: Mathematical and Theoretical* **43**, arXiv:1004.5252 [math-ph, physics:quant-ph], 385306 (2010).
- ⁹¹C. Spengler, M. Huber und B. C. Hiesmayr, “Composite parameterization and Haar measure for all unitary and special unitary groups”, en, *Journal of Mathematical Physics* **53**, arXiv:1103.3408 [math-ph, physics:quant-ph], 013501 (2012).
- ⁹²J. A. Nelder und R. Mead, “A Simplex Method for Function Minimization”, *The Computer Journal* **7**, 308–313 (1965).
- ⁹³“Numerical Optimization”, en, in , hrsg. von J. Nocedal und S. J. Wright, *Springer Series in Operations Research and Financial Engineering* (1999).
- ⁹⁴P. K. Mogensen und A. N. Riseth, “Optim: A mathematical optimization package for Julia”, en, *Journal of Open Source Software* **3**, 615 (2018).
- ⁹⁵P. Gawron, D. Kurzyk und . Pawela, “QuantumInformation.jlA Julia package for numerical computation in quantum information theory”, en, *PLOS ONE* **13**, Publisher: Public Library of Science, e0209358 (2018).
- ⁹⁶B. O’Donoghue, E. Chu, N. Parikh und S. Boyd, *Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding*, arXiv:1312.3039 [math], Juli 2016.
- ⁹⁷M. Lubin, O. Dowson, J. D. Garcia, J. Huchette, B. Legat und J. P. Vielma, “JuMP 1.0: recent improvements to a modeling language for mathematical optimization”, en, *Mathematical Programming Computation* **15**, 581–589 (2023).
- ⁹⁸J. Watrous, *The Theory of Quantum Information*, en, 1. Aufl. (Cambridge University Press, Apr. 2018).

A. Preliminaries

We briefly summarize the essential theoretical preliminaries of quantum information theory within the scope of this thesis. For definitions, we mainly adapt the notation introduced in Ref.[21] (cf. Sec.2.5.1) and refer to standard literature as, e.g., Ref.[1] and Ref.[27].

A.1. Notation

Let $\mathcal{H}$ be a Hilbert space of finite dimension $d > 1$ and $\mathcal{L}(\mathcal{H})$ be the set of linear operators acting on $\mathcal{H}$. Given a partition of Hilbert spaces and corresponding objects acting on subsystems are labeled by subscripts accordingly. We mainly consider the two parties Alice and Bob, holding systems labeled A and B , respectively. $\text{Tr}[\cdot]$ denotes the trace operator, $\mathbb{Z}_d := \mathbb{Z}/d\mathbb{Z}$ the ring of integers modulo d and $\omega := e^{2\pi i/d}$.

A.2. Definitions

Definition A.2.1 (Quantum State) [21]

A linear operator $\rho \in \mathcal{L}(\mathcal{H})$ acting on a Hilbert space $\mathcal{H}$ is called quantum state or density operator if it is positive semi-definite ($\rho \geq 0$) and has unit trace ($\text{Tr}(\rho) = 1$). The set of all quantum states on $\mathcal{H}$ is denoted $\mathcal{D}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$.

The set of quantum states is convex [27], i.e., for any alphabet $\mathcal{X}$, probabilities p_x , and quantum states $\rho_x \in \mathcal{D}(\mathcal{H})$, $x \in \mathcal{X}$, one has

$$\sum_{x \in \mathcal{X}} p_x \rho_x \in \mathcal{D}(\mathcal{H}). \quad (\text{A.1})$$

The extremal points $\psi \in \mathcal{D}(\mathcal{H})$ of this convex set are called pure states [1], being rank-one projections onto a normalized vector $|\psi\rangle \in \mathcal{H}$. They can be written as $\psi = |\psi\rangle\langle\psi|$, where $\langle\psi|$ is the Hermitian conjugate of $|\psi\rangle$. A quantum state ρ is pure if and only if $\text{Tr}[\rho^2] = 1$. Any quantum state can be represented in its spectral decomposition

$$\rho = \sum_{i=0}^{r-1} \lambda_i |\psi_i\rangle\langle\psi_i| \quad (\text{A.2})$$

with eigenvectors $|\psi_i\rangle \in \mathcal{H}$ of $\rho \in \mathcal{D}(\mathcal{H})$ corresponding to the eigenvalue λ_i and r being the number of nonzero eigenvalues, i.e., the rank of ρ . Quantum states acting on the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of two Hilbert spaces are called bipartite quantum states.

Definition A.2.2 (Separability and Entanglement) [21]

A bipartite quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called separable if it can be written as a convex combination of product states, i.e., $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ with probabilities $p_i \geq 0$, $\sum_i p_i = 1$, $\rho_A^i \in \mathcal{D}(\mathcal{H}_A)$, and $\rho_B^i \in \mathcal{D}(\mathcal{H}_B)$. If ρ_{AB} is not separable, it is called inseparable or entangled. We denote the set of all separable (entangled) states by SEP (ENT).

To decide whether a state is entangled or separable is a long-standing, unresolved problem of quantum information theory called the separability problem.

Definition A.2.3 (Separability Problem) [21]

Given $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, decide whether or not $\rho_{AB} \in \text{SEP}$.

In general, any physical evolution of a quantum system is described by a quantum channel [48]. One requires that a channel maps one quantum state to another quantum state, even if it only acts on one part of a bipartite system. Additionally, it should act linearly in the convex set of quantum states. These requirements are fulfilled by linear, completely positive, and trace-preserving maps. More precisely:

Definition A.2.4 (Quantum channel)

A quantum channel $\mathcal{N}_{A \rightarrow B}$ is a linear map $\mathcal{N}_{A \rightarrow B} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$ with $\text{Tr}[\mathcal{N}_{A \rightarrow B}(X)] = \text{Tr}[X]$ for every $X \in \mathcal{L}_A$ and $\text{id}_k \otimes \mathcal{N}_{A \rightarrow B} \geq 0$ for all $k \geq 1$.

An often used operational paradigm of a bipartite setting of the two parties named Alice and Bob is that only local operations applied by each party separately are allowed in addition to classical communication between the parties. Hence, global quantum operations or quantum communication between Alice and Bob are forbidden. These operations are described by LOCC channels.

Definition A.2.5 (LOCC-channel) [48]

Let $\mathcal{X}$ be a finite alphabet and $\{\mathcal{M}^x\}_{x \in \mathcal{X}}$ be a set of completely positive and trace non-increasing maps such that $\sum_x \mathcal{M}^x$ is a channel. Let $\{\mathcal{N}^x\}_{x \in \mathcal{X}}$ be a set of quantum channels.

A one-way LOCC channel from Alice to Bob is a channel of the form

$$\mathcal{L}_{AB \rightarrow A'B'}^{\rightarrow} = \sum_{x \in \mathcal{X}} \mathcal{M}_{A \rightarrow A'}^x \otimes \mathcal{N}_{B \rightarrow B'}^x. \quad (\text{A.3})$$

A one-way LOCC channel from Bob to Alice is a channel of the form

$$\mathcal{L}_{AB \rightarrow A'B'}^{\leftarrow} = \sum_{x \in \mathcal{X}} \mathcal{N}_{A \rightarrow A'}^x \otimes \mathcal{M}_{B \rightarrow B'}^x. \quad (\text{A.4})$$

A channel is a LOCC channel if it is a composition of finitely many one-way LOCC channels from Alice to Bob and Bob to Alice.

Any LOCC protocol can be written as the following channel:

$$\mathcal{L}_{AB \rightarrow A'B'}^{\leftrightarrow} = \sum_{y \in \mathcal{Y}} \mathcal{R}_{A \rightarrow A'}^y \otimes \mathcal{S}_{B \rightarrow B'}^y \quad (\text{A.5})$$

for a finite alphabet $\mathcal{Y}$ and sets of completely positive trace non-increasing maps $\{\mathcal{R}\}_{y \in \mathcal{Y}}$ and $\{\mathcal{S}\}_{y \in \mathcal{Y}}$ such that $\sum_{y \in \mathcal{Y}} \mathcal{R}_{A \rightarrow A'}^y \otimes \mathcal{S}_{B \rightarrow B'}^y$ is a channel.

Entanglement distillation is the process of transforming input states to the qubit Bell state $|\Phi^+\rangle := 1/\sqrt{2} (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$ via LOCC. More precisely:

Definition A.2.6 (Entanglement Distillation, Distillability, Distillation Rate) [21]

Entanglement distillation describes any LOCC protocol that transforms N copies of a resource state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ to $M \leq N$ copies of the maximally entangled qubit state $\Phi^+ \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with nonzero probability:

$$\rho^{\otimes N} \xrightarrow[\text{LOCC}]{} (\Phi^+)^{\otimes M}. \quad (\text{A.6})$$

A state ρ is distillable if in the asymptotic limit $N \rightarrow \infty$, a distillation rate $R := \frac{M}{N}$ larger than zero can be achieved via entanglement distillation. If no such operations exist, ρ is called undistillable.

As any LOCC channel is a separable map [98], the set of separable states SEP is closed under LOCC operations. Consequently, any distillable state must be entangled. However, the converse is not true due to the existence of bound entanglement [51]. This phenomenon is present for entangled states with so-called positive partial transposition. The partial transposition $\Gamma : L(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow L(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined via the transposition T as $\rho^\Gamma := T_A \otimes \text{id}_B(\rho)$.

Definition A.2.7 (PPT Property) [21]

A bipartite quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ satisfying $\rho^\Gamma \geq 0$ is said to be positive under partial transposition (denoted $\rho \in \text{PPT}$), else non-positive under partial transposition ($\rho \in \text{NPT}$).

Like the set of separable states, also the set PPT is closed under LOCC operations [51]. Consequently, only NPT states can be distilled, as the maximally entangled state Φ^+ is also NPT. As the set $\text{PPT} \cap \text{ENT}$ is nonempty, there are entangled but undistillable states. These states are called bound entangled.

Definition A.2.8 (Free and Bound Entanglement) [21]

A bipartite quantum state ρ is called free entangled (denoted $\rho \in \text{FE}$) if it is distillable. If ρ is entangled and undistillable, it is called bound entangled ($\rho \in \text{BE}$).

Up to the date of writing this work, only PPT bound entangled states are known. The distillability problem is about the potential existence of undistillable NPT states [54]. More precisely:

Definition A.2.9 (Distillability Problem) [21]

Does the set equivalence

$$\text{SEP} \cup \text{BE} \equiv \text{PPT} \quad (\text{A.7})$$

hold generally?

One family for which bound entanglement can be observed are mixtures of maximally entangled states [62]. Using the partial trace $\text{Tr}_{B/A}$ [1], the reduced or marginal states of a bipartite quantum state ρ_{AB} are $\rho_{A/B} := \text{Tr}_{B/A}[\rho_{AB}]$.

Definition A.2.10 (Maximally Entangled States) [21]

A pure bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$ is said to be maximally entangled if the partial trace on either system yields the maximally mixed state, i.e., $\text{Tr}_A(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_B$ and $\text{Tr}_B(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_A$.

Maximally entangled states on a bipartite system are also called Bell states. Using a generalization of the Pauli operators of $d = 2$ to arbitrary dimension d via the Weyl-Heisenberg operators allows construction of Bell states in any dimension.

Definition A.2.11 (Weyl-Heisenberg Operators) [21]

$$W_{k,l} := \sum_{j=0}^{d-1} \omega^{jk} |j\rangle\langle j+l|, \quad k, l \in \mathbb{Z}_d. \quad (\text{A.8})$$

The Weyl-Heisenberg operators satisfy the Weyl relations, i.e.,

$$\begin{aligned} W_{k_1,l_1} W_{k_2,l_2} &= \omega^{l_1 k_2} W_{k_1+k_2, l_1+l_2}, \\ W_{k,l}^\dagger &= \omega^{kl} W_{-k,-l} = W_{k,l}^{-1}. \end{aligned} \quad (\text{A.9})$$

The Weyl relations (A.9) induce a group structure for the Weyl operators. More precisely, for fixed d , the Weyl relations imply that the set $\mathcal{V}_d := \{\omega^n W_{k,l}\}_{k,l,n \in \mathbb{Z}_d}$ forms a group, the so-called Weyl-Heisenberg group [29]. Identifying two elements of the group if they differ only by a phase yields the group $\mathcal{W}_d = \{W_{k,l}\}_{k,l \in \mathbb{Z}_d}$. Formally, we can define $\mathcal{W}_d$ with the normalizer $N_d := \{\omega^n \mathbb{1}_d\}_{n \in \mathbb{Z}_d}$ as the factor $\mathcal{W}_d := \mathcal{V}_d / N_d$. Moreover, the Weyl relations imply the group isomorphism

$$\mathcal{W}_d = \{W_{k,l}\}_{k,l \in \mathbb{Z}_d} \cong \mathbb{Z}_d \times \mathbb{Z}_d. \quad (\text{A.10})$$

Applying these d^2 unitary operators locally to one side of a maximally entangled state defines d^2 linearly independent, maximally entangled qudit states [30].

Definition A.2.12 (Standard Bell Basis) [21]

The standard Bell basis for a bipartite system of two qudits of dimension d is defined as

$$|\Omega_{k,l}\rangle := (W_{k,l} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle := (W_{k,l} \otimes \mathbb{1}_d) \frac{1}{\sqrt{d}} \sum_{i \in \mathbb{Z}_d} |i\rangle \otimes |i\rangle, \quad (\text{A.11})$$

$$\Omega_{k,l} \equiv |\Omega_{k,l}\rangle\langle\Omega_{k,l}|, \quad k, l \in \mathbb{Z}_d. \quad (\text{A.12})$$

For $d = 2$, $|\Omega_{0,0}\rangle$ is often written as $|\Phi^+\rangle$.

Due to the presence of bound entangled states, a central system of interest in this work are Bell-diagonal states, i.e., quantum states with diagonal density matrix in the basis $\Omega_{k,l}$, which can be consequently interpreted as mixtures of Bell-states [1].

Definition A.2.13 (Standard Bell-diagonal States, Magic Simplex) [21]

$$\mathcal{M}_d := \left\{ \rho = \sum_{k,l \in \mathcal{Z}_d} c_{k,l} \Omega_{k,l} \mid \sum_{k,l \in \mathcal{Z}_d} c_{k,l} = 1, c_{k,l} \geq 0 \right\}. \quad (\text{A.13})$$

The mixing probabilities $c_{k,l}$ form a geometric simplex in real space. The group properties of the Weyl-Heisenberg operators imply certain symmetries, motivating the name “magic simplex” for $\mathcal{M}_d$.

Definition A.2.14 (Entanglement conserving symmetries) [21, 73]

Entanglement conserving symmetries are unitary or anti-unitary transformations that conserve the entanglement class of states in $\mathcal{M}_d$ that can be represented by permutations on the standard Bell states. These symmetries can be generated by the following permutations:

- $t_{p,q} : \Omega_{k,l} \rightarrow \Omega_{k+p,l+q}$, ($p, q \in \mathcal{Z}_d$) (“translation”)
- $m : \Omega_{k,l} \rightarrow \Omega_{-k,l}$ (“momentum inversion”)
- $r : \Omega_{k,l} \rightarrow \Omega_{l,-k}$ (“quarter rotation”)
- $v : \Omega_{k,l} \rightarrow \Omega_{k+l,l}$ (“shear”)

Entanglement conserving symmetries map the sets $\text{SEP} \cap \mathcal{M}_d$, $\text{PPT} \cap \mathcal{M}_d$, $\text{BE} \cap \mathcal{M}_d$ and $\text{NPT} \cap \mathcal{M}_d$ onto themselves.

A crucial tool to detect entanglement are entanglement witnesses.

Definition A.2.15 (Entanglement Witness) [21]

A Hermitian operator W is an entanglement witness if the following conditions are met:

- $\text{Tr}[W\sigma] \geq 0 \ \forall \sigma \in \text{SEP}$
- $\exists \rho \in \mathcal{D}(\mathcal{H})$ such that $\text{Tr}[W\rho] < 0$.

An entanglement witness W is called Bell-diagonal if it can be written as

$$W = \sum_{k,l} \kappa_{k,l} \Omega_{k,l}. \quad (\text{A.14})$$

Besides their use in defining a Bell basis, the Weyl operators also provide a so-called nice error operator basis [66]. Acting on N qudits, the set of Weyl errors is given by

Definition A.2.16 (Weyl Errors) [77]

$$\mathcal{E}_N := \left\{ W(e) \mid e = (\vec{k}, \vec{l}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \right\} := \left\{ \bigotimes_{n=1}^N W_{k_n, l_n} \mid k_n, l_n \in \mathcal{Z}_d \right\}. \quad (\text{A.15})$$

Using the Weyl relations (A.9), the group structure of single Weyl operators naturally extends to the set of Weyl errors by identifying elements that differ only by a phase, implying an isomorphism to $\mathcal{Z}_d^N \times \mathcal{Z}_d^N$. Commutative (abelian) subgroups of a group of error operators are called stabilizer groups or stabilizers. For the group of Weyl errors $\mathcal{E}_N$, and its isomorphism to $\mathcal{Z}_d^N \times \mathcal{Z}_d^N$, we define the following objects:

Definition A.2.17 (Stabilizer, generating operators, generating elements) [77]

A stabilizer group, or stabilizer, S is an abelian subgroup of $\mathcal{E}_N$. If $\{W(g_1), \dots, W(g_p)\}$ forms a minimal generating set, each $W(g_j)$ is called a generating operator, and we write $S = \langle W(g_1), \dots, W(g_p) \rangle$. Each $g_j \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ is called a generating element. Given one choice of generating elements, the corresponding subgroup of $\mathcal{Z}_d^N \times \mathcal{Z}_d^N$, i.e., $G_S := \{g \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \mid W(g) \in S\}$ is also denoted as $G_S = \langle g_1, \dots, g_p \rangle$.

The simultaneous eigenspaces of stabilizer operators are called (stabilizer) codespaces and the corresponding eigenvectors are called (stabilizer) codewords.

Definition A.2.18 (Codeword) [77]

Let $\mathcal{H}_A^{\otimes N} \cong \bigoplus_{x \in \mathcal{Z}_d^p} \mathcal{Q}(x)$ be decomposed into d^{N-p} -dimensional codespaces of a stabilizer S . Let $\{|u_{x,k}\rangle\}_{x \in \mathcal{Z}_d^p, k \in \mathcal{Z}_d^{N-p}}$ be an orthonormal basis of $\mathcal{H}_A^{\otimes N}$ with $|u_{x,k}\rangle \in \mathcal{Q}(x)$, i.e., $W(g_j)|u_{x,k}\rangle = w_{x_j}|u_{x,k}\rangle$, $\forall k, j$. The vectors $|u_{x,k}\rangle \in \mathcal{H}_A^{\otimes N}$ are called codewords of S in $\mathcal{H}_A^{\otimes N}$. The codewords of S^* in $\mathcal{H}_B^{\otimes N}$ are denoted by $|u_{x,k}^*\rangle \in \mathcal{H}_B^{\otimes N}$.

Encoding and decoding quantum states in a basis of codewords in combination with projections to the codespaces called stabilizer measurements constitute stabilizer operations.

Definition A.2.19 (Encoding) [77]

Let $\{|x\rangle\}_{x \in \mathcal{Z}_d^p}$ be the computational basis of $\mathcal{H}_A^{\otimes p}$ and $\{|k\rangle\}_{k \in \mathcal{Z}_d^{N-p}}$ be the computational basis of $\mathcal{H}_A^{\otimes N-p}$. A unitary operator U on $\mathcal{H}_A^{\otimes N} \cong \mathcal{H}_A^{\otimes p} \otimes \mathcal{H}_A^{\otimes N-p}$ is an encoding for a stabilizer S in $\mathcal{H}_A^{\otimes N}$ if $\forall x \in \mathcal{Z}_d^p, k \in \mathcal{Z}_d^{N-p} : U(|x\rangle \otimes |k\rangle) =: |u_{x,k}\rangle$ is a codeword of S in $\mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes N}$.

Definition A.2.20 (Stabilizer measurement)

Given a stabilizer S with codespaces $\mathcal{Q}(x)$, let $\mathcal{P}(x) : \mathcal{H}_A^{\otimes N} \rightarrow \mathcal{Q}(x)$ be the projection to the codespace $\mathcal{Q}(x)$. One has $\mathcal{P}(x) = \sum_k |u_{x,k}\rangle\langle u_{x,k}|$ for any basis of codewords $|u_{x,k}\rangle$. A stabilizer measurement corresponds to the projector-valued measure (PVM) $\{\mathcal{P}(x)\}_x$.

B. Publications

B.1. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits

BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits

Christopher Popp ¹

¹ Faculty of Physics, University of Vienna, Währingerstrasse 17, 1090, Vienna Austria

DOI: [10.21105/joss.04924](https://doi.org/10.21105/joss.04924)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Jarvist Moore Frost](#) 

Reviewers:

- [@meandmytram](#)
- [@Roger-luo](#)

Submitted: 04 October 2022

Published: 30 January 2023

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

In the field of quantum information and technology, entanglement in quantum systems called qudits is regarded as resource for quantum-based or quantum-assisted information processing tasks. It allows new ways of information processing like quantum computation or quantum teleportation and provides possibilities for speedups in a variety of algorithms with applications like search or optimization. Despite its theoretical and practical relevance, there is no general method to determine whether a given quantum state is entangled or not due to a special and hard to detect form of entanglement called bound entanglement. Bipartite Bell states are maximally entangled states of two qudits and are of high relevance for application in quantum technologies. Mixtures of those states can be entangled, including its bound form, or separable. Leveraging their special properties, Bell states can be numerically generated and analyzed for their entanglement properties by various methods implemented in this Julia package.

Statement of need

BellDiagonalQudits is an Quantum Information-affiliated Julia package for generation and entanglement analyses of mixed maximally entangled bipartite qudits in general dimension. The API for BellDiagonalQudits provides an user-friendly interface to generate representations of Bell diagonal quantum states and to analyze their entanglement properties with various general or specialized criteria to detect entanglement or separability. Leveraging geometric properties of a certain class of mixed Bell states that are related by Weyl transformations, BellDiagonalQudits combines known analytical results by Baumgartner et al. (2007) and numerical methods for quantum state representation and analysis. It leverages and depends on the Julia package QuantumInformation, the convex sets of LazySets (Forets & Schilling, 2021) and the optimization methods of Optim (Mogensen & Riseth, 2018).

BellDiagonalQudits was designed to be used by researchers in quantum science and quantum information theory. It has already been used in multiple scientific publications, e.g. in Popp & Hiesmayr (2022a) and Popp & Hiesmayr (2022b) in the context of entanglement classification and detection of bipartite qudits in dimension three and four. The combination of efficient state generation via random sampling or deterministic procedures and implementation of both frequently used and specialized entanglement and separability detectors supports the research of entanglement in several ways. From a general point of view, entangled Bell states are well accessible for powerful methods of entanglement and separability detection, leveraging their symmetries and geometric properties. It was shown in that a significant share of the group of mixed Bell states related by Weyl transformations are bound entangled (Hiesmayr, 2021; Popp & Hiesmayr, 2022a; Popp & Hiesmayr, 2022b), offering a systematic way to generate and investigate those states with respect to the separability problem in different dimensions. In addition to general methods applicable to any Bell diagonal state, BellDiagonalQudits provides features to generate the special symmetries of Bell states that are related via Weyl

transformations. These symmetries are leveraged for improved entanglement classification and the numerical generation of specialized entanglement witnesses in any dimension. Furthermore, the implemented methods of `BellDiagonalQudits` can be used in various quantum information processing tasks involving Bell diagonal states in any dimension like Quantum Key Distribution or entanglement verification. `BellDiagonalQudits` uses and integrates well with the general interface of `QuantumInformation` allowing the investigation of Bell states in the context of quantum channels, entanglement measures or entropy.

Relation to research projects

The methods of `BellDiagonalQudits` to generate and analyze Bell diagonal states in general dimension are based on analytical properties summarized by Baumgartner et al. (2007). Extensions of those methods and efficient implementation in `BellDiagonalQudits` enabled the detailed analysis by Popp & Hiesmayr (2022a) of Bell diagonal qudits in three dimensions (qutrits), providing an operational solution to the separability problem for those states. Additionally the relative shares of separable and (bound) entangled states were precisely determined among the Bell diagonal states. In Popp & Hiesmayr (2022b), higher dimensions were considered, focusing on a detailed comparison and geometric properties of separable states in dimension three and four.

Package information

`BellDiagonalQudits` is available on Github at <https://github.com/kungfugo/BellDiagonalQudits.jl>. The package documentation is available at <https://kungfugo.github.io/BellDiagonalQudits.jl/dev/> and provides examples of usage.

Acknowledgments

I acknowledge support from Beatrix C. Hiesmayr for review and validation of implemented methods.

References

- Baumgartner, B., Hiesmayr, B. C., & Narnhofer, H. (2007). A special simplex in the state space for entangled qudits. *Journal of Physics A: Mathematical and Theoretical*, 40, 7919. <https://doi.org/10.1088/1751-8113/40/28/S03>
- Forets, M., & Schilling, C. (2021). LazySets.jl: Scalable symbolic-numeric set computations*. *Proceedings of the JuliaCon Conferences*, 1(1), 97. <https://doi.org/10.21105/jcon.00097>
- Hiesmayr, B. C. (2021). Free versus bound entanglement, a NP-hard problem tackled by machine learning. *Scientific Reports*, 11. <https://doi.org/10.1038/s41598-021-98523-6>
- Mogensen, P. K., & Riseth, A. N. (2018). Optim: A mathematical optimization package for Julia. *Journal of Open Source Software*, 3(24), 615. <https://doi.org/10.21105/joss.00615>
- Popp, C., & Hiesmayr, B. C. (2022a). Almost complete solution for the NP-hard separability problem of bell diagonal qutrits. *Scientific Reports*, 12, 12472. <https://doi.org/10.1038/s41598-022-16225-z>
- Popp, C., & Hiesmayr, B. C. (2022b). Bound entanglement of bell diagonal pairs of qutrits and ququarts: A comparison. arXiv. <https://doi.org/10.48550/ARXIV.2209.15267>

**B.2. Almost complete solution for the NPhard separability
problem of Bell diagonal qutrits**



OPEN

Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits

Christopher Popp[✉] & Beatrix C. Hiesmayr

With a probability of success of 95% we solve the separability problem for Bell diagonal qutrit states with positive partial transposition (PPT). The separability problem, i.e. distinguishing separable and entangled states, generally lacks an efficient solution due to the existence of bound entangled states. In contrast to free entangled states that can be used for entanglement distillation via local operations and classical communication, these states cannot be detected by the Peres–Horodecki criterion or PPT criterion. We analyze a large family of bipartite qutrit states that can be separable, free entangled or bound entangled. Leveraging a geometrical representation of these states in Euclidean space, novel methods are presented that allow the classification of separable and bound entangled Bell diagonal states in an efficient way. Moreover, the classification allows the precise determination of relative volumes of the classes of separable, free and bound entangled states. In detail, out of all Bell diagonal PPT states $81.0\% \pm 0.1\%$ are determined to be separable while $13.9 \pm 0.1\%$ are bound entangled and only $5.1 \pm 0.1\%$ remain unclassified. Moreover, our applied criteria are compared for their effectiveness and relation as detectors of bound entanglement, which reveals that not a single criterion is capable to detect all bound entangled states.

Quantum information theory can be understood as the study of processing tasks that can be accomplished using quantum mechanical systems¹. Quantum technology uses quantum phenomena for improved or new technological applications. Prominent fields include quantum computing, quantum communication, quantum simulation, quantum metrology or quantum cryptography. Most of existing experimental and theoretic applications of quantum technology are based on two-level quantum systems, called “qubits” ($d = 2$). Recently however, interest in higher dimensional systems, i.e. “qutrits” ($d = 3$) or “qudits” (general d) is growing^{2,3}. Although experimentally more challenging, qudits are expected to have certain advantages and show new phenomena. This work wants to explore one of these phenomena.

Entanglement is one of the main resources to realize information processing with better performance than any classical system and has been shown to exist in many physical systems and is intensively explored for potential applications ranging from quantum teleportation to the detection of cancer in human beings^{4–7}. Despite its theoretical and practical relevance, there is no general method to determine whether a bipartite quantum state is entangled or not, in which case the state is separable. This problem is strongly connected to a special form of entanglement called “bound entanglement” for which a first example has been found by the Horodecki family in 1998⁸ and has been studied since then, e.g. Refs.^{9–19}, and has been experimentally verified with photons entangled in the orbital momentum degrees of freedom²⁰. Often, several weakly entangled states can be transformed via local operations and classical communication (LOCC) to a smaller set of strongly entangled states. This process is called entanglement distillation⁸ and is of high relevance for practical applications that require strong distillable or “free” entanglement as a resource. For bound entangled states, this distillation is not possible. However, they can be generated by mixing several free entangled states, binding or preventing their entanglement to be further used in applications that require distillation, thus motivating the name. In order to avoid operations that would bind the resource of entanglement in bound entangled states through this irreversible mixing process, it is important to have information about the structure of bound entangled states. While free entanglement can be efficiently detected by the the Peres–Horodecki criterion (also known as PPT criterion)^{21,22} in a bipartite system, the detection of bound entanglement has been proved to be NP-hard for bipartite systems and thus intractable for higher dimensions²³. Up to now, there is no complete understanding about why this special form of entanglement exists, what its existence implies for nature or what its exact capabilities with respect to quantum information theoretic tasks are. Moreover, there is no general method known to construct bound entangled states.

Faculty of Physics, University of Vienna, Währingerstrasse 17, 1090 Vienna, Austria. ✉email: poppchristopher@pm.me

In this paper we investigate mixtures of maximally entangled, bipartite, 3×3 -dimensional (“Bell”)-states. One significant subset of this class of Bell diagonal states has been shown^{20,24–26} to be especially useful to analyze (bound) entanglement due to a representation, which allows a precise geometrical interpretation of the system. Depending on the mixing probabilities of pure maximally entangled states, the mixed state can be separable, bound entangled or free entangled. Although geometric properties allow analytical insight in the structure of entanglement²⁴, the problem of entanglement classification, especially the detection of bound entanglement, remains a hard problem. Several methods to detect bound entangled states exist^{13–19}, but these methods based on specifically constructed non-decomposable entanglement witnesses or other constructions are strongly limited in the number of states they can detect. Zyczkowski numerically analyzed^{27,28} general higher dimensional systems with respect to the separability problem and entanglement classification and determined approximate relative volumes of separable and bound entangled states for different dimensions. Among other results it was shown that the relative volume of separable and bound entangled states decreases exponentially with the dimension of the system. Recently, Hiesmayr analyzed mixtures of Bell states for $d = 3$ to determine relative volumes of separable and (bound) entangled states and to detect relevant substructures in the set of bound entangled states²⁹.

This work intends to set up a numeric framework that allows the efficient and precise determination of the entanglement class of Bell diagonal states in general dimension and to apply it to the case $d = 3$. An alternative, more precise and efficient way to generate and analyze states in order to estimate the relative volumes of entanglement classes is presented. New methods are developed to efficiently solve the separability problem in the considered system for almost all states. Finally, the methods are compared in their effectiveness for the detection of bound entanglement.

The paper is organized as follows: First, we define the systems to be analyzed and several subsets of relevance for the investigation. Then, a method to efficiently generate states of that system based on random sampling is presented and shown to have significant advantages compared to the methods used in previous related analyses²⁹. We show how symmetries of the system can be numerically generated and leveraged for entanglement classification. We then give a short overview over implemented and used methods for the detection of (bound) entangled and separable states, including a new numerical sufficient criterion for separability. Finally, these methods are applied to the system for $d = 3$ and some relevant subset, allowing to efficiently determine the entanglement class of 95% of all PPT states.

Methods

In this paper we analyze mixtures of maximally entangled orthonormal Bell states $|\Omega_{k,l}\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ for $k, l = 0, 1, \dots, (d-1)$ in the Hilbert space $\mathcal{H}$ of a bipartite system containing two qudits of dimension d . Suppose Alice and Bob share a bipartite maximally entangled state $|\Omega_{00}\rangle \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$. A basis of orthonormal Bell states $|\Omega_{k,l}\rangle$ can be generated by applying the Weyl operators³⁰ $W_{k,l} \equiv \sum_{j=0}^{d-1} w^{j \cdot k} |j\rangle \langle j+l \pmod{d}|$, $w = e^{\frac{2\pi i}{d}}$, to one of the subsystems, without loss of generality Alice’s qudit: $|\Omega_{k,l}\rangle \equiv W_{k,l} \otimes \mathbb{1}_d |\Omega_{00}\rangle$. The density matrices of these pure basis states are then the “Bell projectors” $P_{kl} \equiv |\Omega_{k,l}\rangle \langle \Omega_{k,l}|$. Mixtures of these states define the system of interest $\mathcal{M}_d$ in this work:

$$\mathcal{M}_d \equiv \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right\}. \quad (1)$$

$\mathcal{M}_d$ has also been named “magic simplex”^{24–26}, referring to the “magic Bell basis” introduced by Wootters and Hill³¹ and the fact that it can be represented as simplex in real space, identifying the mixing probabilities $c_{k,l}$ as coordinates with the Bell projectors $P_{k,l}$ lying at the vertices of the simplex. A special property of this set is that the reduced states, i.e. the partial trace with respect to one of the subsystems, of all states in $\mathcal{M}_d$ are maximally mixed. They are therefore said to be “locally maximally mixed” states. Note that for $d > 2$, $\mathcal{M}_d$ does not contain all locally maximally mixed states²⁵.

Depending on the focus of the analysis, several subsets and families of states are of interest and have been discussed in literature^{26,32}. For the entanglement analysis, the following families are especially relevant:

Enclosure polytope. The enclosure polytope is defined as

$$\mathcal{E}_d \equiv \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \in [0, \frac{1}{d}] \right\}. \quad (2)$$

It was shown²⁴ that all states that lie outside of the enclosure polytope are necessarily entangled. Since they can be detected by the PPT or Peres-Horodecki criterion, they can be distilled by local operation and classical communication (LOCC), or equivalently, are free and not bound entangled⁸.

Kernel polytope. The kernel polytope is another geometric object that allows to determine the entanglement class of a given state represented by its coordinates in $\mathcal{M}_d$. It is defined as convex mixture of certain separable states, named “line states” ρ_α ²⁴:

$$\mathcal{K}_d \equiv \left\{ \rho = \sum_{\alpha} \lambda_{\alpha} \rho_{\alpha} \mid \lambda_{\alpha} \geq 0, \sum_{\alpha} \lambda_{\alpha} = 1 \right\}. \quad (3)$$

The line states ρ_α are related to cyclic subgroups (or more general sublattices in higher dimensions) of the linear ring structure induced by the Weyl operators. As those are separable states, each state in the $\mathcal{H}_d$ is separable by construction.

Family A. Family A is defined by a mixture of three Bell states and the maximally mixed state $\rho_{mm} = \frac{1}{d^2} \sum_{k,l=0}^{d-1} P_{k,l} = \frac{1}{d^2} \sum_{i,j=0}^{d-1} |ij\rangle\langle ij|$ for $d = 3$. The three Bell states are required to be on a phase space line²⁴, i.e. one Bell state is chosen and the other two are generated by application of one chosen Weyl transformation. Choosing without loss of generality the Weyl transformation to be $W_{0,1} \otimes \mathbb{1}_d$, the states of Family A can explicitly be written as

$$\mathcal{F}_A \equiv \{\rho_A = \alpha P_{00} + \beta P_{01} + \gamma P_{02} + (1 - \alpha - \beta - \gamma) \rho_{mm} \mid \rho_A \in M_3\}. \quad (4)$$

Leveraging the high symmetry of such states, an optimal entanglement witness was found²⁴ and later related to the quasi-pure approximation of the concurrence criterion³². It is therefore possible to determine the entanglement class of all states in this family.

Numerical generation of states. The representation of the system as $(d^2 - 1)$ -dimensional simplex with the d^2 Bell states as vertices and the probabilities as “barycentric coordinates” allows the efficient geometric representation of states as points in the simplex. In this representation, states for the whole simplex, subsets of finite volume or hyper-planes can be equivalently described and generated as points in d^2 -dimensional Euclidean space. This can be achieved by, e.g., random sampling according to a certain distribution or by some deterministic discretization procedure. In the following, two ways that allow the estimation of relative volumes of separable, bound and free entangled states are shortly described.

Uniform random sampling of states. The linear, real (sub-)spaces presented above allow uniform sampling of points within and therefore uniformly distributed states can be generated in $\mathcal{M}_d$, $\mathcal{E}_d$ or $\mathcal{F}_A$. Classification of these states allows a probabilistic estimation of relative volumes of the entanglement classes by using the relative number/frequency of classified states as estimator. As the number of required states for a valid estimation only depends of the relative volumes of the classes and not on the dimension, this method of state creation is favorable for volume estimation in higher dimensions. Another advantage of random sampling compared to some deterministic procedure is that the expected relative number of states in a given class only depends on the volume and not on the (unknown) specific geometric shape of the set of states within that class.

Let us demonstrate the validity of this method to estimate the relative volumes of entanglement classes for bipartite qubits (i.e. $d = 2$). In this case, the separability problem can be solved analytically. It has been shown that for $d = 2$, all entangled states are free entangled and can be detected with the PPT criterion^{21,22} and therefore $\mathcal{M}_2$ contains no bound entanglement. It is also known that the kernel polytope $\mathcal{K}_2$ contains all separable states and that the relative volumes of both classes in $\mathcal{M}_2$ are exactly 50%. Figure 1 shows the relative frequency of separable states for uniform samples in $\mathcal{M}_2$. For different sample sizes, the empirical mean and standard deviation of 10 runs are presented. One observes that the relative frequency converges with growing number of states to the known equal sized relative volumes of 0.5 with increasing precision. Given a random sample of size N , the number of states within a certain class of relative volume p should be distributed according to the binomial distribution. The expected number of states in that class is then $N \cdot p$ with standard deviation of $\sqrt{Np(1-p)}$. In case of $\mathcal{M}_2$, the probability that a generated state is PPT and in this special case also separable is $p = 0.5$.

States on an equidistant lattice. One way to discretize the space of interest is using a constant increment in all dimensions. Given $k \leq (d^2 - 1)$ independent parameters, a certain range R of values they can assume and a number of steps s to divide the range in, one obtains $(s + 1)^k$ lattice points. Generating states on a fixed lattice offers the advantage that these states have many symmetries, allowing to reduce the number of states to explicitly classify. On the other hand, generating the full lattice is quickly too computationally expensive with growing dimension, due to the exponential dependency of the number of lattice points given a certain number of steps. This makes this discretization computationally inefficient in higher dimensions, because the step size needs to be small enough to capture the entanglement structure. Another disadvantage of this set of states is that they are only approximately uniformly distributed for small increments. One example for that is the over-representation of states in certain areas depending on the chosen discretization. As a consequence, classifying the states and counting the relative occurrences as estimator for the relative volumes of the entanglement classes is unbiased only in the limit of large s . Figure 2 compares the relative frequency of states that are entangled according to the PPT criterion for random samples and states on a fixed lattice. For $d = 2$, the frequencies are relative to the whole simplex $\mathcal{M}_2$, while for $d = 3$ the enclosure polytope $\mathcal{E}_3$ is used. The ranges of the coordinates are divided in a certain amount of steps to generate the states on a lattice. For comparison, an equal number of random states is generated and analyzed. In Fig. 2 (left) one observes that the randomly generated states quickly converge to the known relative volume of 0.5. The lattice states show a strong dependence on the number of steps. While for an even number of steps the relative frequency approaches the volume from below, for odd numbers the convergence is slower and from above. Figure 2 (right) shows that already for $d = 3$, the number of lattice states that can be generated by standard computational means is not high enough to avoid a significant bias compared to the quickly converging randomly generated states.

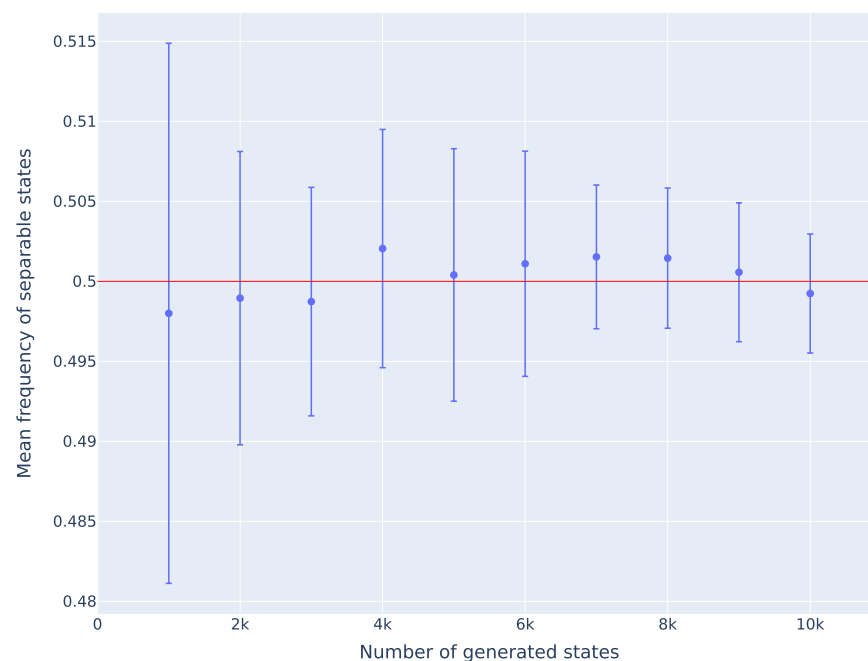


Figure 1. Relative frequencies of separable states in $\mathcal{M}_2$. For each sample size, 10 sets of randomly generated states are generated. The figure shows the empirical mean and standard deviation for the frequencies of separable states.

Generation of symmetry transformations conserving the entanglement property. The phase space structure of $\mathcal{M}_d$ implies a set of “entanglement-conserving” symmetry transformations related to point transformations and translations in the discrete phase space induced by Weyl operators^{24,25}. They act as permutations of the generating Bell projectors, or equivalently, as permutations of the coordinates in $\mathcal{M}_d$ and are generated by elementary transformations, i.e. generators of the involved symmetry group. These generators can be identified with translation, momentum inversion, quarter rotation and vertical shear in the induced discrete phase space (see Supplementary Appendix A3). It was shown that these symmetries conserve entanglement in $\mathcal{M}_d$. A simple argument (see A1 in the Supplementary Appendix) also shows that the entanglement class is conserved, meaning that the classification of free, bound and separable states is not changed by applying any element of this group. The generators and their actions can be represented as specific permutations, allowing to generate all elements by succeeding application of several symmetry transformations. Since the phase space contains a finite number of elements, also the number of permutations is finite. This allows to determine all symmetries generated by the above generators in arbitrary dimension. The number of symmetries that can be generated by those generators grows quickly with the dimension of the subsystems. For $d = 2$, 24 such symmetries exist, while for $d = 3$ already 432 symmetries can be found. Leveraging these symmetries is crucial to the analysis of entanglement in $\mathcal{M}_d$, significantly improving the differentiation of bound entangled from separable states.

Sufficient criteria for entanglement. The problem to decide whether a given state is separable or entangled is generally a NP-hard problem²³. No general solution by polynomial in time algorithms is known and it is often denoted the “separability problem”. Likewise, there is no efficient method to solve the separability problem for the mixtures of maximally entangled states $\mathcal{M}_d$ in general dimension d . However, several sufficient criteria to detect entanglement are known and can be used for entanglement classification. The most effective ones for states in $\mathcal{M}_d$ are shortly stated below and enumerated by E1, E2,....

E1: PPT criterion. The “Positive Partial Transpose (PPT)” criterion²¹, also named “Peres–Horodecki” criterion states that if the partial transpose of a bipartite state has at least one negative eigenvalue (in this case the state is said to be “NPT”), it is entangled. The partial transpose Γ acts on the basis states of a bipartite state as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^\Gamma \equiv |i\rangle\langle j| \otimes |l\rangle\langle k|$. It can be calculated efficiently, but for $d \geq 3$ it is only sufficient, not necessary, for entanglement. In this case, also states that have a positive partial transpose can be entangled. Only entangled NPT-states can be used for entanglement distillation are therefore denoted as free entangled.

E2: realignment criterion. The realignment criterion³³ is structurally similar to the PPT criterion. The realignment operation R acts as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)_R \equiv |i\rangle\langle k| \otimes |j\rangle\langle l|$. The criterion states that if the sum of singular values

of the realigned state σ_R are larger than 1, i.e. if $\text{tr} \sqrt{\sigma_R^\dagger \sigma_R} > 1$, then σ is entangled. Like the PPT criterion, it can also be computed efficiently, but is only sufficient for entanglement. It is neither stronger or weaker than the PPT criterion, so it can detect PPT entangled states, but generally does not detect all NPT-states.

E3: quasi-pure concurrence criterion. The quasi-pure approximation C_{qp} of the concurrence³⁴ provides another sufficient entanglement criterion. States for which the concurrence is positive are entangled, but for $d > 2$ the concurrence allows only numerical estimates for mixed states. The quasi-pure approximation³² provides an easy way to compute lower bound for the concurrence and can therefore be used for entanglement detection. Considering states of $\mathcal{M}_d$, an explicit form for the approximation can be derived. A state $\rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \in \mathcal{M}_d$ is entangled according to the quasi-pure concurrence criterion if $C_{qp}(\rho) = \max(0, S_{nm} - \sum_{(k,l) \neq (n,m)} S_{k,l}) > 0$ where (n, m) is a multi-index of the coordinate of the largest mixing probability $\{c_{k,l}\}$ and $S_{k,l}$ are explicitly given by³²:

$$S_{k,l} = \sqrt{\frac{d}{2(d-1)} c_{k,l} \left[\left(1 - \frac{2}{d}\right) c_{n,m} \delta_{k,n} \delta_{l,m} + \frac{1}{d^2} c_{(2n-k) \bmod d, (2m-l) \bmod d} \right]}. \quad (5)$$

E4: MUB criterion. In a Hilbert space of dimension d , a set of orthonormal bases $\{B_k\}$ and $B_k = \{|i_k\rangle \mid i = 0, \dots, (d-1)\}$ is called “mutually unbiased bases (MUB)” if $\forall k \neq l$:

$$|\langle i_k | j_l \rangle|^2 = \frac{1}{d} \quad \forall i, j = 0, \dots, (d-1). \quad (6)$$

Given m MUBs, it was shown^{20,35} that the sum of all “mutual predictabilities” I_m is bounded from above for separable states ρ_s :

$$I_m(\rho_s) = \sum_{k=1}^m C_k = \sum_{k=1}^m \sum_{i=0}^{d-1} \langle i_k | \otimes \langle i_k | \rho_s | i_k \rangle \otimes | i_k \rangle \leq 1 + \frac{m-1}{d}. \quad (7)$$

At most $d+1$ MUBs exist^{36,37} in which case $I_{d+1}(\rho_s) \leq 2$ for all separable states ρ_s . Conversely, if an unclassified state exceeds this upper bound, it is entangled. In order to detect bound entanglement in $d=3$, the C_k can be modified to the following form:

$$C_1 = \sum_{i=0}^2 \langle i_1 | \otimes \langle (i_1 + 2 \bmod 3)^* | \rho_s | i_1 \rangle \otimes | i_1 + 2 \bmod 3 \rangle^*, \quad (8)$$

$$C_k = \sum_{i=0}^2 \langle i_k | \otimes \langle i_k^* | \rho_s | i_k \rangle \otimes | i_k^* \rangle, \quad k = 2, 3, 4. \quad (9)$$

Here, i_k^* denotes the complex conjugate vector element. In this form, like the realignment and quasi-pure concurrence criteria, the MUB criterion allows the detection of PPT entangled states, which was also experimentally demonstrated for entangled photons²⁰.

E5: numerically generated entanglement witnesses. Leveraging the fact that separable states form a convex set, entanglement witnesses³⁸ (“EWs”) are an important tool to detect entangled states. An EW W is an observable which implies an upper bound U and, as recently shown³⁹, also a lower, mostly nontrivial, bound L ($U, L \in \mathbb{R}$), for separable states ρ_s :

$$L \leq \text{tr}[\rho_s W] \leq U. \quad (10)$$

If a state ρ suffices $\text{tr}[\rho W] \notin [L, U]$ for some W , then it is entangled and said to be “detected by W ”. Originally EWs were introduced considering the upper bound $U = 0$ only. It has been shown²⁴ that for each entangled state $\rho \in \mathcal{M}_d$ there exists an EW W_ρ such that W_ρ detects ρ . In this sense, EWs are universal, because each entangled state can be detected by an, unfortunately generally unknown, EW. If $\rho \in \mathcal{M}_d$, it suffices to consider EWs of the form $W = \sum_{k,l=0}^{d-1} \kappa_{k,l} P_{k,l}$ with $\kappa_{k,l} \in [-1, 1]$ to detect all entangled states²⁴. In this case $\rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \in \mathcal{M}_d$ and we have $\text{tr}[\rho W] = \sum_{k,l=0}^{d-1} c_{k,l} \kappa_{k,l} \equiv c \cdot \kappa$ where the dot indicates the standard scalar product of the d^2 -dimensional vectors c and κ collecting the coefficients $c_{k,l}$ and $\kappa_{k,l}$. Thus, in the geometric representation of $\mathcal{M}_d$, any EW defines two $(d^2 - 1)$ -dimensional hyper-planes via $c_L \cdot \kappa = L$ and $c_U \cdot \kappa = U$ and corresponding halfspaces. Any state that is represented by a point outside of the intersection of these halfspaces is detected as entangled. The difficulty of using EWs lies of course in determining the bounds L and U for the set of separable states. An efficient parameterization of unitaries⁴⁰ and numerical optimization⁴¹ is used to determine the bounds for our case (see Supplementary Appendix 2 for more details).

Sufficient criteria for separability. Since the presented criteria for entanglement are only sufficient and generally not necessary, it is highly desirable to develop methods to directly classify separable states as well. Few analytical sufficient criteria for separability exist, but the geometric characterization and properties of $\mathcal{M}_d$

allow the development of an effective procedure to identify states of $SEP \cap \mathcal{M}_d$. Here we present a new method based on the extension of the convex hull of known separable states (S1) and a criterion with close relation to the states of $\mathcal{M}_d$ (S2).

S1: extended kernel criterion. Given a finite set of known separable states and their representation in $\mathcal{M}_d$, a convex polytope can be constructed by building the convex hull of those vertices. All states of $\mathcal{M}_d$ with coordinates within this polytope are separable as well. The problem to decide whether a given Bell diagonal state is a convex combination of known existing separable states is then equivalent to the standard problem of linear programming to decide whether a given point lies within a convex polytope. Several numerical implementations to solve that problem exist, e.g. Ref.⁴², which was used for this work. Obviously, in order to check an unknown state $\rho \in \mathcal{M}_d$ for separability, a set of known separable states as vertices to build the convex hull are required. Since the effectivity of the check depends on the volume of the spanned polytope, it is advantageous to use vertices that cover uniformly distributed spatial angles and are as close to the boundary of the convex set of separable states as possible. The line states ρ_α that build the kernel polytope $\mathcal{K}_d$ meet those requirements, as they are known to be on the surface of SEP^{24} . However, more separable states/vertices are needed to cover approximately all separable states in $\mathcal{M}_d$. For the results presented in this work, additional separable states were generated and multiplied by using the entanglement-class-conserving symmetries to generate more vertices to extend the separable kernel.

S2: Weyl/spin representation criterion. The Weyl operators $W_{k,l}$ satisfy useful relations to simplify required calculations for states of $\mathcal{M}_d$. Two of those are:

$$W_{k_1,k_2} W_{l_1,l_2} = w^{k_2 l_1} W_{k_1+l_1,k_2+l_2}, \quad (11)$$

$$W_{k_1,k_2}^\dagger = w^{k_1 k_2} W_{-k_1,-k_2} = W_{k_1,k_2}^{-1}. \quad (12)$$

It follows that the Weyl operators form an orthogonal family of unitary matrices with respect to the trace norm $\langle A|B \rangle \equiv \text{tr}[A^\dagger B]$ and as such is a basis for the space of $d \times d$ matrices. Any density matrix σ can then be represented as $\sigma = \frac{1}{d} \sum_{k,l=0}^{d-1} s_{k,l} W_{k,l}$. The coefficients of this Weyl representations are $s_{k,l} = \text{tr}[W_{k,l}^\dagger \sigma]$. Accordingly, $W_{\mu,\nu} \equiv W_{\mu_1,\nu_1} \otimes W_{\mu_2,\nu_2}$ can be used to represent a bipartite state ρ with coefficients $s_{\mu,\nu}$. A sufficient criterion for separability was derived⁴³ that is named “Weyl” or “Spin representation criterion” here: If $\sum_{\mu,\nu} |s_{\mu,\nu}| \leq 2$, where $|s_{\mu,\nu}|$ are the coefficients of the Weyl representation of the bipartite state ρ , then ρ is separable.

Symmetry analysis. Leveraging the rich symmetries of $\mathcal{M}_d$ is another crucial factor to detect both entanglement and separability with the presented methods. The generation of all elements of this symmetry group greatly improves the detection capabilities of available detectors. Given an unknown state, all symmetric states can be generated by application of the according transformation for all generated symmetries. This set is then analyzed with respect to the available criteria. If entanglement class is determined for one state, all states of the set of symmetric states are known to be of the same class as well, due to the entanglement class conserving property of the symmetries.

Results

We determine the share of separable, bound and free entangled states with high accuracy using the methods presented above. After summarizing arguments for the generation of states via random sampling, two exemplary analyses are presented. First, states of the subset $\mathcal{F}_A$ (Eq. (4)) are analyzed before the main results of the entanglement classification for the enclosure polytope $\mathcal{E}_3$ (Eq. (2)) for the $d = 3$ are presented. The analysis of $\mathcal{F}_A$ serves as validation and comparison of applied methods, because here the borders between separable, bound entangled and free entangled states are analytically known^{25,32}. Since we know that all states outside of the enclosure polytope $\mathcal{E}_d$ are free entangled, the analysis of $\mathcal{E}_3$ suffices to know the entanglement classification of the whole simplex $\mathcal{M}_3$. Entanglement and separability criteria are compared for their effectiveness and their relations are discussed. The states are classified according to the labels “SEP” (separable states), “BOUND” (bound/PPT entangled states), “FREE” (free/NPT entangled states) and “PPT-UNKNOWN” (PPT states that could not be shown to be separable or bound entangled). For that purpose, more than 16,000 entanglement witnesses have been generated numerically. Additionally, the kernel of separable states has been extended by new separable states leveraging the generated symmetries to generate new vertex states for the convex hull. The symmetries have also been used for all methods to further increase the detection capabilities (see “Symmetry analysis” section above).

Generation of states. The generation of uniformly distributed random states offers significant advantages compared to states on a fixed lattice for estimating the relative volumes of entanglement classes in subsets of $\mathcal{M}_d$.

First, the relative frequencies of states in a certain class provides an unbiased estimator for the relative volume of that class if the states are randomly sampled but not if they are generated on a lattice with fixed increment. As shown in Fig. 2 (right), the relative frequencies of randomly generated states converge quickly with increasing sample size, while the convergence of lattice states is slower for $d = 3$ so that a significant bias remains for a number of states of the same magnitude.

Second, the expected relative frequencies of random states only depend on the relative volume of the classes and not on their specific geometric shape. A fixed lattice, however, may not distribute its states equally between the classes (according to their volume) as can be observed in Fig. 2 (left). For $d = 2$, the relative frequencies depend strongly on the increment of the lattice and show different convergence rates and directions for even and

Sample set	#PPT states
1	4869
2	5021
3	4901
4	5019
5	4982
6	4996
7	5014
8	5022
9	4969
10	5079

Table 1. Number of PPT states in sample sets of size 10,000. Table shows ten sample sets containing 10,000 states each and the number of PPT states within. The empirical mean is 4987.2 with empirical standard deviation of 61.7. These results are in agreement with the expected mean of 5000 and standard deviation of 50 according to the binomial distribution. For the exemplary analysis below we use a sample size of 10,000 states. The determined empirical values imply that the obtained numbers for the size of relative volumes have an statistical error in the magnitude of $10^{-3} = 0.1\%$.

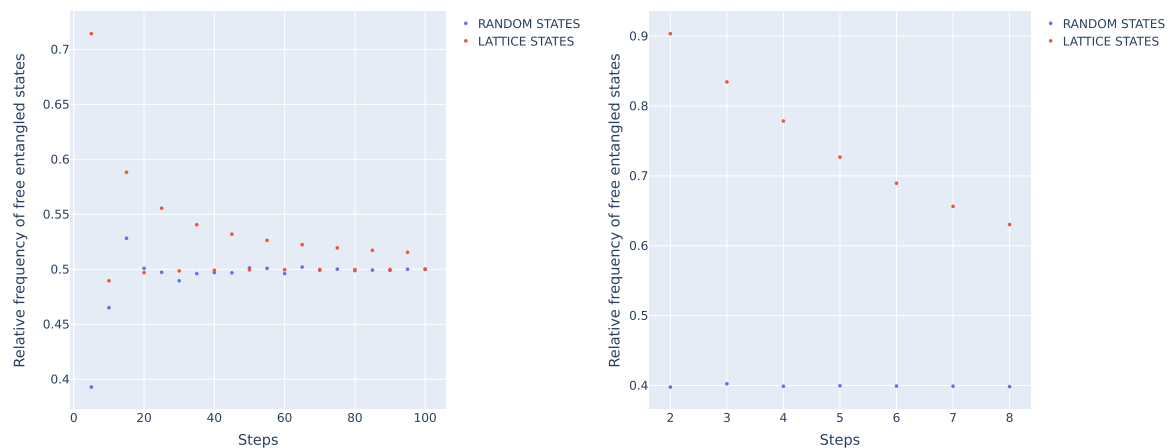


Figure 2. Relative frequency of free entangled states generated on a lattice (red) and by random sampling (blue) in dependence of the number of steps for $\mathcal{M}_2$ for $d = 2$ (left) and $\mathcal{E}_3$ for $d = 3$ (right).

odd number of lattice points in each dimension of $\mathcal{M}_2$. Since the geometric shape of entanglement classes in $\mathcal{M}_d$ is generally unknown for $d \geq 3$, it is not possible to make quantitative statements about the estimation and its error. Here, the uniform generation of random states is advantageous as well, since the probability of finding a random state to be in a given entanglement class is equal to its relative volume. The expected number of states in that class and the standard deviation thereof is then given by the binomial distribution, which was demonstrated for $d = 2$ in Table 1. Even if the exact size of the classes (and therefore the distribution) is not known *a priori*, the obtained estimations can be used to approximate the expected deviation of the true volume and the estimation.

Finally, the validity of this method allows the extension for higher dimensions. While already for $d = 3$ it is not possible to generate enough states based on a lattice to estimate the relative volumes correctly, the problem will be even larger for $d > 3$ due to the exponential growth of the Hilbert space. For randomly generated states, however, the estimation and its variance depend only on the size of the classes and not on the dimension. This implies that the sample size does not need to be increased for an estimation of the same quality in higher dimensions.

Entanglement classification for $\mathcal{F}_A$. To analyze Family A, the parameters α, β, γ of Eq. (4) are uniformly sampled within the range $[-1, 1]$ and the corresponding states are checked for positivity. This way, 10,000 uniformly distributed states of $\mathcal{F}_A$ are generated and analyzed.

Figure 3 shows that more than 99.5% of the states could be classified. The majority of states in $\mathcal{F}_A$ are free entangled (81.7%), 17.7% are separable and only 0.1% of all states are found to be bound entangled. For the remaining 0.5%, the states are known to have a positive partial transpose, but none of the criteria detected separability or bound entanglement.

Entanglement Class	# states	# relative to $\mathcal{F}_A$
Total $\mathcal{F}_A$	10000	100%
SEP	1774	17.7%
BOUND	7	0.1%
FREE	8171	81.7%
PPT-UNKNOWN	48	0.5%

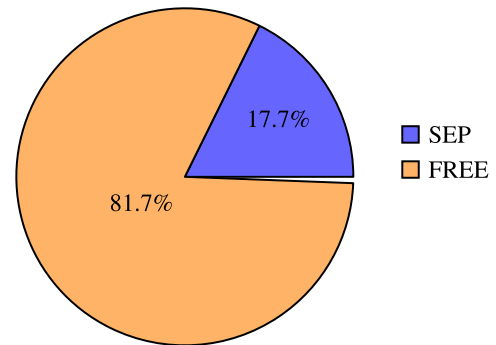


Figure 3. Entanglement classes and their relative volumes in $\mathcal{F}_A$.

Entanglement class	Criterion	#Detected	#Relative to class (%)
SEP	S1	1774	100
SEP	S2	457	25.8
BOUND	E2	0	0
BOUND	E3	7	100
BOUND	E4	0	0
BOUND	E5	0	0

Table 2. BOUND and SEP detectors in $\mathcal{F}_A$ and their relative number in class.

Entanglement Class	# states	# relative to $\mathcal{E}_3$
Total $\mathcal{E}_3$	10000	100%
SEP	4858	48.6%
BOUND	835	8.4%
FREE	3999	40.0%
PPT-UNKNOWN	308	3.1%

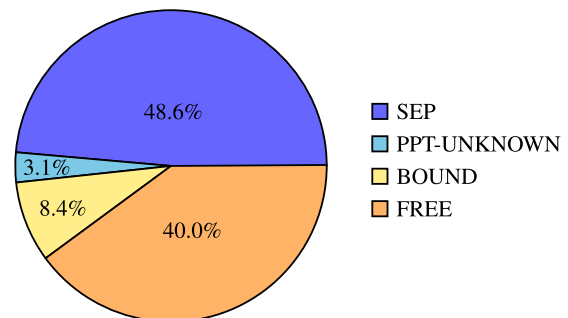


Figure 4. Entanglement classes and their relative volumes in $\mathcal{E}_3$.

As mentioned above, the criterion $E3$ was shown to be optimal for this family. This is also reflected in our results in Table 2, showing that only $E3$ detects bound entangled states. It can therefore be concluded, that the PPT-UNKNOWN states should be separable as well. Concerning separable states, both criteria $S1$ and $S2$ detect a significant number of states in SEP. However $S1$ is clearly stronger than $S2$, because all states detected by $S2$ are also detected by $S1$.

Entanglement classification for $\mathcal{E}_3$. The main results of this contribution is the determination of the entanglement class for arbitrary mixtures of Bell diagonal states of which $\mathcal{F}_A$ is a small subset. We restrict the analysis to the enclosure polytope $\mathcal{E}_3$ since all other states of $\mathcal{M}_3$ are known to be free entangled according to the PPT criterion.

For this analysis, 10,000 uniformly distributed random states are generated and analyzed. The results are summarized in Fig. 4, showing the number of states and the relative volume for each class. Almost half of the states (48.6%) are shown to be separable states via the separability criteria. 48.4% are entangled, of which 40.0% are found to be free entangled according to the PPT criterion as already expected with the analysis of Fig. 2. 8.4% of the states are bound entangled and only 3.1% of the states are known to be PPT-states, but none of the criteria allowed the certain detection of entanglement or separability, so the entanglement class remains unknown. To confirm the classification, we applied the detection criteria for the other classes also to already classified states, which all failed.

Entanglement class	Criterion	#Detected	#Relative to class (%)
SEP	S1	4858	100
SEP	S2	76	1.7
BOUND	E2	625	74.9
BOUND	E3	160	19.1
BOUND	E4	113	13.5
BOUND	E5	724	86.6

Table 3. BOUND and SEP detectors and their relative number in class.

Criterion (A)	#Detected (A)	Criterion (B)	#Detected (B)	#Detected (A) and (B)
E2	625	E3	160	107
E2	625	E4	113	113
E2	625	E5	724	545
E3	160	E4	113	19
E3	160	E5	724	120
E4	113	E5	724	113

Table 4. Pairwise comparison by criterion of detected bound states.

These results extend the previous related investigation²⁹ in several ways. First, the presented methods allow a more exact estimation of the relative volumes of the entanglement classes in $\mathcal{M}_3$. As argued above, the random generation of states provides an unbiased estimator for the relative volumes. In the cited analysis, the states were generated on a lattice with constant increment, but as shown in Fig. 2 the increment is not small enough to ignore the systematic bias due to the fixed lattice. In the cited analysis, higher values for the relative volumes of FREE and BOUND entangled states are given, while less states are classified as SEP. Figure 2 indicates that the lattice construction generates a higher density of states in areas in which FREE and BOUND entangled states dominate. Consequently, the generated states are not uniformly distributed, but, as can be easily seen, have a higher density for larger distances to the center of the kernel $\mathcal{K}_3$. This introduces a shift in the relative frequencies of the entanglement classes of generated states, since on average, separable states are closer to the center than entangled states. Second, using the new method S1 and S2 allows the direct detection of separable states beyond the kernel polytope $\mathcal{K}_3$. Together with the creation and consideration of all available symmetries it is possible to significantly reduce the number of PPT-UNKNOWN states.

The main challenge in the classification is of course the differentiation of separable and bound entangled states. For this reason we are strongly interested in the detection capabilities of different criteria for these classes, which are summarized in Table 3.

Although the analytical separability criterion S2 detects some states, the share (1.7%) is very small and the criterion S1, leveraging the geometric representation of the system and its symmetries, is clearly stronger. It detects all states for which separability is implied by S2 and many additional ones. Concerning the detection of bound entangled states, two criteria are especially powerful: First the analytical criterion E2, which detects 74.9% of all determined bound entangled states and second E5, using the set numerically generated EWs to detect 86.6% of the bound states. The later especially leverages the symmetry analysis to detect additional states. E3 and E4, although less effective than the other criteria, still detect a significant portion (19.1% and 13.5%) of PPT entangled states. Note also that E3 is the criterion that detects all bound entangled states for the family $\mathcal{F}_A$.

The question arises, whether some of the applied criteria to detect bound entangled states are stronger than others for E_3 . A criterion (A) is said to be stronger than criterion (B), if (A) detects all states detected by (B), as well. Table 4 relates a pair of applied entanglement criteria by comparing the number of exclusively detected bound entangled states to the number of bound entangled states that were detected by both criteria. The results are visualized in Fig. 5.

The set of bound entangled states detected by the two most effective criteria, E2 and E5, have a significant intersection, but each criterion also detects states that the other does not. 53 of 160 states (33%) that are detected by E3 are not found by E2, showing that E3 is not really weaker than E2, but seems to detect states of different areas of the geometrical polytope $\mathcal{E}_3$. In this sense, the analytical criteria E2 and E3 complement each other to a certain degree, as they together detect 81% of all detected bound entanglement. A similar, but slightly weaker conclusion can be made for the relation of E3 and E5, where 25% of the states found by E3 are not found by E5. This is in contrast to the results of E4. Here, both E2 and E5 detect all of the states found to be entangled according to E4, too. However, only 17% of the states detected by E4 are also found by E3, showing again a structural difference in the detection capabilities of the criteria E3 and E2. Also note that more than two entanglement criteria are needed to detect all bound entangled states.

In summary, the analytical criterion E2 and the numerically generated E5 are the most effective detectors. Their detected states have a significant intersection, but none is stronger than the other. In principle, all bound

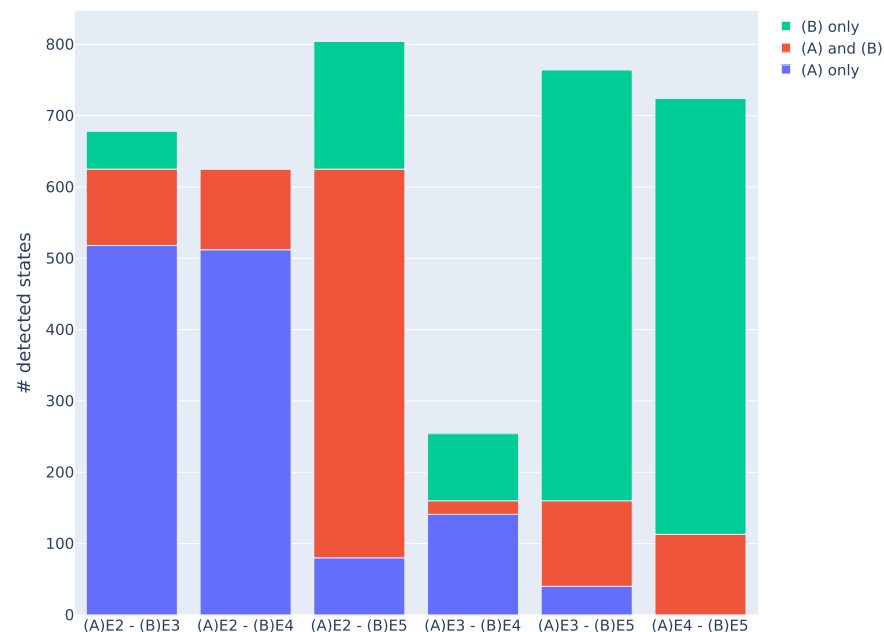


Figure 5. Pairwise comparison of number of exclusively (blue and green) and jointly (red) detected states.

entangled states can be detected by E5 with sufficiently many generated EWs. Thus, the states detected by E2 and not by E5 correspond to states in the space between the convex set of separable states and the enclosing hyper-planes of generated linear EWs. E3 detects less states than E2, but has a high share of exclusively detected states. E4 can be considered as weaker than E2 and has a very small overlap with E3, which further supports the hypothesis that E3 has a structural difference compared to the other criteria.

Discussion

We investigated mixtures of maximally entangled, bipartite Bell states in three dimensions. The states are locally maximally mixed, so all information is in the correlation of the subsystems, while no information about the individual systems is available. Depending on the mixing probabilities, the mixed state can be separable or entangled. Curiously, also bound entangled states, i.e. states that cannot be distilled by LOCC, occur, motivating a detailed analysis of this system to study this exotic form of entanglement.

The investigated system allows a geometric representation in which its states and their properties can be effectively geometrically analyzed. Leveraging these properties, the NP-hard problem of determining whether a state is separable, bound or free entangled can be almost completely solved for Bell diagonal qutrits with new efficient methods. Three main aspects contribute to this result: First, an efficient creation of states to estimate the relative volumes of entanglement classes. Second, the numerical creation of symmetries and their utilization for entanglement classification. Third, the development of independent separability and entanglement detection criteria, which can be applied efficiently to the to be classified states. A new collection of computational methods has been implemented to realize these aspects. It has been shown that the random sampling of states can be used to estimate the relative volumes of entanglement classes without systematic bias and that is has significant advantages compared to using states on a fixed lattice²⁹. Further utilizing the generators of the considered symmetries explicitly, all elements of the related group could be generated and applied to improve applied criteria to analyze the entanglement structure. The rapid growth of the number of distinct symmetry transformations with the dimension suggests that our methods can also be applied to higher dimensions despite the exponential growth of the Hilbert space.

Leveraging an efficient parameterization of separable states and the generation of symmetries allowed to use a sufficient criteria to directly detect separable states as well as bound entangled states using numerically generated entanglement witnesses. Several other well known entanglement criteria such as the PPT or realignment criterion have been implemented and applied to a representative sample of the system, allowing to compare their effectiveness.

Two relevant subsets of bipartite states in $\mathcal{M}_d$ for $d = 3$ were analyzed. First family $\mathcal{F}_A$, which contains mixtures of three Bell states and the maximally mixed state and whose symmetries allow to show that all bound entangled states can be detected by an analytical witness (E3)^{25,32}. More than 99% of the states could be classified including the detection of bound entanglement, but only 0.1% are found to be of that class. The numerical implementation of E3 confirms its effectiveness, since all bound entangled states are detected by this criterion. Interestingly, none of the other criteria detect any bound entanglement for this family. The second analysis and

main result of this work is the entanglement classification of the enclosure polytope $\mathcal{E}_3$, which is known to contain all PPT states. Again 10,000 uniformly distributed states were generated as representative sample of the system of which 96.9% could be classified. Only for 3.1% it remains unknown whether the state is separable or bound entangled. Almost half (48.6%) are found to be certainly separable, 40.0% are free entangled and can therefore be distilled by LOCC and at least 8.4% are bound entangled. The developed numerical methods S1 and E5 leveraging the special properties of the investigated system are the most effective criteria to distinguish separable and bound entangled states. However, also general analytical criteria detect significant shares of the bound entangled states, especially the “realignment criterion” E2 is very effective as well. All methods used to detect bound entanglement were investigated in pairwise comparison. Although not being the most effective method in terms of total number of detected states, the nonlinear criteria E3 (“quasi-pure concurrence”), detects a significant amount of states that are not detected by other criteria, especially the linear witness E4 and the collection of linear witnesses E5. This indicates that the states detected by E3 and not by E4 and E5 relate to states close to the nonlinear surface of separable states. It is interesting that E3 is such a strong detector of bound entanglement in $\mathcal{F}_A$, while other criteria like E2, which is a strong detector in $\mathcal{E}_3$, fail for this specific subset.

The developed methods and constructed related objects (i.e. generated entanglement witnesses and extended kernel) can be repeatedly applied to new states of interest in an efficient way. In this sense, the NP-hard problem of entanglement classification is efficiently solved for this system with an accuracy of approximately 94.9%, because any Bell diagonal PPT state can be efficiently classified with high probability. Only for 5.1% of the analyzed PPT states we fail to classify it as separable or bound entangled. Moreover, we can deduce from our studies that least 13.9% of PPT Bell diagonal states are bound entangled while 81.0% are separable. The implemented framework can equivalently be applied to any dimension d . It will be interesting to see the results in higher dimensions $d = 4$ and $d = 5$ and what structural similarities and differences in the classification of entanglement can be identified.

Data availability

All analyzed datasets were generated during the current study and are available from the corresponding author on reasonable request.

Received: 19 May 2022; Accepted: 6 July 2022

Published online: 21 July 2022

References

- Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
- Cozzolino, D., Da Lio, B., Bacco, D. & Oxenlowe, L. K. High-dimensional quantum communication: Benefits, progress, and future challenges. *Adv. Quantum Technol.* **2**, 1900038. <https://doi.org/10.1002/qute.201900038> (2019).
- Wang, Y., Hu, Z., Sanders, B. C. & Kais, S. Qudits and high-dimensional quantum computing. *Front. Phys.* **8**, 479. <https://doi.org/10.3389/fphy.2020.589504> (2020).
- Moskal, P. et al. Time resolution of the plastic scintillator strips with matrix photomultiplier readout for j-PET tomograph. *Phys. Med. Biol.* **61**, 2025–2047. <https://doi.org/10.1088/0031-9155/61/5/2025> (2016).
- Moskal, P. & Stepien, E. Prospects and clinical perspectives of total-body pet imaging using plastic scintillators. *PET Clin.* **15**, 439. <https://doi.org/10.1016/j.cpet.2020.06.009> (2020).
- Hiesmayr, B. C. & Moskal, P. Genuine multipartite entanglement in the 3-photon decay of positronium. *Sci. Rep.* **7**, 15349. <https://doi.org/10.1038/s41598-017-15356-y> (2017).
- Hiesmayr, B. C. & Moskal, P. Witnessing entanglement in compton scattering processes via mutually unbiased bases. *Sci. Rep.* **9**, 8166. <https://doi.org/10.1038/s41598-019-44570-z> (2019).
- Horodecki, M., Horodecki, P. & Horodecki, R. Mixed-state entanglement and distillation: Is there a bound entanglement in nature? *Phys. Rev. Lett.* **80**, 5239–5242. <https://doi.org/10.1103/PhysRevLett.80.5239> (1998).
- Bej, P. & Halder, S. Unextendible product bases, bound entangled states, and the range criterion. *Phys. Lett. A* **386**, 126992 (2021).
- Lockhart, J., Gühne, O. & Severini, S. Entanglement properties of quantum grid states. *Phys. Rev. A* **97**, 062340. <https://doi.org/10.1103/PhysRevA.97.062340> (2018).
- Bruß, D. & Peres, A. Construction of quantum states with bound entanglement. *Phys. Rev. A* **61**, 030301. <https://doi.org/10.1103/PhysRevA.61.030301> (2000).
- Slater, P. B. Jagged islands of bound entanglement and witness-parameterized probabilities. Preprint at <http://arxiv.org/abs/Quantum> (2019).
- Choi, M.-D. Some assorted inequalities for positive linear maps on C^* -algebras. *J. Oper. Theory* **4**, 271 (1980).
- Chruściński, D. & Sarbicki, G. Entanglement witnesses: Construction, analysis and classification. *J. Phys. A Math. Theor.* **47**, 483001. <https://doi.org/10.1088/1751-8113/47/48/483001> (2014).
- Kalev, A. & Bae, J. Optimal approximate transpose map via quantum designs and its applications to entanglement detection. *Phys. Rev. A* **87**, 062314. <https://doi.org/10.1103/PhysRevA.87.062314> (2013).
- Bae, J. Designing quantum information processing via structural physical approximation. *Rep. Prog. Phys.* **80**, 104001. <https://doi.org/10.1088/1361-6633/aa7d45> (2017).
- Korbicz, J. K., Almeida, M. L., Bae, J., Lewenstein, M. & Acín, A. Structural approximations to positive maps and entanglement-breaking channels. *Phys. Rev. A* **78**, 062105. <https://doi.org/10.1103/PhysRevA.78.062105> (2008).
- Huber, M., Mintert, F., Gabriel, A. & Hiesmayr, B. C. Detection of high-dimensional genuine multipartite entanglement of mixed states. *Phys. Rev. Lett.* **104**, 210501. <https://doi.org/10.1103/PhysRevLett.104.210501> (2010).
- Augusiak, R., Bae, J., Tura Brugués, J. & Lewenstein, M. Checking the optimality of entanglement witnesses: An application to structural physical approximations. *J. Phys. A Math. Theor.* **47**, 065301. <https://doi.org/10.1088/1751-8113/47/6/065301> (2014).
- Hiesmayr, B. C. & Löffler, W. Complementarity reveals bound entanglement of two twisted photons. *New J. Phys.* **15**, 083036. <https://doi.org/10.1088/1367-2630/15/8/083036> (2013).
- Peres, A. Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413–1415. <https://doi.org/10.1103/PhysRevLett.77.1413> (1996).
- Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: Necessary and sufficient conditions. *Phys. Lett. A* **223**, 1–8 (1996).
- Gurvits, L. Classical deterministic complexity of edmonds’ problem and quantum entanglement. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, STOC ’03*, 10–19 (Association for Computing Machinery, 2003). <https://doi.org/10.1145/780542.780545>.

24. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. A special simplex in the state space for entangled qudits. *J. Phys. A Math. Theor.* **40**, 7919. <https://doi.org/10.1088/1751-8113/40/28/S03> (2007).
25. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. State space for two qutrits has a phase space structure in its core. *Phys. Rev. A* **74**, 032327. <https://doi.org/10.1103/PhysRevA.74.032327> (2006).
26. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. The geometry of bipartite qutrits including bound entanglement. *Phys. Lett. A* **372**, 2190–2195 (2008).
27. Życzkowski, K., Horodecki, P., Sanpera, A. & Lewenstein, M. Volume of the set of separable states. *Phys. Rev. A* **58**, 883–892. <https://doi.org/10.1103/PhysRevA.58.883> (1998).
28. Życzkowski, K. Volume of the set of separable states. II. *Phys. Rev. A* **60**, 3496–3507. <https://doi.org/10.1103/PhysRevA.60.3496> (1999).
29. Hiesmayr, B. C. Free versus bound entanglement, a np-hard problem tackled by machine learning. *Sci. Rep.* **11**, 19739. <https://doi.org/10.1038/s41598-021-98523-6> (2021).
30. Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899. <https://doi.org/10.1103/PhysRevLett.70.1895> (1993).
31. Hill, S. A. & Wootters, W. K. Entanglement of a pair of quantum bits. *Phys. Rev. Lett.* **78**, 5022–5025. <https://doi.org/10.1103/PhysRevLett.78.5022> (1997).
32. Bae, J. *et al.* Detection and typicality of bound entangled states. *Phys. Rev. A* **80**, 022317. <https://doi.org/10.1103/PhysRevA.80.022317> (2009).
33. Chen, K. & Wu, L.-A. A matrix realignment method for recognizing entanglement. *Quantum Information and Computation* **3** (2002). <https://doi.org/10.48550/arXiv.quant-ph/0205017>.
34. Wootters, W. K. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248. <https://doi.org/10.1103/PhysRevLett.80.2245> (1998).
35. Spengler, C., Huber, M., Brierley, S., Adaktylos, T. & Hiesmayr, B. C. Entanglement detection via mutually unbiased bases. *Phys. Rev. A* **86**, 022311. <https://doi.org/10.1103/PhysRevA.86.022311> (2012).
36. Wootters, W. K. & Fields, B. D. Optimal state-determination by mutually unbiased measurements. *Ann. Phys.* **191**, 363–381 (1989).
37. Bandyopadhyay, S., Boykin, P. O., Roychowdhury, V. P. & Vatan, F. A new proof for the existence of mutually unbiased bases. *Algorithmica* **34**, 512–528. <https://doi.org/10.1007/s00453-002-0980-7> (2002).
38. Terhal, B. M. Bell inequalities and the separability criterion. *Phys. Lett. A* **271**, 319–326 (2000).
39. Bae, J., Chruściński, D. & Hiesmayr, B. C. Mirrored entanglement witnesses. *NPJ Quant. Inf.* **6**, 1–7. <https://doi.org/10.1038/s41534-020-0242-z> (2020).
40. Spengler, C., Huber, M. & Hiesmayr, B. C. A composite parameterization of unitary groups, density matrices and subspaces. *J. Phys. A Math. Theor.* **43**, 385306. <https://doi.org/10.1088/1751-8113/43/38/385306> (2010).
41. Mogensen, P. K. & Riseth, A. N. Optim: A mathematical optimization package for Julia. *J. Open Source Softw.* **3**, 615. <https://doi.org/10.21105/joss.00615> (2018).
42. Forets, M. & Schilling, C. Lazysets.jl: Scalable symbolic-numeric set computations*. *Proc. JuliaCon Conf.* **1**, 97 (2021).
43. Chruściński, D. & Pittenger, A. O. Generalized circulant densities and a sufficient condition for separability. *J. Phys. A Math. Theor.* **41**, 385301. <https://doi.org/10.1088/1751-8113/41/38/385301> (2008).

Acknowledgements

B.C. Hiesmayr acknowledges gratefully the Austrian Science Fund (FWF-P26783).

Author contributions

C.P. implemented the software, developed the new methods and analyzed the data. B.C.H. revised the analyses and results and proposed improvements. C.P. and B.C.H. both edited the manuscript.

Funding

The funding was provided by Austrian Science Fund (FWF-P26783).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16225-z>.

Correspondence and requests for materials should be addressed to C.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

B.3. Comparing bound entanglement of Bell-diagonal pairs of qutrits and ququarts



OPEN Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts

Christopher Popp[✉] & Beatrix C. Hiesmayr

We compare the classification as entangled or separable of Bell diagonal bipartite qudits with positive partial transposition (PPT) and their properties for different dimensions. For dimension $d \geq 3$, a form of entanglement exists that is hard to detect and called bound entanglement due to the fact that such entangled states cannot be used for entanglement distillation. Up to this date, no efficient solution is known to differentiate bound entangled from separable states. We address and compare this problem named separability problem for a family of bipartite Bell diagonal qudits with special algebraic and geometric structures and applications in quantum information processing tasks in different dimensions. Extending analytical and numerical methods and results for Bell diagonal qutrits ($d = 3$), we successfully classify more than 75% of representative Bell diagonal PPT states for $d = 4$. Via those representative states we are able to estimate the volumes of separable and bound entangled states among PPT ququarts ($d = 4$). We find that at least 75.7% of all PPT states are separable, 1.7% bound entangled and for 22.6% it remains unclear whether they are separable or bound entangled. Comparing the structure of bound entangled states and their detectors, we find considerable differences in the detection capabilities for different dimensions and relate those to differences of the Euclidean geometry for qutrits ($d = 3$) and ququarts ($d = 4$). Finally, using a detailed visual analysis of the set of separable and bound entangled Bell diagonal states in both dimensions, qualitative observations are made that allow to better distinguish bound entangled from separable states.

Quantum technology leverages quantum phenomena for better performance than classical methods for applications like computing, communication, simulation, metrology and cryptography^{1–5}. Quantum information theory provides the theoretical formalism for processing tasks using quantum mechanical systems⁶. One of the characteristic properties of a quantum system that allows realizing information processing with superior performance compared to classical systems is entanglement. Besides its relevance for our general understanding of nature and the interpretation of quantum theory^{7–9} it provides one of the main resources to realize applications in various fields ranging from quantum teleportation to medical applications for the detection of cancer cells^{10–13}. The simplest system to observe entanglement is the bipartite system of two two-level quantum systems, called qubits. Currently, most applications are based on these quantum systems of dimension $d = 2$, but recently, interest in higher dimensional systems like “qutrits” for $d = 3$, “ququarts” for $d = 4$ or “qudits” for general d is growing due to potential advantages and new observable phenomena^{14,15}. Bell states¹⁶ are special sets of entangled states which can be used as basis for the corresponding Hilbert space. They are highly relevant for applications due to the fact that they are maximally entangled states. Originally introduced for $d = 2$, they can be generalized for higher dimensions^{17–19}.

In this paper we analyze mixtures of maximally entangled bipartite Bell states, with focus on $d = 3$ and $d = 4$. Those states are locally maximally mixed, meaning that there is no correlation in the respective subsystems. Depending on the mixing probabilities of the d^2 pure Bell basis states, a general mixed state can be entangled or not, in which case it is separable. The Peres-Horodecki criterion, also known as PPT (positive partial transposition) criterion^{20,21}, provides an efficient method to detect a state as entangled, if the partially transposed density matrix of a given quantum state has at least one negative eigenvalue, in which case the state is called “NPT”. Otherwise it is called “PPT”. For $d = 2$, all entangled states are NPT, but for $d \geq 3$, also PPT entangled states exist²². While NPT entangled states can be “distilled”²² to result in fewer strongly entangled states, this process is not possible for PPT entangled ones. For this reason PPT entanglement is also called “bound” entanglement, which has been extensively investigated since its discovery by the Horodecki family, e.g. in Refs.^{23–33}. 2014, it has also been observed in experiment, using photons entangled in their orbital angular momentum³⁴. Many applications

Faculty of Physics, University of Vienna, Währingerstrasse 17, 1090 Vienna, Austria. ✉email: christopher.popp@univie.ac.at

like teleportation or superdense coding⁶ require strongly entangled states for reliable performance. However, if a given state is transformed to a bound entangled one, this resource is bound for immediate application, since it cannot be used to distill strongly entangled states. For this reason, it is important to know about structure of bound entanglement in a given system and to be able to detect those states reliably, so that operations that result in binding the resource entanglement in certain states can be avoided. However, the “separability problem” to differentiate separable and PPT entangled states has been proved to be NP-hard^{35,36} in general and lacks an efficient solution if the dimension of the system is not small. Existing methods to detect PPT entangled states^{27–33} are often strongly limited in the number of states they can detect and are not efficient in higher dimensions. Likewise, no efficient method to decide whether a PPT state is separable or bound entangled states is known for Bell diagonal qudits, which are known to be highly relevant for practical applications³⁷. However, special families of these states have strong symmetries that can be leveraged for the analytical and numerical analysis of its properties regarding the entanglement structure^{19,34,38}. In particular, the analytical structure of mixed Bell states generated by Weyl-Heisenberg transformations¹⁸ allows to derive several criteria to detect separability and entanglement^{19,38}. Furthermore, an efficient geometric representation of the states, symmetries and entanglement witnesses^{39,40} makes the system well applicable for numerical methods.

Recently, analytical and numerical methods were combined to solve the separability problem for the system in three dimensions in an “almost complete” way⁴¹. Given any unknown PPT state, the developed methods allow the classification of this state as separable or bound entangled with a probability of success of 95%. Moreover the classification allows the determination of the relative volumes of separable, PPT and NPT entangled states. It was further shown, that a significant share of the PPT states of bipartite, Bell diagonal qudits are bound entangled (at least 13.9%), making this system exceptionally well suited to study this exotic form of entanglement regarding its detection, use in information processing tasks and implications for nature. It is expected that the dimension of the system has a large influence on the structure and the relative shares of entanglement classes, which is a focus of this contribution. While approximations of the relative volumes of separable states in general systems in dependence on the dimension exist^{42,43}, the precise numbers depend on the specific system and are not known.

The aim of this work is to extend and apply those methods, used to successfully characterize the system for qudits⁴¹, for $d = 4$, to draw conclusions about the structure of entangled and separable states as well as the effectiveness of their detectors and to compare the results to $d = 2$ and $d = 3$. The paper is organized as follows: First, the system to be analyzed is defined and relevant methods to generate states and to investigate its entanglement structure are presented for general dimension. Second, we analyze the set of PPT states for $d = 4$. We quantify the share of this set in the total system and the relative volumes of separable and (bound) entangled states within and compare to other dimensions. Then, the applied criteria to detect separability and entanglement are compared for their effectiveness in different dimensions. Finally, we leverage the special properties of the system to visualize the set of separable and bound entangled states for $d = 3$ and $d = 4$. The visual analysis demonstrates relations between the algebraic structure of Bell diagonal mixtures and geometric restrictions on the set of separable and bound entangled states. This might be leveraged for the detection of bound entanglement and separability.

Methods

Consider the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ for the bipartite system of two qudits of dimension d . In this work we analyze mixtures of maximally entangled orthonormal Bell states $|\Omega_{k,l}\rangle \in \mathcal{H}$ with $k, l = 0, 1, \dots, (d-1)$ generated by applying the Weyl operators¹⁸ $W_{k,l}$ to one qudit of the shared maximally entangled state $|\Omega_{00}\rangle \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$:

$$|\Omega_{k,l}\rangle \equiv W_{k,l} \otimes \mathbb{1}_d |\Omega_{00}\rangle \quad (1)$$

where $W_{k,l} \equiv \sum_{j=0}^{d-1} w^{j \cdot k} |j\rangle \langle j+l \pmod{d}|$, $w = e^{\frac{2\pi i}{d}}$. Mixing the density matrices, or “Bell projectors”, $P_{kl} \equiv |\Omega_{k,l}\rangle \langle \Omega_{k,l}|$ with mixing probability $c_{k,l}$ defines Bell diagonal states with respect to the above defined Weyl operators and the system of interest for this work:

$$\mathcal{M}_d \equiv \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right\} \quad (2)$$

By taking the partial trace with respect to one of the subsystems, the reduced state of any state in $\mathcal{M}_d$ is maximally mixed, so all information is in the correlation of the combined state and not in the subsystems themselves. States with this property are called “locally maximally mixed”. Any state of $\mathcal{M}_d$ is equivalent to a point in d^2 dimensional Euclidean space by identifying the mixing probabilities $c_{k,l}$ with coordinates in real space. Due to the normalization of the $c_{k,l}$, the set of these points forms a standard simplex. Referring to the “magic Bell basis” of Wootters and Hill⁴⁴ $\mathcal{M}_d$ is also known as “magic simplex”^{19,38,45}.

The properties of the Weyl operators $W_{k,l}$ imply a linear ring structure or “discrete phase space” for operators indexed by the tuples (k, l) based on succeeding application of these operators¹⁹. This can be seen via the Weyl relations¹⁸ (addition defined modulo d):

$$W_{k_1,l_1} W_{k_2,l_2} = w^{l_1 k_2} W_{k_1+k_2,l_1+l_2} \quad (3)$$

$$W_{k,l}^\dagger = w^{kl} W_{-k,-l} = W_{k,l}^{-1} \quad (4)$$

In Figs. 1 and 2 we visualize this phase space as lattice of $d \times d$ vertices, each vertex corresponding to the Weyl operator (and thus also to a Bell state via eq.(1)) with according indices (k, l) . Depending on the dimension d ,

several subgroups exist and for this work subgroups containing d elements are of special relevancy, as they can be related to the structure of the sets of separable and bound entangled states. In general, a subgroup is defined by its generating elements. In case d is prime, all subgroups of d elements are generated by one of the Weyl operators. Highlighting the vertices (here red and blue) corresponding to a subgroup generated by $W_{k,l}$ induces “lines” in the discrete phase space (see Fig. 1). This is different for non-prime dimensions, where subgroups of d elements can additionally be generated by two Weyl operators whose indices contain proper divisors of d in which case “sublattices” are formed in the phase space (see e.g. Fig. 2). For more details, consult Ref.¹⁹. We can also use Figs. 1 and 2 to specify “subgroup states” by assigning each highlighted subgroup to a state which consists of equal mixtures of all corresponding Bell states.

Relevant subsets. The properties of $\mathcal{M}_d$ allow the definition of special subsets related to the entanglement properties of contained states, which have been investigated with respect to separability and (bound) entanglement^{45,47}.

Enclosure polytope. The enclosure polytope is a superset of all states with positive partial transposition. It was shown¹⁹ that all states that have at least one mixing probability $c_{k,l}$ exceeding $1/d$ are necessarily entangled and can be detected by the Peres-Horodecki criterion (PPT criterion). As NPT entangled states they can be distilled by local operation and classical communication (LOCC), and are therefore called “free” and not “bound” entangled²². The enclosure polytope is defined as:

$$\mathcal{E}_d \equiv \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \in [0, \frac{1}{d}] \right\} \quad (5)$$

Using the representation of $\mathcal{M}_d$ in Euclidean space, $\mathcal{E}_d$ forms a bounded polytope.

Kernel polytope. For each of the subgroups of d elements induced by the Weyl operators (indexed by α), a special “subgroup” or “sublattice state” ρ_α can be defined, which is known to be a separable state¹⁹. In general, these subgroup states are equal mixtures of d Bell states corresponding to a subgroup with probability $1/d$, e.g. the line state $\rho_{\alpha_1} = \frac{1}{d} \sum_{k=0}^{d-1} P_{k,0}$ (see also Figs. 1 and 2). This gives rise to a kernel polytope $\mathcal{K}_d$, which is defined as convex mixture of these separable line or sublattice states with d elements:

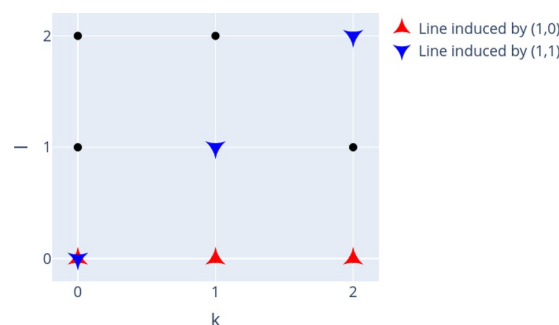


Figure 1. Phase space and exemplary induced subgroups called “lines” for $d = 3$. Figure created with Ref.⁴⁶.

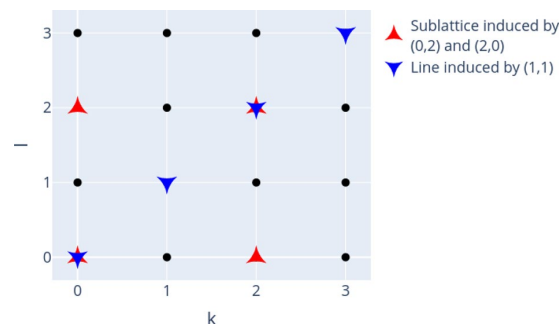


Figure 2. Phase space and exemplary induced subgroups called “lines” and “sublattices” for $d = 4$. Figure created with Ref.⁴⁶.

$$\mathcal{K}_d \equiv \{\rho = \sum_{\alpha} \lambda_{\alpha} \rho_{\alpha} \mid \lambda_{\alpha} \geq 0, \sum_{\alpha} \lambda_{\alpha} = 1\} \quad (6)$$

All states in the kernel polytope are convex combinations of separable states, consequently each state in $\mathcal{K}_d$ is by construction separable and the center corresponds to the maximally mixed state.

Bell diagonal state generation in arbitrary dimension. Using the geometric representation of $\mathcal{M}_d$ in Euclidean space, any state can be generated by specifying its d^2 coordinates $c_{k,l}$. Random sampling as well as deterministic procedures can be used to generate states. It was shown⁴¹ that via random sampling, uniformly distributed states in $\mathcal{M}_d$ can be generated and used to estimate the relative volumes of the entanglement classes of separable, bound entangled and free entangled states in $\mathcal{M}_3$. The same method can be used for $d > 3$ to generate states in $\mathcal{M}_d$ and $\mathcal{E}_d$ by drawing the first $d^2 - 1$ coordinates from a uniform distribution in the range $[0, 1]$ for $\mathcal{M}_d$ or $[0, 1/d]$ for $\mathcal{E}_d$ respectively. The remaining coordinate is then chosen according to the normalization condition. If this is not possible with a non-negative probability, the coordinates do not represent a normalized physical state and is therefore rejected. This form of rejection sampling becomes less effective with growing dimension d because the probability of a random state being rejected increases rapidly for $d \geq 5$. However, several methods exist to sample uniformly distributed points on the standard simplex and thus in $\mathcal{M}_d$ in any dimension (see Ref.⁴⁸ and the references therein), but these methods are not accessible for sampling of states located only in the polytopes $\mathcal{E}_d$ or $\mathcal{K}_d$.

Symmetries and their generation. The ring structure of the Weyl operators $W_{k,l}$ can be used to define linear symmetry transformations^{19,38} acting on states in $\mathcal{M}_d$. These transformations act as permutations on the Bell basis projectors $P_{k,l}$ or equivalently as permutations of the coordinates $c_{k,l}$ of a state in $\mathcal{M}_d$. These symmetry transformations form a group and are known to conserve both the PPT property and entanglement. Together these properties imply the conservation of the entanglement class⁴¹, meaning that the subsets of separable, bound entangled and free entangled states are mapped to themselves. For numerical implementations as well as for understanding their action on a given state, any linear symmetry s can be characterized by their action on the basis projectors $s : P_{k,l} \rightarrow P_{k',l'}$. All elements of this symmetry group can be generated by combined application of the following group generators (all mathematical operations on the indices are defined as(mod d)):

- Momentum inversion: $m : P_{k,l} \rightarrow P_{-k,l}$
- Quarter rotation: $r : P_{k,l} \rightarrow P_{k,-l}$
- Vertical shear: $v : P_{k,l} \rightarrow P_{k+l,l}$
- Translation: $t_{p,q} : P_{k,l} \rightarrow P_{k+p,l+q}$ for $p, q \in (0, \dots, d-1)$

Note that the translation of a Bell diagonal state can also be realized by applying a corresponding Weyl operator to the mixed state. This is not the case for the other generators, as their action on each Bell state that contributes to the mixture depends on that state. Due to the finite number of elements (k, l) in the phase space induced by the Weyl operators $W_{k,l}$, the number of distinct symmetries generated by the generators above is finite as well and can be generated numerically. The number of existing symmetries grows quickly with the dimension. For $d = 2$, 24 distinct symmetries can be generated, for $d = 3$, 432 and for $d = 4$ already 1536 symmetries of this group exist. Taking these symmetries into account is essential for the numerical methods to be effective in the d^2 dimensional space.

Criteria for the detection of entanglement and separability. The separability problem to decide whether a given mixed quantum state is separable or entangled has been shown to be NP-hard with respect to the dimension d as complexity measure^{35,36} and lacks an efficient general solution by polynomial in time algorithms, which is also the case for states in $\mathcal{M}_d$ for general dimension d . It is currently unclear, whether an efficient general solution exists for the separability problem in $\mathcal{M}_d$. For $d = 2$, however, all entangled states can be detected by the PPT criterion^{20,21} and recently, an almost complete solution was presented⁴¹ for $d = 3$, in the sense that any random unknown PPT state in $\mathcal{M}_3$ can be classified numerically with a probability of success of 95%. The methods used in this work can be equivalently used or extended to be applicable for $d > 3$. In the following we introduce those methods shortly. For more detailed information, the reader is referred to Ref.⁴¹ and the references therein.

E1: PPT criterion. The “Positive Partial Transpose (PPT)” or “Peres-Horodecki” criterion²⁰ detects entanglement for a bipartite state if it has at least one negative eigenvalue (in which case it is said to be “NPT”). For $d = 2$ it detects all entangled states, but for $d \geq 3$ it is only sufficient due to the existence of PPT- or bound entangled states. The partial transpose Γ acts on the basis states of a bipartite state as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^{\Gamma} \equiv |i\rangle\langle j| \otimes |l\rangle\langle k|$.

E2: Realignment criterion. The realignment operation R is defined as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)_R \equiv |i\rangle\langle k| \otimes |j\rangle\langle l|$. The realignment criterion⁴⁹ states that if the sum of singular values of the realigned state σ_R are larger than 1, then σ is entangled. Like the PPT criterion it is only sufficient for entanglement. Bound entangled states can be detected by this criterion, but it does not detect all NPT states in general.

E3: Quasi-pure concurrence criterion. The quasi-pure approximation⁴⁷ C_{qp} of the concurrence⁵⁰ allows the efficient detection of entanglement including its bound form. The approximation takes an explicit form for states

in $\mathcal{M}_d$: A state $\rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \in \mathcal{M}_d$ is entangled if $C_{qp}(\rho) = \max(0, S_{nm} - \sum_{(k,l) \neq (n,m)} S_{k,l}) > 0$ where the $S_{k,l}$ are explicitly given by⁴⁷

$$S_{k,l} = \sqrt{\frac{d}{2(d-1)} c_{k,l} \left[\left(1 - \frac{2}{d}\right) c_{n,m} \delta_{k,n} \delta_{l,m} + \frac{1}{d^2} c_{(2n-k) \bmod d, (2m-l) \bmod d} \right]} \quad (7)$$

and (n, m) is a multi-index of the coordinate of the largest value $\{c_{k,l}\}$.

E4: MUB criterion. A set of orthonormal bases $\{B_k\}$ and $B_k = \{|i_k\rangle \mid i = 0, \dots, (d-1)\}$ is called “mutually unbiased bases (MUB)” if $\forall k \neq l: |\langle i_k | j_l \rangle|^2 = \frac{1}{d} \quad \forall i, j = 0, \dots, (d-1)$. At most $d+1$ MUBs exist^{51,52}, in which case it was shown^{34,53,54} that the sum of “mutual predictabilities” obeys

$$I_{d+1}(\rho_s) = \sum_{k=1}^{d+1} C_k(\rho_s) \leq 2 \quad (8)$$

for all separable states (ρ_s) , when defining

$$C_1(\rho) = \sum_{i=0}^{d-1} \langle i_1 | \otimes \langle (i_1 + s)^* | \rho | i_1 \rangle \otimes | (i_1 + s)^* \rangle, \quad (9)$$

$$C_k(\rho) = \sum_{i=0}^{d-1} \langle i_k | \otimes \langle i_k^* | \rho | i_k \rangle \otimes | i_k^* \rangle, \quad k = 2, \dots, d+1. \quad (10)$$

Here, $s = 0, 1, \dots, (d-1)$ and i_k^* denotes the complex conjugate vector element. The MUB criterion thus indicates that if any state violates (8), it is entangled. If $s > 0$ and $s \neq d/2$ the MUB criterion allows the detection of PPT entangled states⁵³, which was also experimentally demonstrated for entangled photons³⁴ in the case $d = 3$. Note, that other, inequivalent MUBs exist, including extendible or unextendible sets of bases that contain less than $d+1$ elements⁵⁵. The set of entangled states that are detected by the MUB criterion generally depends on the used MUB. For this work, we use MUBs of $d+1$ elements as given in the Appendix A1 and set $s = 2$ for $d = 3$ and $s = 3$ for $d = 4$.

E5: Numerically generated entanglement witnesses. An entanglement witness³⁹ (“EW”) W is an observable which implies an upper bound U and also a lower bound⁴⁰ L ($U, L \in \mathbb{R}$), for separable states ρ_s :

$$L \leq \text{tr}[\rho_s W] \leq U \quad (11)$$

A state ρ is “detected by W ” to be entangled, if $\text{tr}[\rho W] \notin [L, U]$. For the system $\mathcal{M}_d$, EWs of the form $W = \sum_{k,l=0}^{d-1} \kappa_{k,l} P_{k,l}$ with $\kappa_{k,l} \in [-1, 1]$ can detect all entangled states¹⁹. In this case $\rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \in \mathcal{M}_d$ and $\text{tr}[\rho W] = \sum_{k,l=0}^{d-1} c_{k,l} \kappa_{k,l} \equiv c \cdot \kappa$ using the standard scalar product of the d^2 -dimensional vectors c and κ with coefficients $c_{k,l}$ and $\kappa_{k,l}$. Using the geometric representation of $\mathcal{M}_d$, an EW defines two $(d^2 - 1)$ -dimensional hyper-planes via $c_L \cdot \kappa = L$ and $c_U \cdot \kappa = U$ and induced halfspaces. Any point in the simplex but outside of the intersection of these halfspaces is entangled. A parameterization of unitaries⁵⁶ can be used to numerically determine the bounds for any EW defined by its coefficients $\kappa_{k,l}$ to create EWs for states in $\mathcal{M}_d$ numerically.

Leveraging the geometric characterization of $\mathcal{M}_d$, also sufficient criteria to detect separable states have been developed and used to analyze $\mathcal{M}_3$ ⁴¹. They can be applied for $d > 3$ as well and are shortly stated here:

S1: Extended kernel criterion. The convexity of the set of separable states can be used to check if an unknown state is contained in the separable hull of known separable states via linear programming. For this work, the implementation of Ref.⁵⁷, was used to check an unknown state for separability based on the convex hull of a given set of separable states. Using known separable states in $\mathcal{M}_d$ as vertices, they form a polytope which approximates the convex set of all separable states in $\mathcal{M}_d$. The effectiveness of this criterion depends on the quality of this approximation. More separable vertices for the polytope to improve the approximation increase the probability to detect new separable states, but on the other hand the complexity of the linear program also increases. It is therefore important to use vertices that are spatially uniformly distributed and as close to the surface of the set of separable states as possible. The sublattice states ρ_α of the kernel polytope $\mathcal{K}_d$ meet those requirements¹⁹ and by using the entanglement-class-conserving symmetries, more vertices to extend the separable kernel can be generated.

S2: Weyl/Spin representation criterion. Based on the Weyl relations¹⁸

$$W_{k_1, l_1} W_{k_2, l_2} = w^{l_1 k_2} W_{k_1 + k_2, l_1 + l_2} \quad (12)$$

$$W_{k,l}^\dagger = w^{kl} W_{-k, -l} = W_{k,l}^{-1} \quad (13)$$

one can see that the Weyl operators form an orthogonal basis in the space of $d \times d$ matrices with respect to the trace norm $\langle A, B \rangle \equiv \text{tr}[A^\dagger B]$. Representing a density matrix σ as $\sigma = \frac{1}{d} \sum_{k,l=0}^{d-1} s_{k,l} W_{k,l}$ defines the coefficients

of this “Weyl representation” as $s_{k,l} = \text{tr}[W_{k,l}^\dagger \sigma]$. For bipartite states, $W_{\mu,v} \equiv W_{\mu_1,v_1} \otimes W_{\mu_2,v_2}$ the coefficients are indexed as $s_{\mu,v}$. It was shown⁵⁸ that if $\sum_{\mu,v} |s_{\mu,v}| \leq 2$, then ρ is separable. This criterion for separability is named “Weyl” or “spin representation criterion”.

Symmetry classification. The group of entanglement class conserving symmetries (see “Symmetries and their generation”) provide further methods to determine the entanglement class of a given unknown state in $\mathcal{M}_d$. First, the set of all symmetric states, i.e. the orbit of the unknown state, is generated by application of the according transformation for all generated symmetries. Then, this set is analyzed with respect to the available criteria. If the entanglement class is determined for one of the symmetric states, then all symmetric states are certainly of the same class. This method can additionally be used to generate more states of a certain class for further investigations.

Results

Volume of PPT states. A first application leveraging the presented methods is to determine the relative volume of states with positive partial transposition in $\mathcal{M}_d$. It was shown^{42,43} that the volume of general separable and bound entangled quantum states decreases exponentially with the dimension of the system. Here, we determine the relative volumes for Bell diagonal states.

As described in section “Relevant subsets”, all states with positive partial transposition are necessarily located in the enclosure polytope $\mathcal{E}_d$ (5) when represented in Euclidean space. This property of $\mathcal{M}_d$ yields an upper bound of the relative share of PPT states in the simplex by comparing the total volume of $\mathcal{M}_d$ to the volume of $\mathcal{E}_d$. The enclosure polytope generally contains both PPT and NPT states, but the ratio depends on the dimension d . For $d = 2$, all states of $\mathcal{E}_2$ are known to be separable and thus PPT⁵⁹, no PPT/bound entangled states exist. For $d = 3$ it was numerically shown⁴¹ that approximately 60.0% of the states in $\mathcal{E}_3$ (39% of all states in $\mathcal{M}_3$) are PPT. In order to determine the relative volumes of PPT, we generate a large number of uniformly distributed states for dimensions $d = 2, \dots, 10$ and check if they are in the enclosure polytope and if they are PPT. The results are summarized in Fig. 3.

The analysis demonstrates that despite of the fact that the relative volume of the enclosure polytope grows with increasing dimension, the relative number of PPT states quickly decreases. For $d = 4$, 11.6% are PPT and the enclosure polytope $\mathcal{E}_4$ makes up 79.0% of $\mathcal{M}_4$. For $d = 5$, the relative PPT volume reduces to 7.3% and already for $d = 6$, less than 1% are PPT, although 97.1% of states in $\mathcal{M}_6$ are located in $\mathcal{E}_6$.

Entanglement classification of PPT states in $\mathcal{E}_d$. In order to compare the entanglement properties of mixed Bell diagonal states for the dimension $d = 2, 3, 4$, we use uniformly generated random states in $\mathcal{E}_d$ and classify them with the criteria for separability and entanglement presented above in order to estimate the share of each entanglement class. As mentioned in section “Volume of PPT states”, the classification of bipartite qubits can be completely characterized by the PPT criterion (E1): all states in $\mathcal{E}_2$ are PPT and separable, no PPT entangled states exist. For $d = 3$, we take the results of the previous investigation⁴¹, in which 96.1% of generated states in $\mathcal{E}_3$ have been successfully classified. To determine the relative volumes of entanglement classes among the PPT states for $d = 4$ with comparable precision, 40000 random states are generated in $\mathcal{E}_4$ out of which 96.7% can be successfully classified.

The 60% of states in $\mathcal{E}_3$ and 14.6% of $\mathcal{E}_4$ that have positive partial transposition are labeled as “SEP” or “BOUND” if they are detected as separable or entangled by the criteria E2–E5 or S1–S2. If none of the criteria allows classification, the state is labeled “PPT-UNKNOWN”. Table 1 summarizes the results:

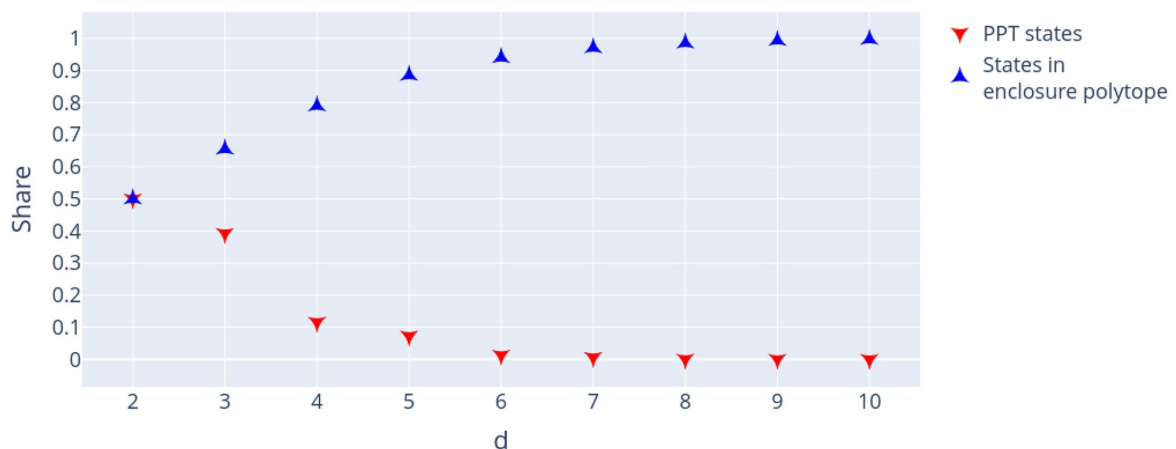


Figure 3. Relative volumes of the enclosure polytope $\mathcal{E}_d$ and PPT states in $\mathcal{M}_d$ for different dimensions d . Figure created with Ref.⁴⁶.

Entanglement class	Share of PPT for $d = 2$ (%)	Share of PPT for $d = 3$ (%)	Share of PPT for $d = 4$ (%)
SEP	100	81.0	75.7
BOUND	0	13.9	1.7
PPT-UNKNOWN	0	5.1	22.6

Table 1. Relative volumes of entanglement classes among the PPT states for $d = 2, 3, 4$.

Comparing the numerical classifications for $d = 3$ and $d = 4$, three noteworthy differences can be seen: First, the relative number of PPT-UNKNOWN states is significantly higher for $d = 4$ (22.6%) than for $d = 3$ (5.1%), although the success rates (96.7% and 96.1%) for the classification with respect to the total number of generated states in $\mathcal{C}_4$ and $\mathcal{C}_3$ are similar. Second, the share of separable states in $d = 4$ (75.7%) among the PPT states is quite high, in spite of the large number of yet to be classified PPT states. Third, the number of detected bound entangled states in $d = 4$ (1.7%) is significantly lower than for $d = 3$ (13.9%). Although it is possible that a large part of the PPT-UNKNOWN states are in fact BOUND and thus could potentially be detected by criterion E5, this shows that the detection capability of the analytical criteria (E2–E4) is more limited for $d = 4$ than for $d = 3$.

Detection capabilities and relations of applied criteria. The detection and thus differentiation between bound entangled and separable states is the core of the separability problem. Hence, the detection capabilities of the presented detectors for the classes SEP and BOUND are of special interest. Table 2 shows for each relevant criterion and dimension d the share of detected states among all SEP, respectively BOUND, classified states.

On the one hand, one notices that the strongest detectors for bound entangled Bell diagonal qutrits are also the most successful detectors in $d = 4$, namely E2 and E5. Relative to the total amount of detected bound entangled states, E2 seems to perform equally well in both dimensions. However, due to the large amount of PPT-UNKNOWN states, the relative detection power could be much worse. Still, E2 is clearly the strongest applied analytical criterion for $d = 3$ and $d = 4$. The criterion based on combining many numerically generated witnesses, E5, detects a large share of the identified BOUND states for both analyzed dimensions, however, the share is lower for $d = 4$ (86.6% for $d = 3$, 68.7% for $d = 4$). Considering the PPT-UNKNOWN states, the true detection capability of this criterion might be even below the determined share of 68.7% for $d = 4$, even though more EWs were used for $d = 4$ (approximately 22700 compared to 16700 for $d = 3$). This indicates that a single randomly generated numerical EW most likely is a weaker detector for the higher dimension.

On the other hand, the other detectors are clearly weaker in $d = 4$ compared to $d = 3$. The second strongest criterion in $d = 3$, E3, detects 19.1% of the BOUND qutrits in, while only 2.2% in $d = 4$. Again, the true detection capability is likely even below that due to the large number of unclassified PPT states. E4 detects a significant share (13.5%) of bound entangled qutrits while no PPT entangled qutrits for $d = 4$. Likewise, S2 detects no states as separable for $d = 4$.

These differences are also clearly reflected when comparing the criteria E2–E5 pairwise as shown in Figs. 4 and 5. The only criteria that have a significant number of jointly detected states in $d = 4$ are the detectors E2 and E5 (44.4% of combined detected states), although the share is smaller than for $d = 3$ (67.7%), confirming the reduced effectiveness of numerical EWs in the higher dimension. For $d = 4$, the other pairs are rather trivial, because of the very low number of detected states by E3 and E4. It should be noted, however that E3 detects one bound state that is neither detected by E2 nor by E5. Interestingly, this criterion also detects significant shares of bound entanglement that are not detected jointly by E2 or E5 in $d = 3$.

A final remark can be made related to the purity $\text{tr} \rho^2$ of detected bound entangled states ρ . For $d = 3$, the least pure states were detected by the criterion E3 but not E2 or E4. The few detected states for $d = 4$ do not allow to confirm this observation, although it can be noted that the least pure bound state is also uniquely detected by E3.

Visual analysis of separable and bound entangled states in $\mathcal{M}_3$ and $\mathcal{M}_4$. In this section we analyze and compare the structure of the set of separable and PPT entangled Bell diagonal states in dimension $d = 3$ and $d = 4$. Analyzing the geometric properties of those sets helps to improve geometric methods of entanglement detection or to provide new insights about the geometry of such quantum states in general (see e.g. Ref.⁶⁰). For

Entanglement class	Criterion	Share in class for $d = 3$ (%)	Share in class for $d = 4$ (%)
SEP	S1	100	100
SEP	S2	1.7	0
BOUND	E2	74.9	74.7
BOUND	E3	19.1	2.2
BOUND	E4	13.5	0
BOUND	E5	86.6	68.7

Table 2. BOUND and SEP detectors and their detection shares for $d = 3$ and $d = 4$.



Figure 4. Pairwise comparison of number of exclusively (blue and green) and jointly (red) detected states for $d = 3$. Figure created with Ref.⁴⁶.

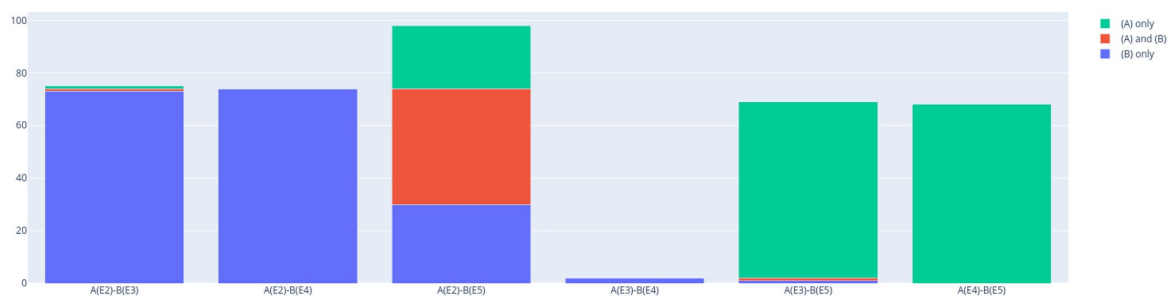


Figure 5. Pairwise comparison of number of exclusively (blue and green) and jointly (red) detected states $d = 4$. Figure created with Ref.⁴⁶.

$d = 2$, the geometric properties have been analyzed in detail before⁵⁹. Here, we visualize d dimensional projections of PPT states in $\mathcal{M}_d$ to demonstrate clear patterns that relate the geometric structure of the set of separable or bound entangled states to the algebraic structure of the Weyl operators and related Bells states. In particular, we make two qualitative observations:

- (Figs. 6, 7, 8 and 9):
The set of separable states in $\mathcal{M}_d$ is geometrically strongly restricted by the d element subgroup structure of the Weyl operators. More precisely, the more probability of a separable mixed state is concentrated on d Bell states, the closer it needs to be to a subgroup state (for definition see Figs. 1, 2 and related discussion).
- (Fig. 10):

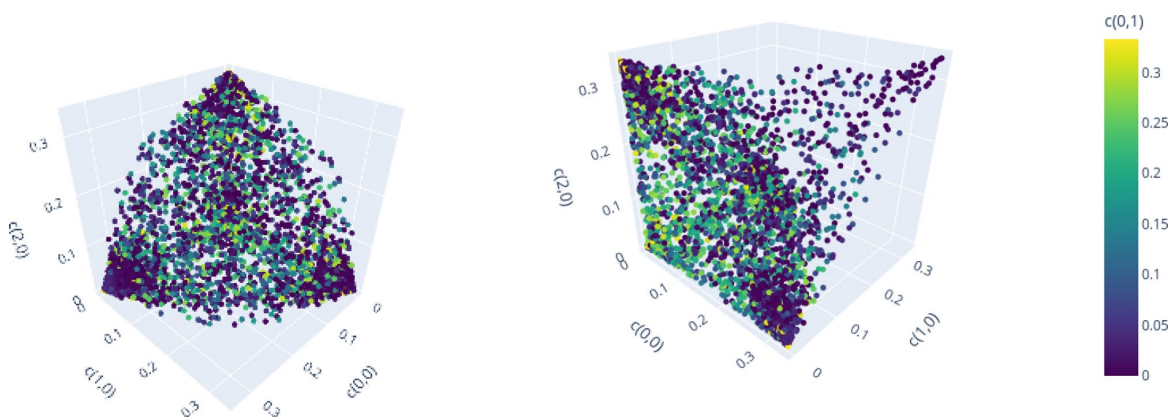


Figure 6. $(c_{0,0}, c_{1,0}, c_{2,0}, c_{0,1})$ for optimized separable states for $d = 3$, demonstrating strong correlations of on-line coordinates. Figure created with Ref.⁴⁶.

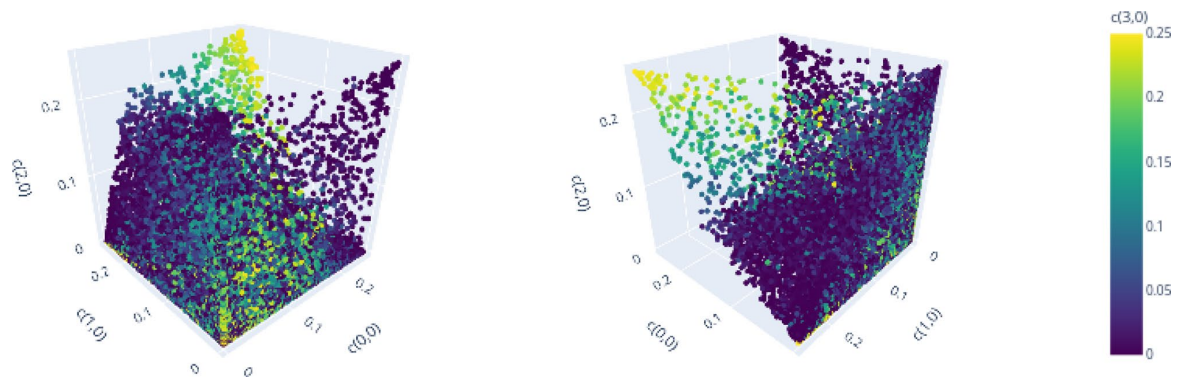


Figure 7. $(c_{0,0}, c_{1,0}, c_{2,0}, c_{3,0})$ for optimized separable states for $d = 4$, demonstrating strong correlations of on-line and additionally on-sublattice coordinates. Figure created with Ref.⁴⁶.

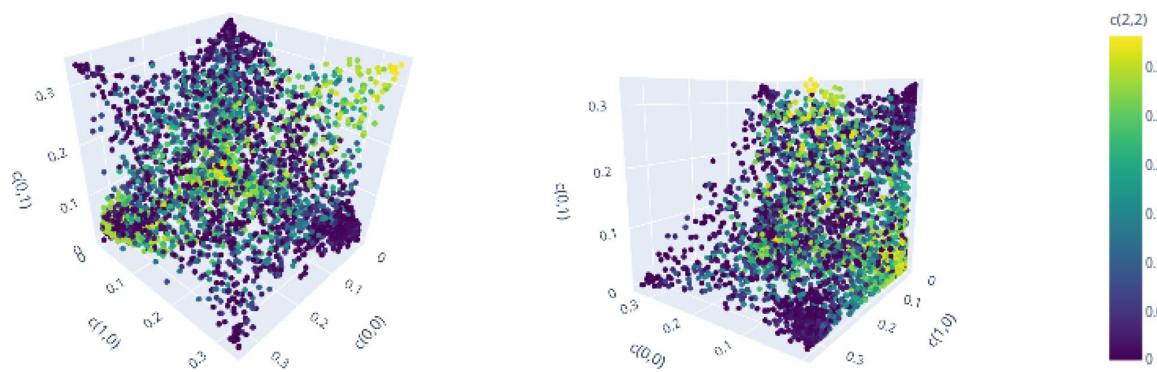


Figure 8. $(c_{0,0}, c_{1,0}, c_{0,1}, c_{2,2})$ for optimized separable states for $d = 3$, demonstrating necessity of full-line mixing for separability of states that concentrate most probability on 3 Bell states. Figure created with Ref.⁴⁶.

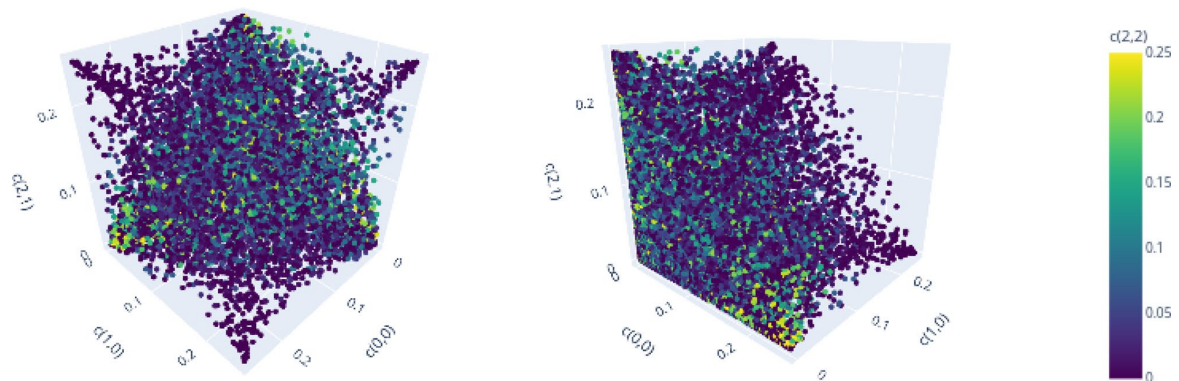


Figure 9. $(c_{0,0}, c_{1,0}, c_{2,1}, c_{2,2})$ for optimized separable states for $d = 4$, demonstrating necessity of full-line mixing for separability of states that concentrate most probability on 4 Bell states. Figure created with Ref.⁴⁶.

Similar restrictions related to this subgroup structure are present for bound entangled states, but certain restrictions are stronger for separable states than for bound entangled states. This difference can potentially be leveraged to detect bound entanglement.

The set of separable states in $\mathcal{M}_d$ forms a convex geometric body when represented as set of points in d^2 dimensional Euclidean space via their coordinates $c_{k,l}$. Naturally, not all coordinates of the d^2 dimensional space can be visualized at once, however, the symmetries (see section “Symmetries and their generation”) allow to capture

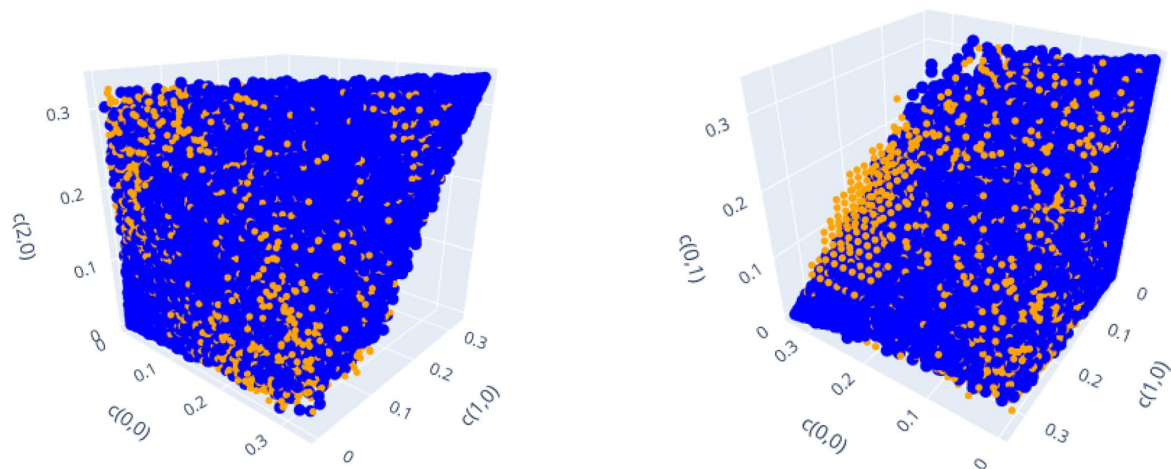


Figure 10. Comparison of projections to on-line (left) vs. off-line (right) coordinates for separable (blue) and bound entangled (orange) states in $d = 3$. Off-line projections show weaker geometric restrictions for bound entangled states in certain regions. Figure created with Ref.⁴⁶.

some essential geometric properties, even if only d coordinates are shown. The reason for this are d -element subgroups that are induced by the underlying ring structure of the Weyl operators (see Figs. 1 and 2). Subgroups can always be mapped onto each other with a corresponding symmetry transformation¹⁹. Since these transformations conserve the entanglement class and act as permutations on the coordinates $c_{k,l}$, different sets of d coordinates will show the same geometric properties if they are symmetric. Here, we consider the special subgroups represented by lines^{19,38} in the discrete phase space that are induced by the consecutive application of simple translations $t_{p,q}$. For example in $d = 3$, the indices (k, l) of the first three coordinates $(c_{0,0}, c_{1,0}, c_{2,0})$ form a line (red markers in Fig. 1), because they are related by the translation $t_{1,0}$. The geometric properties of these coordinates are then equivalent to e.g. those of $(c_{0,0}, c_{1,1}, c_{2,2})$ (blue markers in Fig. 1), as a suitable symmetry transformation relates the collections.

For visualizations we use separable states, which are specifically optimized to be close to the surface of the set of separable states, and random samples of PPT entangled states. We use three coordinates for a 3D-visualization and encode a fourth coordinate by color. Note that all PPT states are contained in the enclosure polytope, so it suffices to limit the range of the coordinates to $[0, 1/d]$.

To demonstrate the first observation, we compare the projections of separable states for $d = 3$ and $d = 4$ to d coordinates that relate either all to the same d -element subgroup or not. Figure 6 shows the geometric distribution of the first three coordinates on a line and a fourth coordinate encoded by the color for $d = 3$ from two point of views. A structure similar to a cone spanned by the corners of the enclosing polytope $\mathcal{E}_3$ $\{(1/3, 0, 0), (0, 1/3, 0), (0, 0, 1/3)\}$ and the subgroup state $\{(1/3, 1/3, 1/3)\}$ is visible, while no correlation with the off-line coordinate $c_{0,1}$ can be identified. Note that there are no separable states, for which two coordinates of the line, i.e., $c_{0,0}$ and $c_{1,0}$, are large, while the remaining line coordinate, i.e., $c_{2,0}$, is small.

Similar observations can be made for $d = 4$, when projecting to the on-line coordinates $(c_{0,0}, c_{1,0}, c_{2,0}, c_{3,0})$ in Fig. 7. Several correlations between the coordinates can be seen. Corresponding to the yellow cone, we see an accumulation of separable states that have high mixing probabilities for all of the four Bell states on the line. The blue cone pointing to the point $(1/4, 0, 1/4)$, on the other hand, relates to a different subgroup with indices $\{(0, 0), (2, 0), (0, 2), (2, 2)\}$ (red markers in Fig. 2). Finally, there is no symmetric cone and thus no separable states in the vicinity of $(0, 1/4, 1/4)$. Note that there is also no corresponding subgroup that contains the indices $(1, 0)$ and $(2, 0)$ but excludes the remaining line elements $(0, 0)$ and $(3, 0)$.

The visualizations above demonstrate that for each d -element subgroup, there is an accumulation of separable states in the vicinity of the corresponding subgroup state.

Consider now that d Bell states share most of the probability of a mixed state. We demonstrate that if those Bell states do not all correspond to the same subgroup, then the mixed state cannot be separable. First, note that if all probability is concentrated on less than d states, then at least one probability must exceed $1/d$, in which case the state cannot be PPT and therefore must be (NPT) entangled (see “Enclosure polytope” on p.4). Below, we visualize projections to coordinates, of which two relate to the same line and remaining coordinates are chosen to be part of different lines. Consider for $d = 3$ the projection to the line coordinates defined by $(c_{0,0}, c_{1,0})$ together with the off-line coordinates $c_{0,1}$ and $c_{0,2}$ (Fig. 8). No separable states are present for large values of both on-line coordinates $(c_{0,0}, c_{1,0})$ and large off-line values for $c_{0,1}$ or $c_{0,2}$.

The same observation can be made for $d = 4$. Choosing again two coordinates $(c_{0,0}, c_{1,0})$ to define a subgroup and two coordinates that are not part of it, one sees a similar structure in Fig. 9. Again, if the values of the line coordinates $c_{0,0}$ and $c_{1,0}$ are large, there is no separable state for large values of the off-line coordinates $c_{2,1}$ or $c_{2,2}$.

Due to the discussed symmetries, these characterizations hold for all coordinates that have a similar relation regarding their corresponding subgroups. We have therefore demonstrated the first observation: Considering mixed states that concentrate most of the probability on d Bell states, similarity to the subgroup states is necessary

for separability. Most mixed separable states have their probabilities distributed on all Bell states and are therefore centered around the maximally mixed state. In this case they are likely contained in the convex hull of the subgroup states that define the kernel polytope. We have therefore shown that this linear approximation also captures the relevant geometrical structure of the whole set of separable states in $\mathcal{M}_d$. In addition, the observed restrictions allow to construct more effective approximations, which focus on the areas, in which the surface of the set of separable states is curved (see e.g. Fig. 6 (left)) and the linear approximation fails. This can help to improve existing methods of detecting separability (e.g. the extended kernel criterion S1).

We conclude this section by demonstrating the second observation, which states that the geometric restrictions induced by the subgroup structure of the Weyl operators are also present for bound entangled states, but can be distinguished from those for separable states in certain cases.

Consider $d = 3$, for which a significant amount of bound entangled states can be classified. In Fig. 10, we again visualize d dimensional projections of separable states but also include the detected bound entangled states. Here, we show separable states in blue and bound entangled states in orange. The graphic on the left of Fig. 10 shows the projection to the coordinates $(c_{0,0}, c_{1,0}, c_{2,0})$ that relate to the same subgroup (line). One can see that the projections of the bound entangled states are restricted to the same area as the separable states (also compare Fig. 6 (right)). The described dependence of on-line projections on the related subgroups seems therefore to be a feature of all PPT states.

On the right hand, the visualization of the projection to three coordinates $(c_{0,0}, c_{1,0}, c_{0,1})$, that do not all belong to the same line, shows a relevant difference. The general form of visualized PPT states is still dominated by the convex combination of subgroup states (compare Fig. 8). There are no bound entangled states that concentrate most of the mixing probabilities on three Bell states that do not belong to the same subgroup. Crucially, however, there is a clearly visible region, in which only bound entangled states are present. For given on-line coordinate values $(c_{0,0}, c_{1,0})$, this region is characterized by higher off-line coordinate values $c_{0,1}$ than those accessible for separable projections.

For $d = 3$, this demonstrates the second observation: On the one hand, the subgroup structure imposes restrictions on all PPT states that determine the dominant shape of d dimensional projections of separable and bound entangled states. On the other hand, there exist d dimensional projections under which a subset of bound entangled states is mapped to regions, which do not contain any projected separable states. In principle, knowledge about these regions can be used to construct suitable entanglement witnesses of rank d to detect those PPT entangled states in d^2 dimensions.

For $d = 4$ the number of classified bound entangled states is not large enough to confirm similar characteristics, although none of the classified bound entangled states indicate a qualitative difference to the observations made for $d = 3$.

Discussion and conclusion

In this work we analyzed states of the bipartite system of mixed Bell states, which are related by Weyl transformations, with focus on subsystems with dimension 2, 3 and 4. The entanglement class of these locally maximally mixed states depends on the mixing probabilities and can be separable, NPT/free entangled or for $d > 2$ also PPT/bound entangled.

In order to investigate the properties for $d = 4$, i.e. bipartite ququarts, we extended the applied methods recently used to analyze the system of bipartite qutrits⁴¹. Leveraging a geometric representation, a random sampling of uniformly distributed states can be used to estimate the relative sizes of the entanglement classes via various criteria to detect separability and entanglement, including its bound version. Using this representation together with a group of entanglement class preserving symmetries, related analytical properties and an efficient parameterization⁵⁶ of states allows us to draw several conclusions about the entanglement properties of bipartite ququads and their dependence on the dimension for $d = 2, 3$ and 4.

A first observation is that the relative number of states with positive partial transposition, so either separable or bound entangled states, decreases very quickly with growing dimension, despite of increasing relative volume of the enclosure polytope $\mathcal{E}_d$, known to contain all PPT states. The share of PPT states in the full “magic simplex” $\mathcal{M}_d$ decreases from 50% for $d = 2$ to 39% for $d=3$ to 12% for $d = 4$. For $d > 5$, less than 1% of states are PPT.

The second observation is that significantly less bound entangled states can be detected, while the number of states that cannot be classified is considerably higher for $d = 4$ than for $d = 3$. To determine the relative volumes of entanglement classes for the enclosure polytope in $d = 4$ and compare them to $d = 3$, 40000 states in $\mathcal{E}_4$ have been created and classified with a probability of success of 96.7%. In principle, the probability of success could further be improved by extension of the numerical analyses and more states could easily be classified. In order to compare to $d = 3$, however, the extend of the numerical analysis and the number of states are chosen to result in a similar probability of success and number of PPT states. Limited to the set of PPT states, 77.4% of states could be successfully differentiated between separable and bound entangled states. The developed methods can be efficiently and repeatedly applied to new unknown states to solve the NP-hard “separability problem” with a probability of success of 77.4% for Bell diagonal ququarts in the magic simplex. Out of all PPT states in the system $\mathcal{M}_4$, at least 75.7% are determined to be separable and 1.7% are classified as bound entangled. The share of detected bound entangled states is clearly smaller than for $d = 3$ (13.9%), however, compared to the results of $d = 3$, a higher share (22.6 % vs 5.1%) of PPT states could not be classified and it remains unclear whether they are separable or bound entangled.

A third result can be stated regarding the detection capabilities of the applied criteria. The applied detectors for separability (S1, S2) or bound entanglement (E2–E5) are either based on deterministic, analytical conditions (S2, E2, E3, E4) or on a combined collection of numerically generated objects, i.e., vertices for the extension of the kernel polytope for S1 or EW-defining hyperplanes for E5. It can be seen from the large share of unclassified

PPT states, as well as from the relative shares of detected states by the criteria in each class, that both types of detectors are less powerful for classification in $d = 4$. Out of the analytical criteria for entanglement detection (E2–E4), for $d = 4$, only E2 detects a significant amount of bound entangled states, while E3 detects very few and E4 none at all, though it is known that E4 can detect bound entangled Bell diagonal states in $d = 4$ ^{53,55}. This is in strong contrast to $d = 3$, where the later two criteria allow the detection of 19.1% and 13.5% of all BOUND classified states. Interestingly, E3 still detects bound entanglement in $d = 4$ that cannot be detected with E2. This can also be observed for $d = 3$, where E3 can detect more strongly mixed entangled states than E2. The numerical criteria S1 and E5 also show reduced detection capability. Although S1 detects more than 75% of the PPT states as separable, the large number of unclassified states suggests that many separable states might not be detected by the used kernel extension. In addition to the share of classified states, the reduced number of states that are both detected by the analytical criterion E2 and the numerical E5 is also an indication of lower detection power of a single randomly generated EW for $d = 4$. Another striking difference between $d = 3$ and $d = 4$ is that S2 does not detect or is very unlikely to detect any separability for $d = 4$.

Many BOUND states that are detected by E2 are thus not confirmed by the criterion E5, which clearly shows that the number of generated EWs is not high enough, although more EWs were used than for $d = 3$. Two main reasons are likely responsible for the weaker performance in $d = 4$: First, the higher dimension of the Euclidean space and second, the different geometric properties of the set of separable states in $\mathcal{M}_d$ related to the properties of the Weyl operators and their induced phase space in non-prime dimensions. Both criteria represent approximations of this convex set: S1 represents an polytope approximation from within by identifying separable vertices close to the surface of separable states, while E5 represents an enclosing approximation with the hyperplanes defined by the upper and lower bounds of the EWs. The higher the dimension of the Euclidean space, the more objects (vertices/hyperplanes) are needed for a sufficient approximation and a generated set of objects may not be sufficient to achieve a comparable probability of success. The geometric properties of the (unknown) convex body formed by separable states are also relevant, as they determine the results of optimization procedures over the set of separable states in the whole Hilbert space, on which the generation of EWs and separable vertices rely.

Finally, we used classified separable and bound entangled states to enable visual analyses concerning the structure of PPT states in $\mathcal{M}_d$. We argued that relevant information can be extracted by considering projections to d coordinates, due to the special symmetries in $\mathcal{M}_d$ for $d = 3$ and $d = 4$. Two main qualitative observations were made that relate the structure of separable and bound entangled states to the algebraic subgroup structure of the Weyl operators.

On the one hand, it was shown that the states defined by the induced subgroup structure of the Weyl operators, which are used to define the kernel polytope $\mathcal{K}_d$, also determine the dominant geometric shape of the body of separable states. It was demonstrated for both $d = 3$ and $d = 4$ that a mixed state that concentrates most of the probability on d Bell states is separable, only if all Bell states relate to the same subgroup. This insights can be used to construct better approximations of the set of separable states and thus to improve methods to detect separability among PPT states.

On the other hand, we have demonstrated that similar restrictions related to the subgroups hold as well for the d dimensional projections of bound entangled states. Importantly, however, for some projections, these restrictions seem to be less strict for bound entangled states than those imposed on separable states. As a consequence, there are d dimensional areas, which are not reachable for projections of separable states. In principle, PPT states that are projected to these areas can be classified as entangled by geometric criteria that are defined in d dimensional space, instead of the full $(d^2 - 1)$ dimensional Hilbert space.

In conclusion, the methods for state generation and entanglement analysis applied to the system of $d = 3$ can also be successfully applied to $d = 4$ (with extensions), although with reduced effectiveness. Nonetheless, the presented methods solve the separability problem for $\mathcal{M}_4$ to a large extent, since any unknown state in $\mathcal{M}_4$ can efficiently be classified as separable or (bound) entangled with high probability of success. Significant differences in the relative volumes of the entanglement classes and in the detection capabilities of criteria for separability and entanglement are observed for $\mathcal{M}_3$ and $\mathcal{M}_4$. Relating the algebraic structure of the Weyl operators to the geometry of $\mathcal{M}_d$, qualitative observations could be made that characterize the structure of PPT states and propose new potential criteria to detect bound entanglement in $\mathcal{M}_d$. These contributions can serve as starting point to further improve the methods for classification and the general understanding of the entanglement structure of Bell diagonal qudits or general quantum states. On the one hand, further numerical investigations, extending the current results in terms of higher dimensions $d \geq 5$ or the structure of bound entangled quantum states, are possible. The applied sampling methods remain efficient for Bell diagonal states and thus allow the confirmation of results concerning the exponential decrease of volume for separable general quantum states with growing dimension as reported in Ref.^{42,43} or the observation of special properties for Bell diagonal states. Recently, a sequentially constrained Monte Carlo sampler (SCMCS) for quantum states was proposed⁶¹, which allows efficient sampling of quantum states subject to constraints like PPT or detection properties for specific entanglement criteria. This method could be used to generate general bound entangled quantum states and compare their properties to those of Bell diagonal states or for the detailed investigation of detection capabilities of certain entanglement witnesses. On the other hand, the reported structures of separable and bound entangled Bell diagonal states in $\mathcal{M}_d$ indicate properties that can be related to those of the Weyl operators and their induced phase space structure. The presented methods help to create, confirm or refute hypothesis about the structure of separable or entangled Bell diagonal states for different dimensions. Thus, they could provide a new accesses to the separability problem, the detection of bound entanglement or application relevant properties of quantum systems using entangled qudits.

Data availability

All analyzed datasets were generated during the current study and are available from the corresponding author on reasonable request. The software used to generate the reported results is published as registered open source package “BellDiagonalQudits.jl”⁶² available at <https://github.com/kungfugo/BellDiagonalQudits.jl>.

Received: 23 November 2022; Accepted: 31 January 2023

Published online: 04 February 2023

References

1. Cirac, J. I., Ekert, A. K., Huelga, S. F. & Macchiavello, C. Distributed quantum computation over noisy channels. *Phys. Rev. A* **59**, 4249–4254. <https://doi.org/10.1103/PhysRevA.59.4249> (1999).
2. Ekert, A. K. Quantum cryptography based on bell’s theorem. *Phys. Rev. Lett.* **67**, 661–663. <https://doi.org/10.1103/PhysRevLett.67.661> (1991).
3. Georgescu, I. M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86**, 153–185. <https://doi.org/10.1103/RevModPhys.86.153> (2014).
4. Deng, F.-G., Long, G. L. & Liu, X.-S. Two-step quantum direct communication protocol using the Einstein–Podolsky–Rosen pair block. *Phys. Rev. A* **68**, 042317. <https://doi.org/10.1103/PhysRevA.68.042317> (2003).
5. Sheng, Y.-B., Zhou, L. & Long, G.-L. One-step quantum secure direct communication. *Sci. Bull.* **67**, 367–374. <https://doi.org/10.1016/j.scib.2021.11.002> (2022).
6. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
7. Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete?. *Phys. Rev.* **47**, 777–780. <https://doi.org/10.1103/PhysRev.47.777> (1935).
8. Bell, J. S. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika* **1**, 195–200. <https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195> (1964).
9. Hensen, B. *et al.* Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
10. Moskal, P. *et al.* Time resolution of the plastic scintillator strips with matrix photomultiplier readout for j-PET tomograph. *Phys. Med. Biol.* **61**, 2025–2047. <https://doi.org/10.1088/0031-9155/61/5/2025> (2016).
11. Moskal, P. & Stepien, E. Prospects and clinical perspectives of total-body pet imaging using plastic scintillators. *PET Clin.* **15**, 439–452. <https://doi.org/10.1016/j.cpet.2020.06.009> (2020).
12. Hiesmayr, B. C. & Moskal, P. Genuine multipartite entanglement in the 3-photon decay of positronium. *Sci. Rep.* **7**, 15349. <https://doi.org/10.1038/s41598-017-15356-y> (2017).
13. Hiesmayr, B. C. & Moskal, P. Witnessing entanglement in compton scattering processes via mutually unbiased bases. *Sci. Rep.* **9**, 8166. <https://doi.org/10.1038/s41598-019-44570-z> (2019).
14. Cozzolino, D., Da Lio, B., Bacco, D. & Oxenlowe, L. K. High-dimensional quantum communication: Benefits, progress, and future challenges. *Adv. Quantum Technol.* **2**, 1900038. <https://doi.org/10.1002/qute.201900038> (2019).
15. Wang, Y., Hu, Z., Sanders, B. C. & Kais, S. Qudits and high-dimensional quantum computing. *Front. Phys.* **8**, 479 (2020).
16. Braunstein, S. L., Mann, A. & Revzen, M. Maximal violation of bell inequalities for mixed states. *Phys. Rev. Lett.* **68**, 3259–3261. <https://doi.org/10.1103/PhysRevLett.68.3259> (1992).
17. Sych, D. & Leuchs, G. A complete basis of generalized bell states. *New J. Phys.* **11**, 013006. <https://doi.org/10.1088/1367-2630/11/1/013006> (2009).
18. Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899. <https://doi.org/10.1103/PhysRevLett.70.1895> (1993).
19. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. A special simplex in the state space for entangled qudits. *J. Phys. A* **40**, 7919. <https://doi.org/10.1088/1751-8113/40/28/S03> (2007).
20. Peres, A. Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413–1415. <https://doi.org/10.1103/PhysRevLett.77.1413> (1996).
21. Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: Necessary and sufficient conditions. *Phys. Lett. A* **223**, 1–8 (1996).
22. Horodecki, M., Horodecki, P. & Horodecki, R. Mixed-state entanglement and distillation: Is there a bound entanglement in nature?. *Phys. Rev. Lett.* **80**, 5239–5242. <https://doi.org/10.1103/PhysRevLett.80.5239> (1998).
23. Bej, P. & Halder, S. Unextendible product bases, bound entangled states, and the range criterion. *Phys. Lett. A* **386**, 126992 (2021).
24. Lockhart, J., Gühne, O. & Severini, S. Entanglement properties of quantum grid states. *Phys. Rev. A* **97**, 062340. <https://doi.org/10.1103/PhysRevA.97.062340> (2018).
25. Bruß, D. & Peres, A. Construction of quantum states with bound entanglement. *Phys. Rev. A* **61**, 030301. <https://doi.org/10.1103/PhysRevA.61.030301> (2000).
26. Slater, P. B. Jagged islands of bound entanglement and witness-parameterized probabilities. *arXiv Quantum Physics* (2019). <https://doi.org/10.48550/arXiv.1905.09228>.
27. Choi, M.-D. Some assorted inequalities for positive linear maps on c^* -algebras. *J. Oper. Theory* **4**, 271–285 (1980).
28. Chruściński, D. & Sarbicki, G. Entanglement witnesses: Construction, analysis and classification. *J. Phys. A* **47**, 483001. <https://doi.org/10.1088/1751-8113/47/48/483001> (2014).
29. Kalev, A. & Bae, J. Optimal approximate transpose map via quantum designs and its applications to entanglement detection. *Phys. Rev. A* **87**, 062314. <https://doi.org/10.1103/PhysRevA.87.062314> (2013).
30. Bae, J. Designing quantum information processing via structural physical approximation. *Rep. Prog. Phys.* **80**, 104001. <https://doi.org/10.1088/1361-6633/aa7d45> (2017).
31. Korbicz, J. K., Almeida, M. L., Bae, J., Lewenstein, M. & Acín, A. Structural approximations to positive maps and entanglement-breaking channels. *Phys. Rev. A* **78**, 062105. <https://doi.org/10.1103/PhysRevA.78.062105> (2008).
32. Huber, M., Mintert, F., Gabriel, A. & Hiesmayr, B. C. Detection of high-dimensional genuine multipartite entanglement of mixed states. *Phys. Rev. Lett.* **104**, 210501. <https://doi.org/10.1103/PhysRevLett.104.210501> (2010).
33. Augusiak, R., Bae, J., Tura Brugués, J. & Lewenstein, M. Checking the optimality of entanglement witnesses: An application to structural physical approximations. *J. Phys. A: Math. Theor.* **47**, 065301 (2014).
34. Hiesmayr, B. C. & Löffler, W. Complementarity reveals bound entanglement of two twisted photons. *New J. Phys.* **15**, 083036. <https://doi.org/10.1088/1367-2630/15/8/083036> (2013).
35. Gurvits, L. Classical deterministic complexity of Edmonds’ problem and quantum entanglement. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, STOC ’03* 10–19 (Association for Computing Machinery, 2003). <https://doi.org/10.1145/780542.780545>.
36. Gharibian, S. Strong np-hardness of the quantum separability problem. *Quantum Inf. Comput.* **10**, 343–360 (2008).
37. Werner, R. F. All teleportation and dense coding schemes. *J. Phys. A* **34**, 7081–7094. <https://doi.org/10.1088/0305-4470/34/35/332> (2001).

38. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. State space for two qutrits has a phase space structure in its core. *Phys. Rev. A* **74**, 032327. <https://doi.org/10.1103/PhysRevA.74.032327> (2006).
39. Terhal, B. M. Bell inequalities and the separability criterion. *Phys. Lett. A* **271**, 319–326 (2000).
40. Bae, J., Chruściński, D. & Hiesmayr, B. C. Mirrored entanglement witnesses. *npj Quantum Inf.* **6**, 15. <https://doi.org/10.1038/s41534-020-0242-z> (2020).
41. Popp, C. & Hiesmayr, B. C. Almost complete solution for the np-hard separability problem of bell diagonal qutrits. *Sci. Rep.* **12**, 12472. <https://doi.org/10.1038/s41598-022-16225-z> (2022).
42. Życzkowski, K., Horodecki, P., Sanpera, A. & Lewenstein, M. Volume of the set of separable states. *Phys. Rev. A* **58**, 883–892. <https://doi.org/10.1103/PhysRevA.58.883> (1998).
43. Życzkowski, K. Volume of the set of separable states. II. *Phys. Rev. A* **60**, 3496–3507. <https://doi.org/10.1103/PhysRevA.60.3496> (1999).
44. Hill, S. A. & Wootters, W. K. Entanglement of a pair of quantum bits. *Phys. Rev. Lett.* **78**, 5022–5025. <https://doi.org/10.1103/PhysRevLett.78.5022> (1997).
45. Baumgartner, B., Hiesmayr, B. C. & Narnhofer, H. The geometry of bipartite qutrits including bound entanglement. *Phys. Lett. A* **372**, 2190–2195 (2008).
46. Plotly graphing libraries (v0.18.6). <https://plotly.com/julia/>.
47. Bae, J. *et al.* Detection and typicality of bound entangled states. *Phys. Rev. A* **80**, 022317. <https://doi.org/10.1103/PhysRevA.80.022317> (2009).
48. Willms, A. R. Uniform Sampling on the Standard Simplex. *Mo. J. Math. Sci.* **33**, 119–124 (2021).
49. Chen, K. & Wu, L.-A. A matrix realignment method for recognizing entanglement. *Quantum Inf. Comput.* **3**, 193–202. <https://doi.org/10.48550/arXiv.quant-ph/0205017> (2002).
50. Wootters, W. K. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248. <https://doi.org/10.1103/PhysRevLett.80.2245> (1998).
51. Wootters, W. K. & Fields, B. D. Optimal state-determination by mutually unbiased measurements. *Ann. Phys.* **191**, 363–381 (1989).
52. Bandyopadhyay, S., Boykin, P. O., Roychowdhury, V. P. & Vatan, F. A new proof for the existence of mutually unbiased bases. *Algorithmica* **34**, 512–528. <https://doi.org/10.1007/s00453-002-0980-7> (2002).
53. Bae, J., Bera, A., Chruściński, D., Hiesmayr, B. C. & McNulty, D. How many measurements are needed to detect bound entangled states? (2021).
54. Spengler, C., Huber, M., Brierley, S., Adaktylos, T. & Hiesmayr, B. C. Entanglement detection via mutually unbiased bases. *Phys. Rev. A* **86**, 022311. <https://doi.org/10.1103/PhysRevA.86.022311> (2012).
55. Hiesmayr, B. C. *et al.* Detecting entanglement can be more effective with inequivalent mutually unbiased bases. *New J. Phys.* **23**, 093018. <https://doi.org/10.1088/1367-2630/ac20ea> (2021).
56. Spengler, C., Huber, M. & Hiesmayr, B. C. A composite parameterization of unitary groups, density matrices and subspaces. *J. Phys. A: Math. Theor.* **43**, 385306. <https://doi.org/10.1088/1751-8113/43/38/385306> (2010).
57. Forets, M. & Schilling, C. Lazysets.jl: Scalable symbolic-numeric set computations*. *Proc. JuliaCon Conf.* **1**, 97 (2021).
58. Chruściński, D. & Pittenger, A. O. Generalized circulant densities and a sufficient condition for separability. *J. Phys. A: Math. Theor.* **41**, 385301. <https://doi.org/10.1088/1751-8113/41/38/385301> (2008).
59. Bertlmann, R. A., Narnhofer, H. & Thirring, W. Geometric picture of entanglement and bell inequalities. *Phys. Rev. A* **66**, 032319. <https://doi.org/10.1103/PhysRevA.66.032319> (2002).
60. Bengtsson, I. & Życzkowski, K. *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2006).
61. Li, W., Han, R., Shang, J., Ng, H. K. & Englert, B.-G. *Sequentially Constrained Monte Carlo Sampler for Quantum States*. [arXiv.org/abs/2109.14215](https://arxiv.org/abs/2109.14215) (2021).
62. Popp, C. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits. *J. Open Source Softw.* **8**(81), 4924. <https://doi.org/10.21105/joss.04924> (2023).

Acknowledgements

B.C.H. and C.P. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) project P36102-N. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

C.P. implemented the software, developed the new methods and analyzed the data. B.C.H. revised the analyses and results and proposed improvements. C.P. and B.C.H. both edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-29211-w>.

Correspondence and requests for materials should be addressed to C.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

B.4. Special features of the WeylHeisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states

PAPER • OPEN ACCESS

Special features of the Weyl–Heisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states

To cite this article: Christopher Popp and Beatrix C Hiesmayr 2024 *New J. Phys.* **26** 013039

View the [article online](#) for updates and enhancements.

You may also like

- [Distributing bipartite quantum systems under timing constraints](#)
Kfir Dolev, Alex May and Kianna Wan
- [An optimized encoding method for secure key distribution by swapping quantum entanglement and its extension](#)
Gan Gao and
- [Controlled quantum secure direct communication with twice identity authentications based on the entanglement swapping of Bell states](#)
Zheng Wang, Xia Li and Tian-Yu Ye



OPEN ACCESS

RECEIVED

22 September 2023

REVISED

21 December 2023

ACCEPTED FOR PUBLICATION

10 January 2024

PUBLISHED

22 January 2024

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



PAPER

Special features of the Weyl–Heisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states

Christopher Popp*  and Beatrix C Hiesmayr 

University of Vienna, Faculty of Physics, Währingerstrasse 17, 1090 Vienna, Austria

* Author to whom any correspondence should be addressed.

E-mail: christopher.popp@univie.ac.at**Keywords:** quantum information, entanglement, bound entanglement, Bell states, qudits, Weyl–Heisenberg, separability problem

Abstract

Bell states are of crucial importance for entanglement based methods in quantum information science. Typically, a standard construction of a complete orthonormal Bell-basis by Weyl–Heisenberg operators is considered. We show that the group structure of these operators has strong implication on error correction schemes and on the entanglement structure within Bell-diagonal states. In particular, it implies an equivalence between a Pauli channel and a twirl channel. Interestingly, other complete orthonormal Bell-bases do break the equivalence and lead to a completely different entanglement structure, for instance in the share of positive partial transposition (PPT)-entangled states. In detail, we find that the standard Bell basis has the highest observed share on PPT-states and PPT-entangled states compared to other Bell bases. In summary, our findings show that the standard Bell basis construction exploits a very special structure with strong implications to quantum information theoretic protocols if a deviation is considered.

1. Introduction

Leveraging quantum phenomena in technology offers new resources that enable methods with better performance than classically limited methods in the field of information theory and related applications like communication, computation, simulation, metrology or cryptography [1–5].

Entanglement is one of the main resources for novel methods for quantum information processing like super-dense coding [6] or teleportation [7], as well as in other fields like e.g. medical sciences [8–12]. Despite its relevance for our quantum-theoretic understanding of nature [13–15] and for practical applications in quantum technology, this phenomenon is far from being well understood. Two entangled qubits are the simplest system to observe entanglement. The qubit Bell states [16] are a set of maximally entangled states, which form a basis of this bipartite Hilbert space. A bipartite pure state is called ‘maximally entangled’, if the reduced state for each subsystem is the maximally mixed state, so all information about the state lies in the correlations between the subsystems. Most applications leveraging entanglement as a resource use these special states. The amount of entanglement two parties possess if they share a maximally entangled qubit Bell state is named ‘ebit’ [17]. Recently, multi-level quantum systems called ‘qudits’ (with the ‘ d ’ indicating the general dimension of the system) have been shown to offer potential advantages for applications [18, 19]. The notion of Bell states can be generalized to the qudit systems [20], in which case two parties holding a maximally entangled bipartite qudit state are said to possess one ‘edit’.

A frequently used standard construction for a basis of bipartite Bell states of arbitrary dimension is based on a set of operators called ‘Weyl–Heisenberg’ operators, which can be seen as generalization of the two-dimensional Pauli matrices to other dimensions [7]. Applying these operators locally to the standard maximally entangled state, a basis of Bell states can be generated [21]. Due to their properties, this specific ‘standard’ Bell basis is often used in applications, but as we will show in this contribution, other Bell bases exist, which differ in some relevant properties.

For real applications, decoherence due to interaction with the environment generally disturbs a pure state and the resource in form of an edit is destroyed or cannot be used effectively. For this reason, quantum error

correction [22, 23] and entanglement purification/distillation [24] are required. While error correction transforms a certain disturbed state back to some logical pure state, entanglement purification processes several weakly entangled or disturbed states to produce fewer but stronger or maximally entangled states. Both concepts have been generalized to qudit systems. The Weyl–Heisenberg operators thereby play a crucial role in the form of defining a ‘nice’ error basis [25–27] or to use the corresponding Bell states as target states [28–30]. One phenomenon, which does not appear for bipartite qubits but for dimension of the subsystems $d \geq 3$ is bound entanglement [31–36]. Bound entangled states are entangled states, but they cannot be used for entanglement purification. It is known that all entangled states with positive partial transposition (PPT) are bound entangled, while the existence of bound entangled states with negative partial transposition (NPT) is still an open problem. In general, the separability problem to decide whether a given PPT quantum state is entangled or separable is an NP-hard problem [37, 38].

Bipartite quantum states that are diagonal in the standard Bell basis are called ‘Bell-diagonal’. Bell-diagonal states arise naturally in noise environments affecting Bell states or in several applications like entanglement distillation or quantum key distribution [3, 39]. If they are constructed via the Weyl–Heisenberg operators, special algebraic and geometric properties of the states can be leveraged for some insight concerning this complex structure of entanglement [21, 36, 40]. Recently, the authors combined analytical and numerical methods to investigate the systems of bipartite qutrits ($d = 3$) and ququarts ($d = 4$) in detail [41–44]. As a result, 95%/77% of all PPT (standard) Bell-diagonal qutrits/ququarts can be classified as PPT-entangled or separable. Moreover, it was shown that the group structure in the set of the standard Weyl–Heisenberg-constructed Bell states is highly relevant for the entanglement structure.

In this work, we demonstrate properties of the standard Bell basis constructed via the Weyl–Heisenberg operators and show that this basis has special properties among a set of generalized Bell bases with significant implications for the corresponding systems of Bell-diagonal states. The paper is organized as follows: First, we define a set of operators called ‘Weyl-Twirl’ operators, for which all elements of the standard Bell basis are eigenstates and demonstrate an equivalence between the map to the set of Bell-diagonal states and the randomized application of those operators. Then, we present some applications of this ‘Weyl-Twirl’ with relevance for the separability problem of Bell-diagonal states and a simple error correction scheme for Bell states. Second, by generalizing the construction based on the Weyl–Heisenberg operators, we present a family of Bell bases, which contains the standard Bell basis, but are generally not unitarily equivalent. We show that, even though the generalized Bell bases also consist of maximally entangled orthonormal Bell states, several characteristics and properties of the standard Bell basis relevant for applications do not exist in the generalized case. These differences significantly affect the entanglement structure of Bell-diagonal states and the detection of PPT-entangled states. Finally, we conclude with a summary of our findings.

2. A stabilizing group for maximally entangled states

Here we define the Weyl–Heisenberg operators, introduce the channels and show their equivalency. Then we discuss applications in quantum information theory.

2.1. Weyl–Heisenberg and Weyl-Twirl operators

Let $\mathcal{H} = \mathcal{H}_d \otimes \mathcal{H}_d$ be the Hilbert space of two qudits of dimension d . In this bipartite system, we consider maximally entangled states, also called ‘Bell states’. An orthonormal basis of d^2 Bell states spanning the Hilbert space can be defined by applying certain local operators on one of the subsystems to a seed state, e.g. the maximally entangled state $|\Omega_{00}\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$. Applying the Weyl–Heisenberg operators [7]

$$W_{k,l} := \sum_{j=0}^{d-1} w^{jk} |j\rangle \langle j+l|, \quad k, l = 0, \dots, d-1 \quad (1)$$

with $w := e^{\frac{2\pi i}{d}}$ to the (w.l.o.g.) first subsystem of $|\Omega_{00}\rangle$ defines the ‘standard’ Bell basis of bipartite qudits:

$$|\Omega_{k,l}\rangle := W_{k,l} \otimes \mathbb{1}_d |\Omega_{00}\rangle, \quad k, l = 0, \dots, d-1. \quad (2)$$

In equation (1) and in the following, addition and subtraction are always to be understood mod d . The indices k and l in $W_{k,l}$ relate to phase and shift operations, respectively. To highlight this property, $W_{k,l}$ can also be written as $W_{k,l} = Z(k)X(l)$, where $Z(k) = \sum_j w^{jk} |j\rangle \langle j|$ is the phase operator and $X(l) = \sum_j |j-l\rangle \langle j|$ denotes the shift operator.

The Weyl–Heisenberg operators obey the following algebraic relations:

$$\begin{aligned} W_{k_1,l_1} W_{k_2,l_2} &= w^{l_1 k_2} W_{k_1+k_2,l_1+l_2} \\ W_{k,l}^\dagger &= w^{kl} W_{-k,-l} = W_{k,l}^{-1} \\ W_{k,l}^* &= W_{-k,l} \\ W_{k,l}^T &= w^{-kl} W_{k,-l}. \end{aligned} \quad (3)$$

Here, $(\dagger)$, $(*)$ and (T) denote the adjoint, complex conjugation and transposition with respect to the computational basis, respectively. These relations imply a linear structure for the Bell states, which was recently shown [42] to be highly relevant for the geometric properties of the set of separable, PPT entangled and NPT entangled states that are diagonal in the basis defined in equation (2).

Consider now the set of d^2 unitary operators, which we call ‘Weyl-Twirl operators’,

$$T_{i,j} := W_{i,j} \otimes W_{i,j}^*, \quad i, j = 0, \dots, d-1. \quad (4)$$

Applying the Weyl relations (3), one observes that this set of unitary operators $\{T_{i,j} | i, j = 0, \dots, d-1\}$ forms an Abelian group under multiplication with neutral element $T_{0,0} = \mathbb{1}_d \otimes \mathbb{1}_d$:

$$\begin{aligned} T_{i_1,j_1} T_{i_2,j_2} &= T_{i_1+i_2,j_1+j_2} = T_{i_2,j_2} T_{i_1,j_1} \\ T_{i,j}^{-1} &= T_{-i,-j}. \end{aligned} \quad (5)$$

Using the well-known property of the maximally entangled state $\mathbb{1}_d \otimes M |\Omega_{0,0}\rangle = M^T \otimes \mathbb{1}_d |\Omega_{0,0}\rangle$ for any matrix M together with the Weyl relations (3), we can calculate the action of $T_{i,j}$ on the Bell state $|\Omega_{k,l}\rangle$:

$$\begin{aligned} T_{i,j} |\Omega_{k,l}\rangle &= (W_{i,j} \otimes W_{i,j}^*) (W_{k,l} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle \\ &= W_{i,j} W_{k,l} \otimes W_{-i,j} |\Omega_{00}\rangle \\ &= W_{i,j} W_{k,l} W_{-i,j}^T \otimes \mathbb{1}_d |\Omega_{00}\rangle \\ &= w^{jk-il} W_{k,l} \otimes \mathbb{1}_d |\Omega_{00}\rangle \\ &= w^{jk-il} |\Omega_{k,l}\rangle. \end{aligned} \quad (6)$$

Each Bell state $|\Omega_{k,l}\rangle$ is thus an eigenvector for each $T_{i,j}$, with the phase of the eigenvalue being determined by the corresponding indices (k, l) and (i, j) .

2.2. Pauli projectors and the Weyl-Twirl channel

The properties of quantum channels, i.e. completely positive maps from one quantum state to another, lie at the heart of the dynamics of state propagation (see e.g. [17] for an overview). For instance, it is of interest to characterize channels which transform entangled states to separable states, so called entanglement-breaking channels (for a recent study see [45]). Here we consider bipartite ‘Pauli-type’ channels, i.e. channels whose Kraus operators are of the form $P_{k,l}$, that transform general states to Bell-diagonal ones.

Let us define the Bell projectors by $P_{k,l} := |\Omega_{k,l}\rangle \langle \Omega_{k,l}|$. The set $\mathcal{M}_d$ of Bell-diagonal states is given as mixtures of the projectors $\{P_{k,l}\}$ with mixing probabilities $\{c_{k,l}\}$:

$$\mathcal{M}_d := \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right\}. \quad (7)$$

This object forms a mathematical simplex and is also known as the magic simplex [21, 40, 46], where ‘magic’ refers to the ‘magic’ Bell basis for bipartite qubits introduced by Hill and Wootters [47]. Let

$\rho = \sum_{k,l=0}^{d-1} \sum_{m,n=0}^{d-1} \rho_{k,l,m,n} |\Omega_{k,l}\rangle \langle \Omega_{m,n}| \in \mathcal{H}$ be a bipartite state, represented in the Bell basis. We define the Pauli channel $\mathcal{P} : \mathcal{H} \rightarrow \mathcal{M}_d$ as a map from the total Hilbert space $\mathcal{H}$ to the set of Bell-diagonal states as follows:

$$\begin{aligned} \mathcal{P}(\rho) &:= \sum_{k,l=0}^{d-1} P_{k,l} \rho P_{k,l} = \sum_{k,l=0}^{d-1} \langle \Omega_{k,l} | \rho | \Omega_{k,l} \rangle P_{k,l} \\ &= \sum_{k,l=0}^{d-1} \rho_{k,l,k,l} P_{k,l} := \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l}. \end{aligned} \quad (8)$$

Using the operators $T_{i,j}$ (4), we can define another channel, which we name ‘Weyl-Twirl’ channel:

$$\mathcal{T}(\rho) := \frac{1}{d^2} \sum_{i,j=0}^{d-1} T_{i,j} \rho T_{i,j}^\dagger = \frac{1}{d^2} \sum_{i,j=0}^{d-1} W_{i,j} \otimes W_{i,j}^* \rho (W_{i,j} \otimes W_{i,j}^*)^\dagger. \quad (9)$$

We will show now that these two channels are equivalent, i.e.

Theorem 1.

$$\mathcal{P}(\rho) \equiv \mathcal{T}(\rho) \quad \forall \rho \in \mathcal{H}. \quad (10)$$

Proof. Consider the action of $\mathcal{T}$ on a basis state $|\Omega_{k,l}\rangle\langle\Omega_{m,n}|$:

$$\begin{aligned} \mathcal{T}(|\Omega_{k,l}\rangle\langle\Omega_{m,n}|) &= \frac{1}{d^2} \sum_{i,j} w^{j(k-m)-i(l-n)} |\Omega_{k,l}\rangle\langle\Omega_{m,n}| \\ &= \delta_{k,m} \delta_{l,n} |\Omega_{k,l}\rangle\langle\Omega_{m,n}|. \end{aligned}$$

Here, the first equality follows from (6) and the second equality from the identity $\sum_{j=0}^{d-1} w^{jx} = d\delta_{x,0}$. Let $\rho = \sum_{k,l,m,n=0}^{d-1} \rho_{k,l,m,n} |\Omega_{k,l}\rangle\langle\Omega_{m,n}| \in \mathcal{H}$. The equation above implies that only diagonal elements remain, i.e.:

$$\mathcal{T}(\rho) = \sum_{k,l=0}^{d-1} \rho_{k,l,k,l} |\Omega_{k,l}\rangle\langle\Omega_{k,l}| = \sum_{k,l=0}^{d-1} \rho_{k,l,k,l} P_{k,l} = \mathcal{P}(\rho).$$

□

Calling the application of the channel $\mathcal{T}$ ‘Weyl-Twirl’ is motivated by the fact that equation (9) represents the random application of (bi-local) operators $W_{i,j} \otimes W_{i,j}^*$. Operations of the form $\rho \rightarrow \int (U \otimes U^*) \rho (U \otimes U^*)^\dagger dU$ with U being a local unitary operator and dU being the according Haar measure are generally named ‘Twirl’. These operations leave certain diagonal elements invariant, while eliminating off-diagonal elements. Operations of this kind transform a state to Bell-diagonal form and are often relevant for certain applications, e.g. entanglement purification schemes [24]. The channel equivalence of the Pauli channel (equation (8)) and the finite ‘Weyl-twirl’ (equation (9)) shows that the operators $T_{i,j}$ have the special property that they leave all Bell-diagonal elements invariant under random application, while eliminating any off diagonal elements.

2.3. Applications of Weyl-Twirl operators and channel

The properties of the Weyl-Twirl operators $T_{i,j}$ and the channel equivalence (equation (10)) imply several properties that are relevant for applications related to maximally entangled states. In the following, we will briefly mention two examples.

2.3.1. The separability problem

A bipartite quantum state $\rho \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is separable ($\rho \in SEP$), if it can be written as convex combination of pure product states, that is $\rho = \sum_i q_i \rho_1^i \otimes \rho_2^i$ with $\rho_{1/2}^i \in \mathcal{H}_{1/2}$, $q_i > 0$ and $\sum_i q_i = 1$. Due to the existence of PPT-entangled states, the decision problem, whether a given bipartite state ρ is separable ($\rho \in SEP$) or entangled ($\rho \notin SEP$), is generally a NP-hard problem [37, 38] called ‘separability problem’. While all bipartite states with NPT are known to be entangled according to the Peres–Horodecki/PPT-criterion [48], for subsystems of dimension $d \geq 3$, also entangled states with non-NPT (PPT) exist [31]. No general and efficient criterion is known that detects all PPT-entangled states nor a general method to construct those. Consequently, given a bipartite PPT state, it is generally unknown whether the state is entangled or not. However, for the set of Bell-diagonal states $\mathcal{M}_d$ (7) in dimensions $d = 3$ and $d = 4$ the problem has efficiently been solved. In detail, for $d = 3$ (4), 95% (77%) of all PPT states have been classified as entangled or separable, using a combination of analytical criteria and numerical methods [41, 42]. Thus, for those dimension and the standard Bell basis the structure of PPT-entangled states is known.

As demonstrated in referenced works, the combination of two numerical methods are especially effective to distinguish separable and PPT-entangled Bell-diagonal states:

1. The construction of separable states $\rho_s \in \mathcal{M}_d \cap SEP$, close to the border of the convex set of separable states, providing an inner approximating polytope of $\mathcal{M}_d \cap SEP$.

2. The construction of optimal entanglement witnesses (see definition below), providing an outer approximation of $\mathcal{M}_d \cap \text{SEP}$.

Both methods rely on an efficient parameterization of separable Bell-diagonal states $\rho_s \in \mathcal{M}_d \cap \text{SEP}$ that can be used to optimize corresponding target quantities. Given an efficient parameterization of unitaries and pure states [49, 50], a parameterization of separable and Bell-diagonal states follows from the following corollaries of theorem 1:

Corollary 1. Any separable state ρ_s remains separable under application of the Pauli channel $\mathcal{P}$, that is:

$$\rho_s \in \text{SEP} \Rightarrow \mathcal{P}(\rho_s) \in \mathcal{M}_d \cap \text{SEP}.$$

Proof. Due to linearity, it suffices to show the corollary for a pure product state $\rho_s = |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2|$. By theorem 1, we have:

$$\begin{aligned} \mathcal{P}(\rho_s) &= \mathcal{T}(\rho_s) \\ &= \frac{1}{d^2} \sum_{i,j=0}^{d-1} W_{i,j} \otimes W_{i,j}^* \rho_s \left(W_{i,j} \otimes W_{i,j}^* \right)^\dagger \\ &= \frac{1}{d^2} \sum_{i,j=0}^{d-1} W_{i,j} |\psi_1\rangle\langle\psi_1| W_{i,j}^\dagger \otimes W_{i,j}^* |\psi_2\rangle\langle\psi_2| (W_{i,j}^*)^\dagger. \end{aligned}$$

$W_{i,j}$ is unitary, so $W_{i,j} |\psi_1\rangle\langle\psi_1| W_{i,j}^\dagger$ and $W_{i,j}^* |\psi_2\rangle\langle\psi_2| (W_{i,j}^*)^\dagger$ are pure states. Therefore, the above expression represents an equal mixture of pure product states with mixing probability $\frac{1}{d^2}$ and thus is a separable mixed state. By definition, we also have $\mathcal{P}(\rho_s) \in \mathcal{M}_d$. $\square$

The second corollary states that optimal and Bell-diagonal entanglement witnesses are also optimal for $\mathcal{M}_d$. An entanglement witness K is a Hermitian operator, for which the expectation $\text{Tr}(K\rho_s)$ is non-negative for all separable states ρ_s , while at least one state ρ_e exists with $\text{Tr}(K\rho_e) < 0$. In this case K is said to detect the entangled state ρ_e . K is said to be optimal, if there exists a separable state ρ_0 , such that $\text{Tr}(K\rho_0) = 0$. It has been shown [40] that Bell-diagonal entanglement witnesses can detect all entangled states in $\mathcal{M}_d$.

Corollary 2. If a Bell-diagonal entanglement witness K is optimal, then it is also optimal for $\mathcal{M}_d$, i.e. $\exists \rho_0 \in \mathcal{M}_d \cap \text{SEP}$ s.th. $\text{Tr}(K\rho_0) = 0$.

Proof. Let $\rho_0 = \sum_{k,l,m,n=0}^{d-1} \rho_{0(k,l,m,n)} |\Omega_{k,l}\rangle\langle\Omega_{m,n}| \in \text{SEP}$ be a state so that $\text{Tr}(K\rho_0) = 0$. Define $\tilde{\rho}_0 := \mathcal{P}(\rho_0)$. By corollary 1, $\tilde{\rho}_0 \in \mathcal{M}_d \cap \text{SEP}$. K is of Bell-diagonal form and Hermitian, so $K = \sum_{k,l=0}^{d-1} \kappa_{k,l} P_{k,l}$, $\kappa_{k,l} \in \mathbb{R}$. This implies:

$$0 = \text{Tr}(K\rho_0) = \sum_{k,l=0}^{d-1} \kappa_{k,l} \rho_{0(k,l,k,l)} = \text{Tr}(K\tilde{\rho}_0).$$

$\square$

Using the corollaries, a parameterization of SEP together with convex optimization methods can then be applied to $\mathcal{M}_d \cap \text{SEP}$, by either optimizing for the convex set of SEP and mapping the result to $\mathcal{M}_d$ (equation (8)), or by directly considering the action of $\mathcal{P}$ for the parameterization.

2.3.2. Error correction for a maximally entangled qudit

The Weyl-Twirl operators $T_{i,j}$ (equation (4)) allow for a simple error identification and correction scheme for the process of sharing a maximally entangled state without access to the state itself. Assume that in some information processing task, the initial state $|\Omega_{0,0}\rangle$ is transformed to $|\Omega_{k,l}\rangle$ with probability $p_{k,l}$ so that the error channel state $\mathcal{E}(P_{0,0})$ is represented as

$$\mathcal{E}(P_{0,0}) = \sum_{k,l=0}^{d-1} p_{k,l} P_{k,l}. \quad (11)$$

Another possibility that leads to the form of equation (11) is a lost or unknown measurement outcome in the Bell basis. The task is now to identify in which Bell states $|\Omega_{k,l}\rangle$ the system is without measuring or disturbing it and optionally to transform it back to the initial state $|\Omega_{0,0}\rangle$. The idea is to use equation (6) to store the

state dependent phase $\Phi = jk - il$ in an ancilla qudit. After the phase has been identified, the state is known and can be transformed back to the initial state.

Consider a system to be in the initial state $|\psi_0\rangle$, where the Bell state $|\Omega_{k,l}\rangle$ is unknown and the ancilla qudit in a fixed state, i.e.

$$|\psi_0\rangle = |0\rangle \otimes |\Omega_{k,l}\rangle.$$

Suppose further that the following operations are available:

- Generalized Hadamard/Fourier gate F acting as $F|j\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} w^{kj}|k\rangle$.
- Controlled Weyl-Twirl gate $CT_{i,j}$ acting as $CT_{i,j} |m\rangle \otimes |n\rangle = |m\rangle \otimes T_{i,j}^m |n\rangle$.

Consider the application of the Fourier gate to the ancilla, followed by the controlled Weyl-Twirl gate controlled by the ancilla with the unknown Bell state as target, followed by the adjoint Fourier gate on the ancilla qudit. These operations act as follows:

$$\begin{aligned} |\psi_0\rangle &\rightarrow (F^\dagger \otimes \mathbb{1}) CT_{i,j} (F \otimes \mathbb{1}) (|0\rangle \otimes |\Omega_{k,l}\rangle) \\ &= (F^\dagger \otimes \mathbb{1}) CT_{i,j} \left(\frac{1}{\sqrt{d}} \sum_k |k\rangle \otimes |\Omega_{k,l}\rangle \right) \\ &= (F^\dagger \otimes \mathbb{1}) \left(\frac{1}{\sqrt{d}} \sum_k |k\rangle \otimes T_{i,j}^k |\Omega_{k,l}\rangle \right) \\ &= (F^\dagger \otimes \mathbb{1}) \left(\frac{1}{\sqrt{d}} \sum_k |k\rangle \otimes w^{k\Phi} |\Omega_{k,l}\rangle \right) \\ &= \left(F^\dagger \frac{1}{\sqrt{d}} \sum_k w^{k\Phi} |k\rangle \right) \otimes |\Omega_{k,l}\rangle \\ &= F^\dagger F |\Phi\rangle \otimes |\Omega_{k,l}\rangle = |\Phi\rangle \otimes |\Omega_{k,l}\rangle. \end{aligned}$$

Measurement of the ancilla qudit yields the phase $|\Phi\rangle = |jk - il\rangle$, depending on the applied Weyl-Twirl operator through i and j and the unknown Bell state indexed by k and l . Due to the stabilizing property of the operators $T_{i,j}$, the Bell state is not disturbed and the operation can be repeatedly applied or run in parallel with additional ancilla qudits. The applied Weyl-Twirl operators can be chosen according to the expected error. If only phase or only shift errors appear, a single measurement with suitably chosen (i, j) can identify the error. If both shift and phase errors need to be identified, then two measurements of Φ obtained with different (i, j) for $T_{i,j}$ are required to identify the phase/shift error via k/l . The correction operation to recover $|\Omega_{0,0}\rangle$ is then $W_{k,l}^\dagger \otimes \mathbb{1} : |\Omega_{k,l}\rangle \rightarrow |\Omega_{0,0}\rangle$.

3. Generalized Bell-diagonal systems

Now we investigate cases, in which we still have a complete orthonormal Bell basis, but not via the standard construction introduced in the last section. In particular, this leads in general to a breaking of the channel equivalence $\mathcal{P} \equiv \mathcal{T}$.

3.1. Generalized bases of Bell states

Bell states are characterized by the fact that the reduced state for any of its subsystems is maximally mixed. Those states are called ‘locally maximally mixed’ and they imply that all information about the state is in the correlations between the subsystems. The standard Bell basis defined in (2) is of course an example of those states. However, there are bases that share this property, but are not unitarily equivalent to the standard basis [40].

For $k, l = 0, \dots, (d-1)$ consider the d^2 states

$$|\phi_{k,l}^\alpha\rangle := \frac{1}{\sqrt{d}} \sum_{s=0}^{d-1} w^{k(s-l)} \alpha_s |s-l\rangle \otimes |s\rangle \quad (12)$$

with $\alpha_s = e^{i\phi_s}$ being a phase factor. For a suitable choice of phase vectors $\alpha = \{\alpha_s\}_{s=0}^{d-1}$, these states form an orthonormal set of locally maximally mixed basis states. Consider the corresponding projection operators

$$|\phi_{k,l}^\alpha\rangle\langle\phi_{k,l}^\alpha| = \frac{1}{d} \sum_{s,t=0}^{d-1} w^{k(s-t)} \alpha_s \alpha_t^* |s-l\rangle\langle t-l| \otimes |s\rangle\langle t|,$$

for which any partial trace of the first or second subsystem yields the maximally mixed state:

$$\text{Tr}_{1/2}(|\phi_{k,l}^\alpha\rangle\langle\phi_{k,l}^\alpha|) = \frac{1}{d} \sum_{s=0}^{d-1} \alpha_s \alpha_s^* |s\rangle\langle s| = \frac{1}{d} \sum_{s=0}^{d-1} |s\rangle\langle s|. \quad (13)$$

Furthermore, the orthonormality condition for $|\phi_{m,n}^\beta\rangle$ and $|\phi_{k,l}^\alpha\rangle$ reads:

$$\delta_{m,k} \delta_{n,l} = \langle\phi_{m,n}^\beta|\phi_{k,l}^\alpha\rangle = \frac{1}{d} \sum_{s,t} \alpha_s \beta_t^* w^{k(s-l)-m(t-n)} \langle t-n|s-l\rangle\langle t|s\rangle = \delta_{n,l} \frac{1}{d} \sum_s w^{s(k-m)} w^{mn-kl} \alpha_s \beta_s^*. \quad (14)$$

This shows that for different shift indices (n, l) , the Bell states (12) are orthogonal, independent of the phases (α, β) . For equal shift indices, we now require $\alpha_s = \beta_s$, which implies:

$$\langle\phi_{m,l}^\beta|\phi_{k,l}^\alpha\rangle = \frac{1}{d} w^{l(m-k)} \sum_s w^{s(k-m)} \alpha_s \alpha_s^* = \delta_{m,k}. \quad (15)$$

Thus, in order to meet the requirements for orthonormality and being locally maximally mixed, we define the matrix $(\alpha_{s,t})_{s,t=0,\dots,d-1}$, $|\alpha_{s,t}| = 1 \forall s, t$ and use it to define the ‘generalized’ Bell basis $\{|\Phi_{k,l}^\alpha|k, l = 0, \dots, d-1\}$ with

$$|\Phi_{k,l}^\alpha\rangle := \frac{1}{\sqrt{d}} \sum_s w^{k(s-l)} \alpha_{s,l} |s-l\rangle \otimes |s\rangle. \quad (16)$$

In accordance with equations (1) and (2), we define unitary operators that map the maximally entangled state $|\Omega_{0,0}\rangle$ to the generalized Bell basis states:

$$\begin{aligned} V_{k,l}^\alpha &:= \sum_j w^{jk} \alpha_{j+l,l} |j\rangle\langle j+l|, \quad k, l = 0, \dots, d-1 \\ |\Phi_{k,l}^\alpha\rangle &= V_{k,l}^\alpha \otimes \mathbb{1} |\Omega_{0,0}\rangle \\ P_{k,l}^\alpha &:= |\Phi_{k,l}^\alpha\rangle\langle\Phi_{k,l}^\alpha|. \end{aligned} \quad (17)$$

Note that standard Weyl–Heisenberg operators $W_{k,l}$ and Bell basis states $|\Omega_{k,l}\rangle$ are a special case of equation (16), namely for $\alpha_{s,t} = 1 \forall s, t$. In this sense we can talk about a generalization of the standard Weyl–Heisenberg operators and related Bell states.

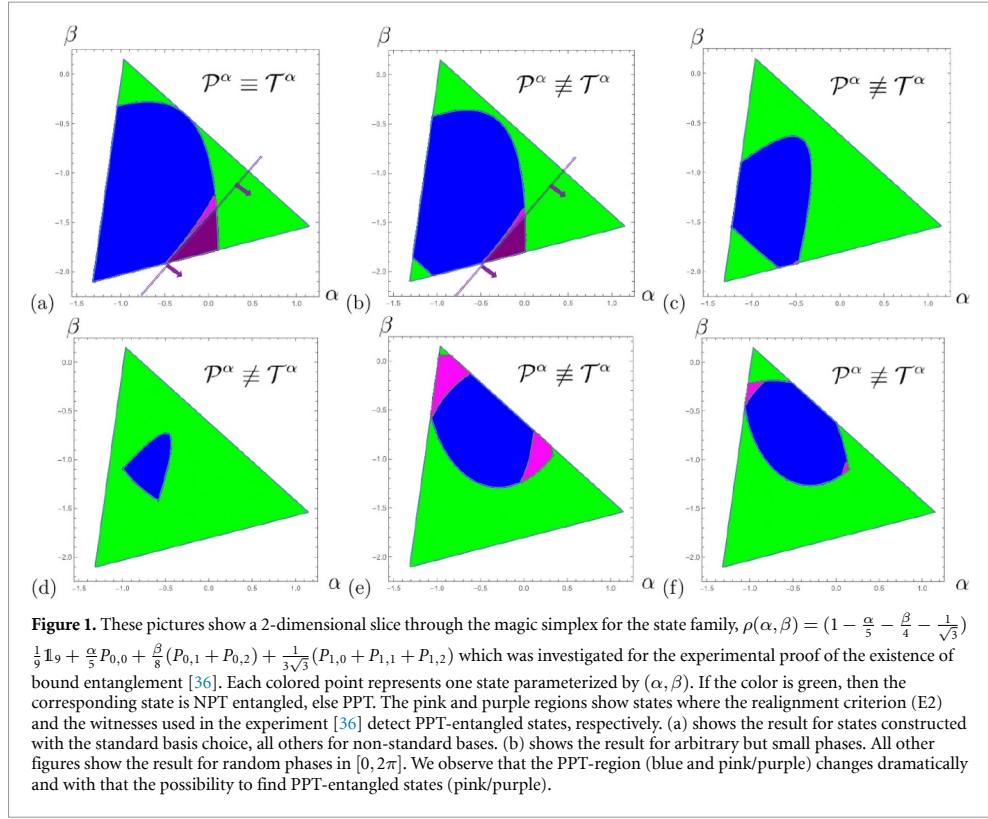
3.2. Properties of the generalized Bell bases

The construction of generalized Bell basis introduced in the previous section (equation (17)) looks rather similar to the standard case (equations (1) and (2)) and generates also a set of orthonormal, maximally entangled and locally maximally mixed states. However, these Bell bases show significant differences with implications for the entanglement properties of diagonal states. We first state those differences and then numerically demonstrate their effect on the entanglement structure of corresponding systems of Bell-diagonal states.

The following differences to the standard construction hold for general transformation matrices α and can be checked by simple calculation or counter examples:

1. The linear group structure does generally not exist: $V_{k_1,l_1}^\alpha V_{k_2,l_2}^\alpha \not\propto V_{k_1+k_2,l_1+l_2}^\alpha$.
2. The stabilizing property does generally not hold: $V_{i,j}^\alpha \otimes V_{i,j}^{\alpha*} |\Phi_{k,l}\rangle \neq w^\phi |\Phi_{k,l}\rangle$.
3. Let the simplex [40]

$$\mathcal{M}_d^\alpha := \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l}^\alpha \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right\} \quad (18)$$



be the set of states that are diagonal in the generalized Bell basis. The generalized Pauli channel

$$\mathcal{P}^\alpha : \mathcal{H} \rightarrow \mathcal{M}_d^\alpha, \quad \mathcal{P}^\alpha(\rho) := \sum_{k,l=0}^{d-1} P_{k,l}^\alpha \rho P_{k,l}^\alpha$$

and the generalized Weyl-Twirl channel

$$\mathcal{T}^\alpha(\rho) := \frac{1}{d^2} \sum_{i,j=0}^{d-1} V_{i,j}^\alpha \otimes V_{i,j}^{\alpha*} \rho (V_{i,j}^\alpha \otimes V_{i,j}^{\alpha*})^\dagger$$

are generally not identical: $\mathcal{P}^\alpha \neq \mathcal{T}^\alpha$.

4. The generalized Pauli channel does generally not conserve separability, so corollary 1 does not hold, i.e.: $\rho_s \in \text{SEP} \not\Rightarrow \mathcal{P}^\alpha(\rho_s) \in \mathcal{M}_d^\alpha \cap \text{SEP}$.
5. Due to the loss of the linear structure, also entanglement conserving symmetries of $\mathcal{M}_d$ (see [42] and the references therein) are lost.

3.3. Explicit counter examples for bipartite qutrits

In figure 1 we present a visualization of the drastic change in the separability structure of a state family within $\mathcal{M}_d^\alpha$ for dimension $d = 3$ for non-standard Bell basis choices compared to the standard Bell basis choice. The state family consists of mixtures of Bell projectors, which were experimentally investigated in [36]. This family is constructed to show the greatest violation for a specific entanglement witness based on mutually unbiased bases [51], which can in principle be realized in experiments. Figure 1(a) shows the result of the standard basis choice, including the region detected by the entanglement witness. In figure 1(b) the phases defining a generalized Bell basis (equation (17)) were randomly chosen but close to zero. One observes that the PPT-region shrinks as well as the region detected by the realignment criterion E2 (defined in the appendix, named E2 according to [41]) and the entanglement witness. The other figures (c)–(f) are based on random choices of the phases for defining the generalized Bell projectors and show how drastically the PPT region and the region detected by the realignment criterion (E2) changes. The Weyl–Heisenberg structure is thus very important for the separability structure within $\mathcal{M}_3^\alpha$.

Table 1. Minimum, maximum and mean statistics for 1000 sample systems $\mathcal{M}_3^\alpha$ and the reference values for the standard system $\mathcal{M}_3$. Relative volume of PPT states (rPPT) and the share of PPT-entangled states among them for the realignment criterion (E2/PPT), the quasipure approximation criterion (E3/PPT) and the combined criterion ((E2&E3)/PPT).

	Std. Basis	Min	Max	Mean
rPPT	60.0%	49.0%	59.0%	51.6%
E2/PPT	10.4%	2.1%	9.6%	3.6%
E3/PPT	2.7%	0.6%	2.3%	1.0%
(E2&E3)/PPT	1.8%	0.1%	1.4%	0.4%

Table 2. Correlation coefficient for the relative volume of PPT states (rPPT) and the share of PPT-entangled states among the PPT states as detected by the realignment criterion (E2/PPT), the quasipure concurrence criterion (E3/PPT) and the combined criterion ((E2&E3)/PPT).

	rPPT
E2/PPT	0.99
E3/PPT	0.95
(E2&E3)/PPT	0.96

3.4. General entanglement structure changes for bipartite qutrits

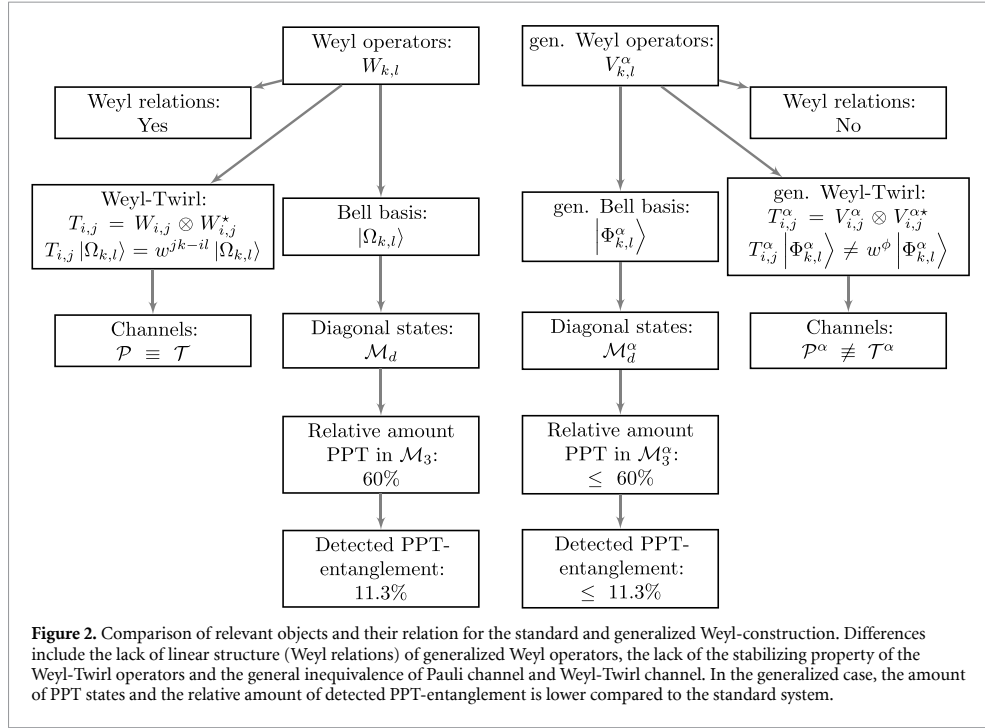
In the last section we have presented particular examples, but we can also come up with general results for $d = 3$ by exploiting and adapting the tools developed in [41, 42]. A combination of analytical and numerical methods [43] exploiting the Heisenberg–Weyl structure were used to analyze the entanglement structure of Bell-diagonal qutrits for $d = 3$ and $d = 4$, leading to an efficient solution of the separability problem in this particular case. As a consequence of the stated differences above, most of the tools cannot be applied for mixtures of the generalized Bell basis states (16). However, some methods are still applicable that allow us to compare the standard basis with the generalized Bell bases for dimension $d = 3$.

In order to compare the system $\mathcal{M}_3$ to $\mathcal{M}_3^\alpha$, we use the same uniformly distributed sample set of mixing probabilities $c_{k,l}$ to construct 10 000 states in $\mathcal{M}_3$ and in $\mathcal{M}_3^\alpha$ for various α . In particular, we generate 1000 matrices α with uniformly distributed elements. For each α , we define the Bell basis projectors $P_{k,l}^\alpha$ and construct 10 000 diagonal states according to (18). We then analyze these states for PPT and entanglement detection by the realignment criterion [52] (named E2 in [41]) and the quasipure concurrence criterion [53] (named E3 in [41]) (see appendix for definitions), which do not depend on the entanglement class preserving symmetries of the standard simplex $\mathcal{M}_3$, and are the most successful analytical criteria among those investigated in [41, 42].

Table 1 shows the share of PPT states among all analyzed states (rPPT), the share of PPT-entangled states detected by the realignment criterion among all PPT states (E2/PPT), and the corresponding share for the quasipure concurrence criterion (E3/PPT), as well as the share of entangled states that were detected by both criteria simultaneously (E2&E3)/PPT. It shows that for the 1000 realizations of $\mathcal{M}_3^\alpha$ the relative volume of PPT states is between 49% and 59%. The mean of 51.6% indicates most of the systems are closer to the minimum value than to the maximum. Indeed, each of more than 85% of the systems have less than 54% PPT states. Curiously, the PPT share for the standard system $\mathcal{M}_3$ is higher than for any analyzed $\mathcal{M}_3^\alpha$. For values of all α being sufficiently close to 1, i.e. for Bell basis being close to the standard Bell basis, the observed quantities are arbitrarily close, but never higher than the reference values for the standard system. The same statements hold for the detection capabilities of E2, E3 and (E2&E3), which are always lower than the value for the standard system and are typically closer to the minimal observed value.

Table 2 presents strong positive correlations between the share of PPT states in the systems $\mathcal{M}_3^\alpha$ and the relative detection capabilities of the criteria E2, E3 and simultaneous detection (E2&E3). Interestingly, the more PPT states are present, the larger the relative volume of PPT-entangled states is, which are detected by those criteria. Moreover, the relative number of simultaneously detected states is then also high, in general. The strong correlations imply an almost linear dependence between those quantities.

Based on these observations, one can assume that the more states with PPT are present in a Bell-diagonal system, the larger the share of PPT-entangled states, detected by the two effective entanglement criteria E2 and E3, is. Additionally, a higher share of PPT states seems to imply a higher share of PPT-entangled states that can be detected by both of the criteria simultaneously. While in the standard system only 33% of the states detected by E3 are not also detected by E2, there are Bell-diagonal systems with less PPT states, in which 83% of the entangled states detected by E3 are not detected by E2. Curiously, the standard system shows extreme values for the relative volume of PPT states and the number of entangled states among them.



4. Summary and outlook

In this work, we investigated and compared properties of bipartite maximally entangled and locally maximally mixed Bell states, known as Bell-diagonal states and showed that the special features of the Weyl–Heisenberg Bell basis imply special features in the entanglement structure of Bell-diagonal states.

The frequently used ‘standard’ construction of a d^2 dimensional Bell basis of the joint Hilbert space $\mathcal{H}_d \otimes \mathcal{H}_d$ via the Weyl–Heisenberg operators is presented and properties of this special basis were derived that are strongly related to the separability problem and other applications, for example, error corrections or channel equivalences. In particular, utilizing the Weyl relations of the Weyl–Heisenberg operators $W_{k,l}$, we showed that the ‘Weyl-Twirl’ operators $W_{i,j} \otimes W_{i,j}^*$ are diagonalized by all elements of the standard Bell basis. We then leveraged this stabilizing property to show the equivalence of the ‘Pauli’ channel, $\mathcal{P}$, which projects onto Bell-diagonal states in the standard Bell basis, and the ‘Weyl-Twirl’ channel, $\mathcal{T}$, which represents the randomized application of the Weyl-Twirl operators.

One implication of this channel equivalence, $\mathcal{P} \equiv \mathcal{T}$, is that the separability is conserved under the channels and that optimal Bell-diagonal entanglement witnesses remain optimal for the set of Bell-diagonal states. We then demonstrated several applications of the Weyl-Twirl operators. On the one hand, they allow the efficient parameterization of separable Bell-diagonal states, relevant for polytope approximations of this convex set and the construction of optimal Bell-diagonal entanglement witnesses. These applications, enabled by the special structure of investigated Bell states, allow the effective entanglement classification and detection of PPT-entanglement for Bell-diagonal states. On the other hand, a simple error detection and correction scheme for a maximally entangled qudit was presented.

Highlighting the implication of the channel equivalence, $\mathcal{T} \equiv \mathcal{P}$, we investigated systems of Bell-diagonal states that are constructed from generalized Bell basis states. Those generalized Bell states form an orthonormal basis and are locally maximally mixed, similar to the standard basis based on Heisenberg–Weyl operators. However, the channel equivalence $\mathcal{T} \equiv \mathcal{P}$ and other properties are lost, which has strong implications on the entanglement structure within the family of Bell-diagonal states. We summarize the most relevant objects, their properties and relations in figure 2. Thereby, we compare the results related to the standard Bell basis to the generalized Bell bases.

In detail, we analyzed this for bipartite qutrits. We derived the relative volume of PPT states among Bell-diagonal states. For the standard basis, the relative volume is 60% while it was shown to drop to 49% for some generalized Bell bases. Interestingly, we found that the share of Bell-diagonal PPT states is generally

lower for the generalized Bell bases compared to the standard Bell basis. Moreover, the relative detection rate for PPT-entangled states of two, in the standard system highly effective, entanglement criteria strongly correlates to the share of PPT states. The more states with PPT exist for a system of Bell-diagonal states, the higher is the share of detected PPT-entangled states among them for those two criteria. Consequently, for the system based on the standard Bell basis, which has the highest PPT share, also the highest relative amount of PPT-entanglement is observed. Furthermore, we visualized the dramatic change of the volume of PPT states in figure 1 for a family of states investigated previously in experiment.

These results indicate that, among all in this way generalized Bell bases, Bell-diagonal states related to the standard construction have some very special properties. These properties have relevant implications on the entanglement structure and on practical applications using those Bell states. Numerical results suggest that either the two entanglement criteria are less effective for Bell-diagonal states if the special properties of the standard system are not given or that the structure and amount of PPT entanglement depend on those properties.

In summary, our findings are the starting point to analyze different quantum information theoretic protocols with regards to their dependence on the underlying structure and properties of used Bell states.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://github.com/kungfugo/BellDiagonalQudits.jl>.

Acknowledgment

B C H and C P acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) Project P36102-N. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Code availability statement

The software used to generate the reported results is published as repository and registered open source package ‘BellDiagonalQudits.jl’ [43] available at <https://github.com/kungfugo/BellDiagonalQudits.jl>.

Author contributions statement

C P carried out the analytical and numerical analyses, developed the new methods and implemented the software. B C H revised the analyses and added results and proposed improvements. C P and B C H both edited the manuscript.

Conflict of interest

All authors declare no financial or non-financial competing interests.

Additional information

Correspondence and requests for materials should be addressed to C P.

Appendix

A.1. Applied entanglement criteria

A.1.1. Peres–Horodecki/PPT criterion

The partial transpose Γ acts on the computational basis states of a bipartite state as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^\Gamma := |i\rangle\langle j| \otimes |l\rangle\langle k|$. If the partial transpose for a density matrix of a bipartite state has at least one negative eigenvalue (in which case it is said to be ‘NPT’), it is entangled [48]. For $d = 2$ it detects all entangled states, but for $d \geq 3$ it is only sufficient due to the existence of PPT-entangled states.

A.1.2. E2: Realignment criterion

The realignment operation R is defined for the computational basis states as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)_R := |i\rangle\langle k| \otimes |j\rangle\langle l|$. If the sum of singular values of the density matrix of a realigned state σ_R are larger than 1, then σ is entangled [52].

A.1.3. E3: Quasipure concurrence criterion

The quasipure approximation [53] C_{qp} provides an efficiently computable lower bound on the concurrence [54] C . Let i and j (k and l) enumerate the computational basis vectors of the first (second) partition of the Hilbert space. Define $A := 4 \sum_{i < j, k < l} |ikjl\rangle - |jkil\rangle - |iljk\rangle + |jljk\rangle \times h.c.$ and let a state be represented in its spectral composition $\rho = \sum_i \mu_i |\Psi_i\rangle\langle\Psi_i|$. With a dominant eigenvector $|\Psi_0\rangle$, define $|\xi\rangle \propto A|\Psi_0\rangle \otimes |\Psi_0\rangle$ and

$$T_{i,j} := \sqrt{\mu_i \mu_j} \langle\Psi_i| \otimes \langle\Psi_j|\xi\rangle.$$

The quasipure concurrence is defined by the singular values S_i of the matrix $T_{i,j}$

$$C_{qp}(\rho) := \max\left(0, S_0 - \sum_{i>0} S_i\right) \leq C(\rho)$$

and detects entanglement for positive values.

For the standard system $\mathcal{M}_d$, the approximation takes an explicit form: Here, the singular values $S_{k,l}$ are explicitly given by

$$S_{k,l} = \sqrt{\frac{d}{2(d-1)} c_{k,l} \left[\left(1 - \frac{2}{d}\right) c_{n,m} \delta_{k,n} \delta_{l,m} + \frac{1}{d^2} c_{(2n-k) \bmod d, (2m-l) \bmod d} \right]}$$

where (n, m) is a multi-index of the coordinate of the largest value $\{c_{k,l}\}$ [53].

ORCID iDs

Christopher Popp  <https://orcid.org/0000-0002-2352-3731>

Beatrix C Hiesmayr  <https://orcid.org/0000-0001-9062-6039>

References

- [1] Cirac J I, Ekert A K, Huelga S F and Macchiavello C 1999 Distributed quantum computation over noisy channels *Phys. Rev. A* **59** 4249–54
- [2] Georgescu I M, Ashhab S and Nori F 2014 Quantum simulation *Rev. Mod. Phys.* **86** 153–85
- [3] Ekert A K 1991 Quantum cryptography based on Bell's theorem *Phys. Rev. Lett.* **67** 661–3
- [4] Sheng Y B, Zhou L and Long G L 2022 One-step quantum secure direct communication *Sci. Bull.* **67** 367–74
- [5] Deng F G, Long G L and Liu X S 2003 Two-step quantum direct communication protocol using the Einstein-Podolsky-Rosen pair block *Phys. Rev. A* **68** 042317
- [6] Bennett C H and Wiesner S J 1992 Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states *Phys. Rev. Lett.* **69** 2881–4
- [7] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels *Phys. Rev. Lett.* **70** 1895–9
- [8] Moskal P *et al* 2018 Feasibility studies of the polarization of photons beyond the optical wavelength regime with the J-PET detector *Eur. Phys. J. C* **78** 970
- [9] Hiesmayr B C and Moskal P 2017 Genuine multipartite entanglement in the 3-photon decay of positronium *Sci. Rep.* **7** 15349
- [10] Hiesmayr B C and Moskal P 2019 Witnessing entanglement in compton scattering processes via mutually unbiased bases *Sci. Rep.* **9** 8166
- [11] Krzemien W, Pérez del Río E and Kacprzak K 2020 Feasibility of ortho-positronium lifetime studies with the J-PET detector in context of mirror matter models *Acta Phys. Pol. B* **51** 165
- [12] Shopa R *et al* 2021 Optimisation of the event-based TOF filtered back-projection for online imaging in total-body J-PET *Med. Image Anal.* **73** 102199
- [13] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 777–80
- [14] Bell J S 1964 On the Einstein Podolsky Rosen paradox *Phys. Phys. Fiz.* **1** 195–200
- [15] Hensen B *et al* 2015 Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres *Nature* **526** 682–4
- [16] Braunstein S L, Mann A and Revzen M 1992 Maximal violation of Bell inequalities for mixed states *Phys. Rev. Lett.* **68** 3259–61
- [17] Wilde M M 2017 *Quantum Information Theory* 2nd edn (Cambridge University Press)
- [18] Cozzolino D, Da Lio B, Bacco D and Oxenlowe I K 2019 High-dimensional quantum communication: benefits, progress and future challenges *Adv. Quantum Technol.* **2** 1900038
- [19] Wang Y, Hu Z, Sanders B C and Kais S 2020 Qudits and high-dimensional quantum computing *Front. Phys.* **8** 479
- [20] Sych D and Leuchs G 2009 A complete basis of generalized Bell states *New J. Phys.* **11** 013006
- [21] Baumgartner B, Hiesmayr B C and Narnhofer H 2007 A special simplex in the state space for entangled qudits *J. Phys. A: Math. Theor.* **40** 7919
- [22] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed-state entanglement and quantum error correction *Phys. Rev. A* **54** 3824–51
- [23] Calderbank A R, Rains E M, Shor P W and Sloane N J A 1997 Quantum error correction and orthogonal geometry *Phys. Rev. Lett.* **78** 405–8

- [24] Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J A and Wootters W K 1996 Purification of noisy entanglement and faithful teleportation via noisy channels *Phys. Rev. Lett.* **76** 722–5
- [25] Ashikhmin A and Knill E 2001 Nonbinary quantum stabilizer codes *IEEE Trans. Inf. Theory* **47** 3065–72
- [26] Knill E 1996 Non-binary unitary error bases and quantum codes (arXiv:quant-ph/9608048)
- [27] Nadakarni P J and Garani S S 2021 Non-binary entanglement-assisted stabilizer codes *Quantum Inf. Process.* **20** 258
- [28] Horodecki M and Horodecki P 1999 Reduction criterion of separability and limits for a class of distillation protocols *Phys. Rev. A* **59** 4206–16
- [29] Alber G, Delgado A, Gisin N and Jex I 2001 Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces *J. Phys. A: Math. Gen.* **34** 8821
- [30] Vollbrecht K G H and Wolf M M 2003 Efficient distillation beyond qubits *Phys. Rev. A* **67** 012303
- [31] Horodecki M, Horodecki P and Horodecki R 1998 Mixed-state entanglement and distillation: is there a ‘bound’ entanglement in nature? *Phys. Rev. Lett.* **80** 5239–42
- [32] Bej P and Halder S 2021 Unextendible product bases, bound entangled states and the range criterion *Phys. Lett. A* **386** 126992
- [33] Lockhart J, Gühne O and Severini S 2018 Entanglement properties of quantum grid states *Phys. Rev. A* **97** 062340
- [34] Bruß D and Peres A 2000 Construction of quantum states with bound entanglement *Phys. Rev. A* **61** 030301
- [35] Slater P B 2019 Jagged islands of bound entanglement and witness-parameterized probabilities (arXiv:1905.09228)
- [36] Hiesmayr B C and Löffler W 2013 Complementarity reveals bound entanglement of two twisted photons *New J. Phys.* **15** 083036
- [37] Gurvits L 2003 Classical deterministic complexity of Edmonds’ problem and quantum entanglement *Proc. 35th Annual ACM Symp. on Theory of Computing (STOC 2003)* (Association for Computing Machinery) pp 10–19
- [38] Gharibian S 2008 Strong NP-hardness of the quantum separability problem *Quantum Inf. Comput.* **10** 343–60
- [39] Kaszlikowski D, Oi D K L, Christandl M, Chang K, Ekert A, Kwek L C and Oh C H 2003 Quantum cryptography based on qutrit Bell inequalities *Phys. Rev. A* **67** 012310
- [40] Baumgartner B, Hiesmayr B C and Narnhofer H 2006 State space for two qutrits has a phase space structure in its core *Phys. Rev. A* **74** 032327
- [41] Popp C and Hiesmayr B C 2022 Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits *Sci. Rep.* **12** 12472
- [42] Popp C and Hiesmayr B C 2023 Comparing bound entanglement of Bell diagonal pairs of qutrits and ququarts *Sci. Rep.* **13** 2037
- [43] Popp C 2023 BellDiagonalQudits: a package for entanglement analyses of mixed maximally entangled qudits *J. Open Source Softw.* **8** 4924
- [44] Hiesmayr B C 2021 Free versus bound entanglement, a NP-hard problem tackled by machine learning *Sci. Rep.* **11** 19739
- [45] Sakuldee F and Rudnicki Ł 2023 Bounds on the breaking time for entanglement-breaking channels *Phys. Rev. A* **107** 022430
- [46] Baumgartner B, Hiesmayr B C and Narnhofer H 2008 The geometry of bipartite qutrits including bound entanglement *Phys. Lett. A* **372** 2190–5
- [47] Hill S A and Wootters W K 1997 Entanglement of a pair of quantum bits *Phys. Rev. Lett.* **78** 5022–5
- [48] Peres A 1996 Separability criterion for density matrices *Phys. Rev. Lett.* **77** 1413–5
- [49] Spengler C, Huber M and Hiesmayr B C 2012 Composite parameterization and Haar measure for all unitary and special unitary groups *J. Math. Phys.* **53** 013501
- [50] Spengler C, Huber M and Hiesmayr B C 2010 A composite parameterization of unitary groups, density matrices and subspaces *J. Phys. A: Math. Theor.* **43** 385306
- [51] Hiesmayr B C, McNulty D, Baek S, Roy S S, Bae J and Chruściński D 2021 Detecting entanglement can be more effective with inequivalent mutually unbiased bases *New J. Phys.* **23** 093018
- [52] Chen K and Wu L A 2002 A matrix realignment method for recognizing entanglement *Quantum Inf. Comput.* **3** 193–202
- [53] Bae J, Tiersch M, Sauer S, de Melo F, Mintert F, Hiesmayr B and Buchleitner A 2009 Detection and typicality of bound entangled states *Phys. Rev. A* **80** 022317
- [54] Wootters W K 1998 Entanglement of formation of an arbitrary state of two qubits *Phys. Rev. Lett.* **80** 2245–8

B.5. A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria

A Group-Theoretic Perspective on the PPT and Realignment Entanglement Criteria

Tobias C. Sutter¹, Christopher Popp², and Beatrix C. Hiesmayr³

University of Vienna, Faculty of Physics, Währingerstrasse 17, 1090 Vienna.

¹Tobias.Christoph.Sutter@univie.ac.at

²Christopher.Popp@univie.ac.at

³Beatrix.Hiesmayr@univie.ac.at

Abstract

Entanglement is a key feature in many quantum technologies, including secure communication protocols and quantum computing. However, detecting it in mixed quantum states remains a challenging task. While the PPT and realignment criteria are well-established tools for entanglement detection in general, and are especially effective in Bell-diagonal states, their connection to the underlying group structure of this state family has not been fully explored. In this work, we analyze the PPT and realignment criteria for Bell-diagonal states from a group-theoretic point of view. Our results demonstrate that the group structure of Bell-diagonal states provides a clear framework for analyzing and computing these two entanglement detection criteria. This unified perspective offers new insights into the mathematical and physical properties of entanglement in structured quantum systems, like the connection of the PPT criterion to experimentally accessible entanglement witnesses, thus opening a path for extending the group formalism to more complex quantum systems.

1 Introduction

Many quantum technologies, including quantum computing [1], cryptography [2], and teleportation [3], rely on entanglement as a fundamental resource. As these technologies advance, a deeper understanding of entanglement becomes increasingly essential for creating and manipulating resource states. Still, one of the major open problems in quantum information theory is determining whether a given bipartite quantum state is separable or entangled [4]. For mixed quantum states, which arise naturally in experiments due to decoherence, this so-called *separability problem* is provably hard when considering systems of arbitrary dimensions [5, 6]. Nevertheless, efficient solutions may exist for fixed local dimensions or particular state families.

In the special case of pure states of dimension $d_A \times d_B$, the separability problem has a conclusive and efficiently computable answer [7]. Furthermore, it is also solved for general qubit-qubit ($d_A = d_B = 2$) and qubit-qutrit ($d_A = d_B - 1 = 2$) systems, because the PPT criterion [8, 9] is necessary and sufficient for entanglement in these dimensions. However, for

higher dimensions ($d_A d_B > 6$), the PPT criterion ceases to be necessary for entanglement, leading to the notion of PPT-entangled states [10]. Physically, such states are relevant as they cannot be distilled to maximally entangled form using local operations and classical communication [11]. More generally, there exist multiple other efficiently computable criteria for entanglement detection that are necessary *or* sufficient (cf. [4, 12]). They aim to approximate a solution to the separability problem in higher dimensions, albeit in a one-sided manner. One prominent example is the realignment (or cross-norm) criterion [13, 14].

Besides the possibility of fixing the dimensions of the subsystems, a different way to obtain further results is to consider specific state families that bear desirable properties while still being experimentally relevant. One such family of bipartite quantum states is Bell-diagonal states [15, 16]. From a physical point of view, they appear naturally by considering (generalized) Pauli errors. Thus, they are intimately connected to applications like error correction and entanglement distillation by their mathematical group structure related to the Weyl operators [17, 18].

Regarding their entanglement structure, Ref. [19,

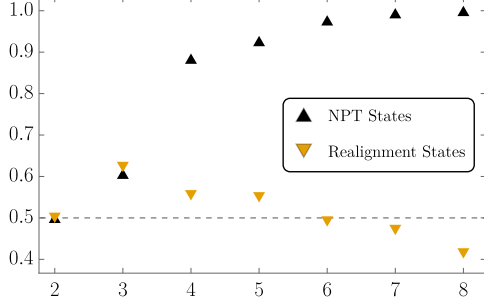


Figure 1: Percentages of Bell-diagonal states detected by the PPT and realignment criterion for $d = 2, \dots, 8$. For each d , a total of 10^5 uniformly sampled states were generated using the Julia package [22]. For qubits ($d = 2$), the PPT criterion is equivalent to the realignment criterion, both detecting around 50% of states as entangled, which are all existing entangled states in this state family. For $d = 3$, the realignment criterion is more effective (62.3%) in detecting entanglement than the PPT criterion (60.7%). In higher dimensions, the vast majority of states are NPT entangled. In contrast, the number of entangled states detected by the realignment criterion decreases with increasing dimension $d > 3$.

[20] found that the set of Bell-diagonal qutrit-qutrit states ($d_A = d_B = 3$) contains a significant share of PPT-entanglement, namely about 6%. This means that the PPT criterion is relatively ineffective for detecting entanglement in this subclass of states, being successful only for roughly 60% of all states (see Fig. 1). In contrast, the realignment criterion appears to be particularly effective in this dimension, detecting around 62% of all states as entangled. Importantly, this involves about 10% of the states where the PPT criterion fails. Further numerical results in Ref. [21] suggest that the group structure of Bell-diagonal states crucially affects the entanglement class. This leads to the central question of how the properties of Bell-diagonal states are related to the performance of the PPT and realignment criteria. Our contribution investigates this relationship analytically.

The work is organized as follows. Sec. 2 formalizes the PPT and realignment criteria. In Sec. 3, we introduce Bell-diagonal states with a focus on their group properties inherited by the Weyl matrices. We furthermore compute their Bloch vectors and find that they carry information about subgroups of the Weyl-Heisenberg group. Based on this, we derive an explicit version of the realignment criterion in Sec. 4, unveiling its dependence on the subgroups for Bell-diagonal

qutrits in Sec. 4.1. Sec. 5 shows that the PPT criterion for Bell-diagonal qutrits can also be understood in terms of the subgroups of the Weyl-Heisenberg group. We furthermore derive a corresponding NPT-witnesses for this subset of states, and prove a statement that can guide the search for PPT-entangled Bell-diagonal qutrits. Finally, Sec. 6 summarizes the results and gives an outlook on future research directions.

2 PPT and Realignment Criteria

Let $\mathcal{L}(\mathcal{H})$ be the set of linear operators mapping the Hilbert space $\mathcal{H}$ into itself, and $\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{L}(\mathcal{H}) \mid \rho \geq 0, \text{Tr}(\rho) = 1\}$ be the set of density matrices on $\mathcal{H}$. For bipartite systems, we set $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, for which $\mathcal{D}(\mathcal{H})$ can be partitioned into the sets SEP and ENT that denote separable and entangled states, respectively. Hence, $\text{SEP} \cup \text{ENT} = \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Any state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ can be written as

$$\rho = \sum_{i,j=0}^{d_A-1} \sum_{k,l=0}^{d_B-1} \rho_{ik,jl} |i\rangle\langle j| \otimes |k\rangle\langle l|, \quad (1)$$

where $\dim(\mathcal{H}_A) = d_A$ and $\dim(\mathcal{H}_B) = d_B$, and $\{|i\rangle\langle j|\}_{i,j=0}^{d_A-1}$ and $\{|k\rangle\langle l|\}_{k,l=0}^{d_B-1}$ constitute a basis of $\mathcal{D}(\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_B)$, respectively.

We define the partial transpose acting on the second subsystem, denoted by a superscript Γ , by its action on the basis states,

$$\begin{aligned} (|i\rangle\langle j| \otimes |k\rangle\langle l|)^\Gamma &:= (\mathbb{1}_A \otimes T_B)(|i\rangle\langle j| \otimes |k\rangle\langle l|) \\ &= |i\rangle\langle j| \otimes |l\rangle\langle k|, \end{aligned} \quad (2)$$

where $\mathbb{1}_A$ is the identity map on $\mathcal{L}(\mathcal{H}_A)$ and T_B is the transposition map on $\mathcal{L}(\mathcal{H}_B)$. The action of Γ extends to general states of the form (1) by linearity.

The PPT criterion (also known as Peres-Horodecki criterion) is a powerful sufficient criterion for entanglement:

Theorem 2.1 (PPT Criterion [8, 9]). *Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, and ρ^Γ be the partial transpose of ρ . Then*

$$\rho^\Gamma \not\geq 0 \implies \rho \in \text{ENT}. \quad (3)$$

If $\rho^\Gamma \not\geq 0$, i.e., it is not positive semidefinite, we say it is NPT entangled and write $\rho \in \text{NPT}$, and otherwise $\rho \in \text{PPT}$. This admits another partition of the space of bipartite density operators, namely $\text{PPT} \cup \text{NPT} = \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. By Thm. 2.1 it holds that $\rho \in \text{NPT} \implies \rho \in \text{ENT}$, but the converse is only true for $d_A d_B \leq 6$ [9, 10]. Hence, $\text{SEP} \neq \text{PPT}$ for higher-dimensional states, which crucially includes qutrit-qutrit systems with $d_A = d_B = 3$.

The realignment criterion (also known as cross-norm criterion) is another sufficient criterion for entanglement. It is based on the realignment (or shuffle) operation on $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$, which is defined by its action on the basis states as

$$\hat{R}(|i\rangle\langle j| \otimes |k\rangle\langle l|) := |i\rangle\langle k| \otimes |j\rangle\langle l|, \quad (4)$$

and extends to general $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by linearity. The realignment criterion reads:

Theorem 2.2 (Realignment Criterion [13, 14]). *Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and define the realigned state $\rho_R := \hat{R}(\rho)$. Then*

$$\text{Tr} \left(\sqrt{\rho_R^\dagger \rho_R} \right) > 1 \implies \rho \in \text{ENT}. \quad (5)$$

Note that $\text{Tr}(\sqrt{\rho_R^\dagger \rho_R})$ corresponds to the trace norm of ρ_R (i.e., the sum of its singular values). Computing the norm scales as $\mathcal{O}(\min(d_A^4 d_B^2, d_A^2 d_B^4))$ because ρ_R is a matrix of size $d_A^2 \times d_B^2$ [23]. The realignment criterion detects some PPT entanglement, even though it does not detect all NPT entangled states. Hence, it is neither stronger nor weaker than the PPT criterion.

Furthermore, the realignment criterion can be expressed in terms of the correlation tensor $T_{a,b} = \text{Tr}(\rho(G_a \otimes G_b))$, where $\{G_a\}_{a=0}^{d_A^2-1}$ and $\{G_b\}_{b=0}^{d_B^2-1}$ form orthonormal matrix bases of $\mathcal{D}(\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_B)$, respectively. The normalization condition is $\text{Tr}(G_a(G_a^\dagger)) = \delta_{a,a'}$ and $\text{Tr}(G_b(G_b^\dagger)) = \delta_{b,b'}$. With this, Thm. 2.2 can be expressed as [24]

$$\|T\|_{\text{Tr}} > 1 \implies \rho \in \text{ENT}, \quad (6)$$

where $\|T\|_{\text{Tr}} = \text{Tr}(\sqrt{TT^\dagger})$ is the trace norm of T .

We conclude this section by noting three similarities between these two entanglement criteria. (i) Both are based on permutations of the basis states. (ii) The defining operations represent generally non-physical transformations of the state ρ : The transposition is a positive but not completely positive map, and realignment does not preserve hermiticity. (iii) Specific other permutations of the basis elements, similar to (2) and (4), recover the criteria: First, both are trivially invariant under exchanging $\mathcal{H}_A$ and $\mathcal{H}_B$, i.e., mapping $\rho \mapsto \mathbb{F} \rho \mathbb{F}^\dagger$ where $\mathbb{F} = \sum_{i=0}^{d_B-1} \sum_{j=0}^{d_A-1} |i\rangle\langle j| \otimes |j\rangle\langle i|$. Furthermore, for the PPT criterion, we can equivalently take the partial transposition on the first subsystem ρ^{T_A} , acting as $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^{T_A} = |j\rangle\langle i| \otimes |k\rangle\langle l|$ on the basis states. This leaves the criterion unchanged as the spectrum of $\rho^\Gamma = (\rho^{T_A})^T$ is unaffected by transposition. Regarding the realignment criterion, we note that $\mathbb{F}_A \hat{R}(\rho)$,

$\hat{R}(\rho) \mathbb{F}_B$, $\mathbb{F}_A \hat{R}(\rho) \mathbb{F}_B$, $\hat{R}(\rho)^T$, $\mathbb{F}_A \hat{R}(\rho)^T$, $\hat{R}(\rho)^T \mathbb{F}_B$, $\mathbb{F}_A \hat{R}(\rho)^T \mathbb{F}_B$ define different permutations of the basis states, where $\mathbb{F}_X = \sum_{i=0}^{d_X-1} \sum_{j=0}^{d_X-1} |i\rangle\langle j| \otimes |j\rangle\langle i|$ with $X \in \{A, B\}$ are the flip operators on $\mathcal{H}_A$ and $\mathcal{H}_B$. As the singular values of $\hat{R}(\rho)$ are invariant under taking the transposition and/or multiplying it by unitaries from the left or right, these permutations again lead to the realignment criterion.

3 Bell-Diagonal States with Weyl-Structure

Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$ be the Hilbert space of a bipartite qudit system with equal local dimensions. Maximally entangled states $|\psi\rangle \in \mathcal{H}$, i.e., pure states satisfying $\text{Tr}_A(|\psi\rangle\langle\psi|) = \text{Tr}_B(|\psi\rangle\langle\psi|) = \frac{1}{d} \mathbb{1}_d$, are called Bell states on $\mathcal{H}$. A set of d^2 linearly independent Bell states on $\mathcal{H}$ is referred to as a Bell basis (of $\mathcal{H}$). A particularly useful one is generated using the d -dimensional unitary Weyl-Heisenberg operators

$$W_{k,l} = \sum_{j=0}^{d-1} \omega^{jk} |j\rangle\langle j+l| \quad (7)$$

where $k, l \in \{0, 1, \dots, d-1\}$ and $\omega := e^{\frac{2\pi i}{d}}$, and, throughout this paper, variables taking values in $\{0, 1, \dots, d-1\}$ are understood as elements of $\mathbb{Z}_d := \mathbb{Z}/d\mathbb{Z}$, i.e., the ring of integers with addition and multiplication modulo d . The matrices (7) satisfy the so-called Weyl relations

$$\begin{aligned} W_{i,j} W_{k,l} &= \omega^{jk} W_{i+k, j+l}, \\ W_{k,l}^\dagger &= \omega^{kl} W_{-k, -l} = W_{k,l}^{-1}, \\ W_{k,l}^* &= W_{-k, l}, \\ W_{k,l}^T &= \omega^{-kl} W_{k, -l}, \end{aligned} \quad (8)$$

where $W_{k,l}^*$, $W_{k,l}^T$, and $W_{k,l}^\dagger$ denote the complex conjugate, transpose, and conjugate transpose of $W_{k,l}$, respectively. Writing $W_{k,l} = Z^k X^l$ with $Z = \sum_{j=0}^{d-1} \omega^j |j\rangle\langle j|$ and $X = \sum_{j=0}^{d-1} |j\rangle\langle j+1|$, the indices k and l in $W_{k,l}$ can be identified with phase and shift operations, respectively. Applying $W_{k,l}$ to one half of the maximally entangled state $|\Omega_{0,0}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle$ yields

$$|\Omega_{k,l}\rangle = (W_{k,l} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle, \quad (9)$$

with which the set $\{|\Omega_{k,l}\rangle\}_{k,l=0}^{d-1}$ constitutes a complete Bell basis¹. Alternatively, $|\Omega_{k,l}\rangle = \frac{1}{\sqrt{d}} \text{vec}(W_{k,l})$,

¹This is also called the standard or “magic” Bell basis due to its special properties given by the Weyl relations (8), and

where the vectorization $\text{vec}(\cdot)$ is defined by its action on the computational basis states of $\mathcal{L}(\mathcal{H})$ by $\text{vec}(|i\rangle\langle j|) = |i\rangle \otimes |j\rangle$ (which extends to general operators $M \in \mathcal{L}(\mathcal{H})$ by linearity). The set of states that are diagonal in this basis, also known as Bell-diagonal states, is defined as

$$\mathcal{M}_d := \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \mid \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right\}, \quad (10)$$

where $P_{k,l} := |\Omega_{k,l}\rangle\langle\Omega_{k,l}|$ are the projectors onto the Bell states (9). Mathematically, equation (10) defines a simplex, which, due to the group structure inherited by the Weyl relations (8), is sometimes referred to as a “magic” simplex [16].

3.1 Group Structure of $\mathcal{M}_d$

Bell-diagonal states $\rho \in \mathcal{M}_d$ possess an underlying group structure² that is known to have a significant impact on their entanglement class [20, 21]. In Sec. 4.1 and 5, we show that this group property is strongly related to entanglement detection via the PPT and realignment criteria for $\rho \in \mathcal{M}_3$. To formalize this, we can arrange the d^2 coefficients $c_{k,l}$ in (10) in a $d \times d$ coefficient matrix $C = (c_{k,l})_{kl}$. For example, for $d = 3$, this matrix is given by

$$C = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} \\ c_{1,0} & c_{1,1} & c_{1,2} \\ c_{2,0} & c_{2,1} & c_{2,2} \end{pmatrix}. \quad (11)$$

Similarly, we can identify the d^2 basis states $P_{k,l}$ in (10) with points in a discrete phase space $\mathbb{Z}_d^2$ (cf. Ref. [25]; see Fig. 2) where the mapping $P_{k,l} \mapsto (W_{i,j} \otimes \mathbb{1}_d) P_{k,l} (W_{i,j}^\dagger \otimes \mathbb{1}_d) = P_{k+i,l+j}$ is an automorphism. In particular, $P_{k,l} = (W_{k,l} \otimes \mathbb{1}_d) P_{0,0} (W_{k,l}^\dagger \otimes \mathbb{1}_d)$, and thus each $P_{k,l} \in \mathcal{M}_d$ can be uniquely associated to one $W_{k,l}$. We define the projective Weyl-Heisenberg group as

$$\mathcal{W}_d := \{W_{k,l}\}_{k,l \in \mathbb{Z}_d} \cong \mathbb{Z}_d^2, \quad (12)$$

where we identify elements that differ only up to a phase³. Here, $\mathbb{Z}_d^2 = \mathbb{Z}_d \times \mathbb{Z}_d$ is the direct product of

²to distinguish it from the generalization considered in Ref. [20]. This generalization, however, lacks many desirable features of the standard Bell basis, e.g., the generating unitaries generally do not form a group.

³An example of the following group-theoretic considerations for $d = 3$ is presented in App. A.

⁴More precisely, let $\mathcal{V}_d := \{\omega^m W_{k,l}\}_{m,k,l \in \mathbb{Z}_d}$ be the Weyl-Heisenberg group [26] with the abelian normalizer $\mathcal{N}_d = \{\omega^m \mathbb{1}_d\}_{m \in \mathbb{Z}_d}$. Then, $\mathcal{W}_d$ is the factor group $\mathcal{V}_d / \mathcal{N}_d$. Thus, there is a slight abuse of notation in (12): The elements of $\mathcal{W}_d$ are not matrices $W_{k,l}$ but rather equivalence classes $[W_{k,l}]$ with $W_{k,l} \sim W_{i,j} \iff W_{k,l} = \omega^m W_{i,j}$ for some $m \in \mathbb{Z}_d$.

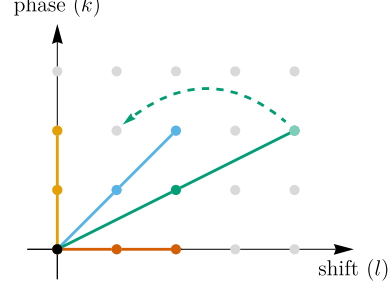


Figure 2: Discrete phase $\mathbb{Z}_3^2$ for a Bell-diagonal qutrit state. The vertices are associated with the Bell states $P_{k,l}$. The colored lines represent the cosets $S(0,0)$ for all $S \in \mathcal{S}_3$, corresponding to the subgroup states ρ_ℓ with $(0,0) \in \ell \in \mathcal{C}_3$. For more details, see App. A.

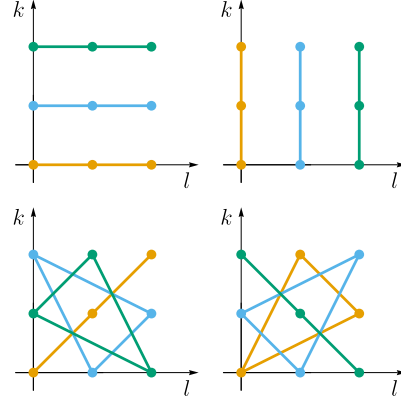


Figure 3: The striations in $\mathbb{Z}_3^2$ corresponding to the cosets $\mathcal{C}(S_{0,1})$ (top left), $\mathcal{C}(S_{1,0})$ (top right), $\mathcal{C}(S_{1,1})$ (bottom left), and $\mathcal{C}(S_{1,2})$ (bottom right), given in (49)-(52) of App. A.

two copies of the cyclic group $\mathbb{Z}_d$. Consequently, the structure of $\mathcal{M}_d$ is inherited from the group structure of $\mathbb{Z}_d^2$, which is the central object of the following considerations.

Let $\mathcal{S}_d$ be the set of d -element subgroups of $\mathbb{Z}_d^2$. Each $S \in \mathcal{S}_d$ gives rise to a line or sublattice in $\mathbb{Z}_d^2$ via the coset $S(i,j) = \{(k,l) + (i,j) \mid (k,l) \in S\}$ for some $(i,j) \in \mathbb{Z}_d^2$ (cf. [16, 25] and Fig. 2). For $S \in \mathcal{S}_d$, we define the set of all cosets by

$$\mathcal{C}(S) := \{S(i,j) \mid (i,j) \in \mathbb{Z}_d^2\}. \quad (13)$$

As each point $(i,j) \in \mathbb{Z}_d^2$ appears exactly once per $\mathcal{C}(S)$, we call $\mathcal{C}(S)$ a “striation” of $\mathbb{Z}_d^2$ generated by $S \in \mathcal{S}_d$ [27] (cf. Fig. 3). The set of all cosets in $\mathbb{Z}_d^2$ is

given by

$$\mathcal{C}_d := \bigcup_{S \in \mathcal{S}_d} \mathcal{C}(S) \quad (14)$$

Bell-diagonal subgroup states $\{\rho_\ell\}_{\ell \in \mathcal{C}_d}$ are highly symmetric separable states in $\mathcal{M}_d$ [16]. For every coset $\ell \in \mathcal{C}_d$, they are defined by having $c_{k,l} = \frac{1}{d}$ for $(k,l) \in \ell$, while all other $c_{k,l}$ are zero, i.e.,

$$\rho_\ell = \frac{1}{d} \sum_{(k,l) \in \ell} P_{k,l} . \quad (15)$$

They form the outermost separable states in $\mathcal{M}_d$, i.e., separable states with highest purity [16], and the convex hull of all subgroup states is known as the (separable) kernel polytope.

3.2 Bloch Vector of $\rho \in \mathcal{M}_d$ and the Discrete Fourier Transform

As the Weyl-Heisenberg operators (7) constitute a basis for the space of d -dimensional density matrices [26], any bipartite qudit state can be written as

$$\rho = \frac{1}{d^2} \sum_{i,j,k,l=0}^{d-1} \beta_{ij,kl} W_{i,j} \otimes W_{k,l} , \quad (16)$$

where $\beta_{ij,kl}$ are the elements defining the generalized Bloch vector (also known as correlation tensor). Using $\beta_{ij,kl} = \text{Tr}(\rho(W_{i,j} \otimes W_{k,l})^\dagger)$ and $\rho \in \mathcal{M}_d$, we obtain

$$\beta_{ij,kl} = \sum_{x,y=0}^{d-1} c_{x,y} \omega^{yi-xj} \delta_{i,-k} \delta_{jl} . \quad (17)$$

With this, we can write [28]

$$\rho = \frac{1}{d^2} \sum_{i,j=0}^{d-1} B_{j,i} W_{i,j} \otimes W_{i,j}^* , \quad (18)$$

where we defined the (reduced) Bloch vector

$$B_{j,i} := \beta_{ij,(-i)j} = \sum_{k,l=0}^{d-1} c_{k,l} \omega^{il-jk} . \quad (19)$$

The form of $\rho \in \mathcal{M}_d$ in (18) can equivalently be derived by applying the Weyl twirl channel to the state (16) (see Ref. [20]).

Note that the Bloch matrix $B = (B_{j,i})_{ji}$ is obtained from the coefficient matrix $C = (c_{k,l})_{kl}$ by a discrete Fourier transform (DFT) (cf. Ref. [28]). In particular, with the unitary DFT matrix

$$F := \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} \omega^{-ij} |i\rangle\langle j| , \quad (20)$$

we get $B = dFCF^\dagger$. Consequently, the Bloch vector B of $\rho \in \mathcal{M}_d$ carries information about the discrete phase space $\mathbb{Z}_d^2$ and the Weyl-Heisenberg group $\mathcal{W}_d$: Coset probabilities in C get mapped to their respective annihilator subgroup cosets via the DFT (cf. Thm. 2.7.1 in Ref. [29]). In other words, while C contains the immediate spatial distribution of ρ in the phase space $\mathbb{Z}_d^2$, the Bloch vector B contains its spectral information.

As preparation for the next section, we conclude here with a short discussion of the properties of B concerning the entanglement class of $\rho \in \mathcal{M}_d$, which is not affected by specific symmetry transformations of C [16]. For example, for prime dimension d , we can understand $\mathbb{Z}_d^2$ as the affine plane $\text{AG}(2, d)$ [30], where these symmetries are given by the affine group $\text{AGL}(2, d)$ (acting as $x \mapsto Ax + b$ with $A \in \text{GL}(2, d)$ and $x, b \in \mathbb{Z}_d^2$). However, as B generally changes under these mappings, the information about the entanglement class contained in B must be invariant under such transformations as well.

4 Realignment Criterion for Bell-Diagonal States

The realignment criterion (Thm. 2.2) can be used to detect entangled quantum states, but it lacks a clear operational and geometric meaning. In this section, we present the realignment criterion for Bell-diagonal qudit-qudit states $\rho \in \mathcal{M}_d$ in a form that admits a clear geometric interpretation in terms of the Bloch vector:

Theorem 4.1. *The realignment criterion for $\rho \in \mathcal{M}_d$ is given by*

$$\|B\|_1 > d \implies \rho \in \text{ENT} , \quad (21)$$

where $B = (B_{j,i})_{ji}$ is the Bloch vector of ρ as defined in (19), and $\|B\|_1 = \sum_{i,j=0}^{d-1} |B_{j,i}|$ denotes the entry-wise 1-norm.

Before we prove the theorem, two remarks are in order. First, note that this theorem reduces the computational complexity of computing the realignment criterion for $\rho \in \mathcal{M}_d$ compared to Thm. 2.2. Particularly, as B is proportional to the DFT of C , evaluating (21) scales only as $\mathcal{O}(d^2 \log(d))$ instead of $\mathcal{O}(d^6)$ for (5) (cf. Ref. [23]). Second, (21) is equivalent to (6) if the correlation tensor is diagonal (as is the case for Bell-diagonal states; cf. (18)), in which case its trace norm reduces to the entry-wise 1-norm. The extra factor d in (21) is due to the unnormalized Weyl matrices satisfying $\text{Tr}(W_{i,j} W_{k,l}^\dagger) = d \delta_{i,k} \delta_{j,l}$.

To prove Thm. 4.1, we need the following lemma:

Lemma 1. *The realigned operator of $\rho \in \mathcal{M}_d$ is given by*

$$\hat{R}(\rho) = \frac{1}{d} \sum_{k,l=0}^{d-1} c_{k,l} W_{k,l} \otimes W_{k,l}^* . \quad (22)$$

Its eigenvectors are $|\Omega_{i,j}\rangle$ for each $(i,j) \in \mathbb{Z}_d^2$ with respective eigenvalues $\frac{1}{d} B_{j,i} = (F C F^\dagger)_{ji} = \frac{1}{d} \sum_{k,l=0}^{d-1} c_{k,l} \omega^{il-jk}$, where F is defined in (20).

Proof. Using (7) and (9), we write $\rho \in \mathcal{M}_d$ as

$$\rho = \frac{1}{d} \sum_{i,j,k,l=0}^{d-1} c_{k,l} \omega^{k(i-j)} |i-l\rangle\langle j-l| \otimes |i\rangle\langle j| . \quad (23)$$

With (4), the realigned operator $\rho_R := \hat{R}(\rho)$ evaluates to

$$\rho_R = \frac{1}{d} \sum_{i,j,k,l=0}^{d-1} c_{k,l} \omega^{k(i-j)} |i-l\rangle\langle i| \otimes |j-l\rangle\langle j| \quad (24)$$

$$= \frac{1}{d} \sum_{k,l=0}^{d-1} c_{k,l} W_{k,l} \otimes W_{k,l}^* . \quad (25)$$

The eigenvectors of $W_{k,l} \otimes W_{k,l}^*$ are the Bell states $|\Omega_{i,j}\rangle$ in (9) with respective eigenvalue ω^{il-jk} [20]. Hence, we obtain

$$\rho_R |\Omega_{i,j}\rangle = \frac{1}{d} \sum_{k,l=0}^{d-1} c_{k,l} \omega^{il-jk} |\Omega_{i,j}\rangle . \quad (26)$$

Proof of Thm. 4.1. Using Lemma 1, we can diagonalize ρ_R via the unitary $U = \sum_{k,l=0}^{d-1} (|k\rangle \otimes |l\rangle) \langle \Omega_{k,l}|$ to obtain $U \rho_R U^\dagger = \frac{1}{d} \sum_{i,j=0}^{d-1} B_{j,i} |i\rangle\langle i| \otimes |j\rangle\langle j|$, and thus

$$\text{Tr} \left(\sqrt{\rho_R^\dagger \rho_R} \right) = \frac{1}{d} \sum_{i,j=0}^{d-1} |B_{j,i}| . \quad (27)$$

Inserting this into Thm. 2.2 completes the proof. $\square$

4.1 Realignment Criterion for Bell-Diagonal Qutrits ($d = 3$)

Thm. 4.1 gives a geometric entanglement condition based on the Bloch vector of $\rho \in \mathcal{M}_d$. Because the Bloch vector B in (19) can be obtained by the DFT of the coefficient matrix C , it encodes the structure of the discrete phase space $\mathbb{Z}_d^2$. In particular, B contains information about the subgroups of $\mathbb{Z}_d^2$.

To exemplify this, we present the explicit version of Thm. 4.1 for Bell-diagonal qutrit states $\rho \in \mathcal{M}_3$. After

some simplifications using the subgroups $S \in \mathcal{S}_3$ and the corresponding cosets $\mathcal{C}(S)$ presented in App. A, (21) can be written as

$$\sum_{S \in \mathcal{S}_3} \sqrt{6 \sum_{\ell \in \mathcal{C}(S)} \left(\sum_{(i,j) \in \ell} c_{i,j} \right)^2} - 2 > 2 \Rightarrow \rho \in \text{ENT} . \quad (28)$$

In words: First, $\sum_{S \in \mathcal{S}_3}$ sums over all subgroups $S \in \mathcal{S}_3$ generating horizontal, vertical, diagonal, and anti-diagonal striations $\mathcal{C}(S)$; second, $\sum_{\ell \in \mathcal{C}(S)}$ sums over all cosets ℓ contained in the striation $\mathcal{C}(S)$; third, $\sum_{(i,j) \in \ell}$ sums the coefficients $c_{i,j}$ of each coset $\ell \in \mathcal{C}(S)$. Consequently, the subgroup structure of $\mathbb{Z}_3^2$ contained in the Bloch vector (19) is crucial for entanglement detection in $\mathcal{M}_3$ using the realignment criterion.

Furthermore, (28) entails a measurement procedure to obtain the realignment value of $\rho \in \mathcal{M}_3$. This requires projecting the state for each $S \in \mathcal{S}_3$ onto the different coset subspaces using the projection operators in $\Pi(S) = \{ \sum_{(i,j) \in \ell} P_{i,j} \mid \ell \in \mathcal{C}(S) \}$. In this way, each subgroup gives rise to a positive operator-valued measure (POVM), corresponding to the striation $\mathcal{C}(S)$.

5 PPT Criterion for Bell-Diagonal Qutrits ($d = 3$)

For bipartite qubit states $\rho \in \mathcal{D}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, the PPT criterion (Thm. 2.1) is necessary and sufficient for entanglement [9], and it even holds that [31]

$$\rho \in \text{ENT} \iff \rho^\Gamma \not\geq 0 \iff \det(\rho^\Gamma) < 0 . \quad (29)$$

For general qudit-qudit states $\rho \in \mathcal{D}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ with $d_A d_B > 6$, these equivalences fail and we get the weaker statement

$$\rho \in \text{ENT} \Leftarrow \rho^\Gamma \not\geq 0 \Leftarrow \det(\rho^\Gamma) < 0 . \quad (30)$$

In this section, we show that for Bell-diagonal qutrit-qutrit states $\rho \in \mathcal{M}_3$, the equivalence for the second implication in (30) is restored, namely, $\rho^\Gamma \not\geq 0 \iff \det(\rho^\Gamma) < 0$. In contrast, even for Bell-diagonal qutrit states, the first implication cannot be promoted to an equivalence because of the existence of PPT-entangled states in $\mathcal{M}_3$, i.e., $\text{ENT} \cap \text{PPT} \cap \mathcal{M}_3 \neq \emptyset$. Furthermore, our result underscores that the subgroup structure of $\mathbb{Z}_3^2$ plays a crucial role in distinguishing between $\text{PPT} \cap \mathcal{M}_3$ and $\text{NPT} \cap \mathcal{M}_3$.

Theorem 5.1. *Let $\rho \in \mathcal{M}_3$ and let ρ^Γ be the partial transpose of ρ . It holds that*

$$\rho^\Gamma \not\geq 0 \iff \det(\rho^\Gamma) < 0 . \quad (31)$$

Explicitly, we have

$$\rho^\Gamma \not\geq 0 \iff 3 \sum_{\ell \in \mathcal{C}_3} \left(\prod_{(i,j) \in \ell} c_{i,j} \right) < \sum_{k,l=0}^2 c_{k,l}^3. \quad (32)$$

The idea of the proof is to reduce the spectrum of the partial transposition of $\rho \in \mathcal{M}_3$ to the spectrum of a 3×3 matrix A_0 with only one non-contradicting assignment of the signs of its eigenvalues if $\rho^\Gamma \not\geq 0$.

Proof. We use (23) to compute the partial transposition $\rho^\Gamma := (\mathbb{1} \otimes T)(\rho)$ and apply the basis transformation $U = \sum_{i,j=0}^2 (|i\rangle \otimes |j\rangle) \langle \Omega_{i,j}|$, finding that ρ^Γ is block-diagonal in the Bell basis, i.e.,

$$U \rho^\Gamma U^\dagger = \sum_{m=0}^2 |m\rangle \langle m| \otimes A_m, \quad (33)$$

where we defined

$$A_m := \frac{1}{9} \sum_{i,j,k,l=0}^2 \omega^{j(m-k)-i(l+j)} c_{k,l} W_{i,j}. \quad (34)$$

The hermitian blocks A_m are known to be unitarily equivalent for odd dimension d [16], and we have $A_m = W_{m,0}^\dagger A_0 W_{m,0}$ for $m \in \{0, 1, 2\}$. Hence, each A_m has the same eigenvalues and thus the spectrum of A_0 completely determines the spectrum of ρ^Γ . Also, the multiplicity of each eigenvalue of ρ^Γ must be divisible by 3. Furthermore, $\text{Tr}(A_0) = \frac{1}{3} \sum_{k,l=0}^{d-1} c_{k,l} = \frac{1}{3}$ and, because $\det(\rho^\Gamma) = \prod_{m=0}^2 \det(A_m) = \det(A_0)^3$, we have

$$\det(\rho^\Gamma) < 0 \iff \det(A_0) < 0. \quad (35)$$

The spectrum of A_0 is restricted as follows: Let us denote the minimal and maximal eigenvalues of a hermitian matrix M by $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$, respectively. The qutrit Bell basis projectors $P_{k,l} = |\Omega_{k,l}\rangle \langle \Omega_{k,l}| \in \mathcal{M}_3$ satisfy $\lambda_{\min}(P_{k,l}^\Gamma) = -\frac{1}{3}$ and $\lambda_{\max}(P_{k,l}^\Gamma) = \frac{1}{3}$ for all $k, l \in \{0, 1, 2\}$. Applying Weyl's inequalities [32] to ρ^Γ for $\rho \in \mathcal{M}_3$ yields

$$\lambda_{\min}(\rho^\Gamma) = \lambda_{\min} \left(\sum_{k,l=0}^2 c_{k,l} P_{k,l}^\Gamma \right) \quad (36)$$

$$\geq \sum_{k,l=0}^2 c_{k,l} \lambda_{\min}(P_{k,l}^\Gamma) = -\frac{1}{3}, \quad (37)$$

and similarly $\lambda_{\max}(\rho^\Gamma) \leq \frac{1}{3}$, i.e., the spectrum of ρ^Γ and A_0 is contained in $[-\frac{1}{3}, \frac{1}{3}]$.

Now, let us show that only $\det(A_0) < 0$ (and thus $\det(\rho^\Gamma) < 0$) is consistent with $\rho^\Gamma \not\geq 0$. The partial

transpose of any $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\dim(\mathcal{H}_A) = d_A$ and $\dim(\mathcal{H}_B) = d_B$ has at most $n_- = (d_A - 1)(d_B - 1)$ negative eigenvalues [33, 34]. For $\rho \in \mathcal{M}_3$ we have $d_A = d_B = 3$ and thus $n_- \leq 4$. By the multiplicity restriction on the eigenvalues of ρ^Γ , we conclude that either $n_- = 0$ or $n_- = 3$. The case $n_- = 0$ is equivalent to $\rho^\Gamma \geq 0$ and $\det(\rho^\Gamma) \geq 0$. For $n_- = 3$ we have $\rho^\Gamma \not\geq 0$, and A_0 has exactly one negative eigenvalue $\lambda_1 < 0$. This leaves three distinct cases for the other two non-negative eigenvalues λ_2 and λ_3 of A_0 . First, if $\lambda_2 = \lambda_3 = 0$, we obtain $\text{Tr}(A_0) = \lambda_1 < 0$, which contradicts $\text{Tr}(A_0) = \frac{1}{3}$. Second, if $\lambda_2 > \lambda_3 = 0$ we have $\frac{1}{3} = \text{Tr}(A_0) = \lambda_1 + \lambda_2$, and thus $\lambda_2 = \frac{1}{3} - \lambda_1 > \frac{1}{3}$. But this contradicts that the spectrum of A_0 is contained in $[-\frac{1}{3}, \frac{1}{3}]$. The third and only consistent case for $\rho^\Gamma \not\geq 0$ is $\lambda_1 < 0 < \lambda_2, \lambda_3$ with $\det(A_0) = \lambda_1 \lambda_2 \lambda_3 < 0$. Thus, we have shown that $\rho^\Gamma \not\geq 0 \implies \det(A_0) < 0$. Together with (35) and the (trivial) implication $\det(\rho^\Gamma) < 0 \implies \rho^\Gamma \not\geq 0$, we get the equivalence (31). Finally, the explicit formula (32) is obtained by computing $\det(A_0)$ using (34). $\square$

Thm. 5.1 has interesting implications regarding the geometric properties of the sets $\text{PPT} \cap \mathcal{M}_3$ and $\text{NPT} \cap \mathcal{M}_3$. To see this, let us define the boundary $\partial(\text{PPT})$ of PPT as states $\rho \in \text{PPT} \cap \mathcal{M}_3$ such that every ϵ -neighborhood around ρ is neither contained in PPT nor NPT. As $\det(\rho^\Gamma)$ is a continuous function of the coefficients $c_{k,l}$, boundary states $\rho \in \partial(\text{PPT})$ satisfy $\det(\rho^\Gamma) = 0$. Using (32), this is equivalent to

$$\rho \in \partial(\text{PPT}) \implies 3 \sum_{\ell \in \mathcal{C}_3} \left(\prod_{(i,j) \in \ell} c_{i,j} \right) = \sum_{k,l=0}^2 c_{k,l}^3. \quad (38)$$

Simple examples include the separable subgroup states $\{\rho_\ell\}_{\ell \in \mathcal{C}_3}$ (cf. Sec. 3). We can reformulate this as a complete geometric characterization of the set $\text{PPT} \cap \mathcal{M}_3$ by tangential hyperplanes, i.e., entanglement witnesses, for which we define

$$W_{\text{NPT}} = \sum_{i,j=0}^2 \kappa_{i,j} P_{i,j}, \quad (39)$$

$$\kappa_{i,j} = -c_{i,j}^2 + \sum_{\{(i,j),(k,l),(m,n)\} \in \mathcal{C}_3} c_{k,l} c_{m,n}. \quad (40)$$

The explicit form of $\kappa_{i,j}$ is presented in App. C. If the coefficients $c_{k,l}$ in ρ and $\kappa_{i,j}$ coincide, it holds that $\rho^\Gamma \not\geq 0 \iff \text{Tr}(\rho W_{\text{NPT}}) < 0$ and $\rho \in \partial(\text{PPT}) \implies \text{Tr}(\rho W_{\text{NPT}}) = 0$. In this regard, W_{NPT} can be considered a ‘‘PPT-optimal’’ entanglement witness. We note, however, that W_{NPT} can be the zero matrix if all $\kappa_{i,j}$ vanish, as is the case for all subgroup states

$\{\rho_\ell\}_{\ell \in \mathcal{C}_3}$; or it can become a positive operator, as is the case for the maximally mixed state with $c_{k,l} = \frac{1}{9}$ for all $k, l \in \{0, 1, 2\}$. Strictly speaking, W_{NPT} does not define an entanglement witness in those cases.

The Weyl-covariant [35] positive map Λ_{NPT} corresponding to W_{NPT} via the Choi-Jamiołkowski isomorphism [36, 37] is given by

$$\Lambda_{\text{NPT}}(\rho) = 3 \text{Tr}_B (W_{\text{NPT}} (\mathbb{1}_A \otimes \rho^T)) \quad (41)$$

$$= \sum_{i,j=0}^2 \kappa_{i,j} W_{i,j} \rho W_{i,j}^\dagger, \quad (42)$$

where $\text{Tr}_B(\cdot)$ denotes the partial trace over the second subsystem.

Lastly, using Thm. 5.1 and the fact that subgroup states are separable [16], we arrive at the following statement (the proof is presented in App. B):

Proposition 5.1. *There is no PPT-entangled $\rho \in \mathcal{M}_3$ with coefficients $c_{k,l}$ corresponding to at least one coset $\ell \in \mathcal{C}_3$ equal to zero.*

This serves as a good guiding principle for finding PPT-entangled states in $\mathcal{M}_3$. For example, the Bell-diagonal states that are defined by the coefficient matrix

$$C = \begin{pmatrix} \frac{1}{3} - x & \frac{1}{3} - x & \frac{1}{3} - x \\ 2x & x & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (43)$$

with $x \in [0, 1/3]$, satisfy the premise of the proposition and are separable for $x = 0$ (reducing to case 3b in App. B) and NPT-entangled for $x \in (0, 1/3]$ (cf. case 5 in App. B). On the contrary, Ref. [38] considered allegedly similar-looking states that are (up to the symmetries of $\mathcal{M}_3$; cf. Sec. 3.2) equivalent to

$$C = \begin{pmatrix} \frac{1}{3} - x & \frac{1}{3} - x & \frac{1}{3} - x \\ 2x & 0 & 0 \\ x & 0 & 0 \end{pmatrix}. \quad (44)$$

These states do not satisfy the requirement of the proposition as no coset $\ell \in \mathcal{C}_3$ has coefficients $c_{k,l}$ equal to zero. Using (28) and (32), it can be checked that this state is PPT-entangled for $x \in (0, \frac{2}{15}]$.

6 Conclusion and Outlook

The goal of this work was to answer the question of how the group structure of Bell-diagonal states relates to the PPT and realignment criteria. To this end, we leveraged the symmetry properties of this state family to provide a unified approach connecting these criteria to the group-theoretic features of the system.

Thm. 4.1 shows that the realignment criterion for Bell-diagonal states is strongly linked to the geometry of the Bloch ball. As the Bloch vector of Bell-diagonal states is obtained by the discrete Fourier transform of their defining coefficient matrix, it contains information about the subgroups of $\mathbb{Z}_d^2$. For the specific case of Bell-diagonal qutrits, we find that in the 8-dimensional coefficient space of $c_{k,l}$, the inequality for the realignment criterion is at most quadratic in $c_{k,l}$. In contrast, the PPT set is defined by a cubic inequality by Thm. 5.1. Additionally, this explicit version of the PPT criterion for Bell-diagonal qutrit states $\rho \in \mathcal{M}_3$ unveils that the equivalence $\det(\rho^\Gamma) < 0 \iff \rho^\Gamma \not\geq 0$ holds, which is generally not valid for all states of the same dimension. Consequently, our results not only demonstrate that the group structure of Bell-diagonal states simplifies the analysis of two specific entanglement detection criteria but also provide deeper insights into the mathematical and physical properties of this state family.

While our analysis focuses on Bell-diagonal states and is limited to $d = 3$ for the PPT criterion, adapting this framework to other dimensions or more general classes of quantum states remains an open challenge. Extending our treatment of the PPT criterion for Bell-diagonal qudits for $d > 3$ is not straightforward, as the spectral properties of $U \rho^\Gamma U^\dagger$ do not satisfy the requirements used in the proof of Thm. 5.1.

Going forward, it is furthermore an open problem whether for general bipartite qudit states, an entanglement criterion based on the entry-wise 1-norm of the Bloch vector, similar to Thm. 4.1, can be derived. Ref. [39, 40] present a sufficient criterion for separability, reading $\|\beta\|_1 \leq 2 \implies \rho \in \text{SEP}$ for β as in (16), but to the authors' knowledge, no corresponding sufficient entanglement criterion exists. A naive approximation using $\|\beta\|_1 \leq d^2 \|\beta\|_2$ for $\beta \in \mathbb{C}^{d^4}$ yields that ρ is entangled if $\|\beta\|_1 > d^3$. However, the maximal value for pure states $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ is also $\|\beta\|_1 = d^3$. A quick numerical optimization for $d = 3$ yields $\|\beta\|_1 \lesssim 25.9735$ for general pure states $|\psi\rangle \in \mathbb{C}^3 \otimes \mathbb{C}^3$, whereas separable pure states seem to satisfy $\|\beta\|_1 \leq 25$. The discrepancy of $25.9735 < 27 = 3^3$ stems from the fact that equality in $\|\beta\|_1 \leq d^2 \|\beta\|_2$ holds if and only if β and the vector $(1, \dots, 1)^T$ are linearly dependent. However, this cannot be as $\frac{1}{d^2} \sum_{i,j,k,l=0}^{d-1} W_{i,j} \otimes W_{k,l}$ does not represent a valid quantum state.

To conclude, by bridging the gap between group theory and entanglement detection, this work provides a novel perspective on the structure of quantum states, specifically for qutrit-qutrit systems, paving the way for further advancements in quantum information science.

Acknowledgments

B.C.H. and C.P. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) project P36102-N (Grant DOI: 10.55776/P36102). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

A Group Structure of $\mathcal{M}_3$

The 3-dimensional projective Weyl-Heisenberg group is given by $\mathcal{W}_3 = \{W_{k,l}\}_{k,l \in \mathbb{Z}_3} \cong \mathbb{Z}_3^2$. As $d = 3$ is prime, all subgroups $S_{k,l} \in \mathcal{S}_3$ of $\mathbb{Z}_3^2$ have d elements and are cyclic, i.e., $S_{k,l} = \{n \cdot (k,l)\}_{n \in \mathbb{Z}_3}$ for some $(k,l) \in \mathbb{Z}_3^2$ with $(k,l) \neq (0,0)$. In particular, they are given by

$$S_{0,1} = \{(0,0), (0,1), (0,2)\}, \quad (45)$$

$$S_{1,0} = \{(0,0), (1,0), (2,0)\}, \quad (46)$$

$$S_{1,1} = \{(0,0), (1,1), (2,2)\}, \quad (47)$$

$$S_{1,2} = \{(0,0), (1,2), (2,1)\}, \quad (48)$$

with the corresponding cosets $\mathcal{C}(S_{k,l})$

$$\begin{aligned} \mathcal{C}(S_{0,1}) = \{ & \{(0,0), (0,1), (0,2)\}, \\ & \{(1,0), (1,1), (1,2)\}, \\ & \{(2,0), (2,1), (2,2)\} \}, \end{aligned} \quad (49)$$

$$\begin{aligned} \mathcal{C}(S_{1,0}) = \{ & \{(0,0), (1,0), (2,0)\}, \\ & \{(0,1), (1,1), (2,1)\}, \\ & \{(0,2), (1,2), (2,2)\} \}, \end{aligned} \quad (50)$$

$$\begin{aligned} \mathcal{C}(S_{1,1}) = \{ & \{(0,0), (1,1), (2,2)\}, \\ & \{(0,1), (1,2), (2,0)\}, \\ & \{(0,2), (1,0), (2,1)\} \}, \end{aligned} \quad (51)$$

$$\begin{aligned} \mathcal{C}(S_{1,2}) = \{ & \{(0,0), (1,2), (2,1)\}, \\ & \{(0,1), (1,0), (2,2)\}, \\ & \{(0,2), (1,1), (2,0)\} \}. \end{aligned} \quad (52)$$

Note that each element $(i,j) \in \mathbb{Z}_3^2$ appears exactly once in each $\mathcal{C}(S_{k,l})$. Fig. 3 shows the horizontal, vertical, diagonal, and anti-diagonal striations corresponding to $\mathcal{C}(S_{0,1})$, $\mathcal{C}(S_{1,0})$, $\mathcal{C}(S_{1,1})$, and $\mathcal{C}(S_{1,2})$, respectively. The set of all cosets $\mathcal{C}_3$ containing all possible lines in $\mathbb{Z}_3^2$ is defined in (14) as the union of (49)-(52). Using (15), the separable subgroup states in $\mathcal{M}_3$ corresponding to the striation $\mathcal{C}(S_{0,1})$ (similarly

for $S_{1,0}$, $S_{1,1}$, and $S_{1,2}$) are

$$\rho_{0,1}^{(1)} = \frac{1}{3}(P_{0,0} + P_{0,1} + P_{0,2}), \quad (53)$$

$$\rho_{0,1}^{(2)} = \frac{1}{3}(P_{1,0} + P_{1,1} + P_{1,2}), \quad (54)$$

$$\rho_{0,1}^{(3)} = \frac{1}{3}(P_{2,0} + P_{2,1} + P_{2,2}). \quad (55)$$

The projectors onto the subgroup subspaces are $\Pi(S_{k,l}) = \{3\rho_{k,l}^{(1)}, 3\rho_{k,l}^{(2)}, 3\rho_{k,l}^{(3)}\}$, which constitutes a POVM measurement for each $S_{k,l}$.

B Proof of Proposition 5.1

The subgroups of $\mathbb{Z}_3^2$ are cyclic and correspond to vertical, horizontal, diagonal, and anti-diagonal lines in the affine plane $\mathbb{Z}_3^2$ (cf. Fig. 3), or equivalently in the coefficient matrix C in (11). Hence, we need to show that if the coefficients $c_{k,l}$ (satisfying $c_{k,l} \geq 0$ and $\sum_{k,l=0}^2 c_{k,l} = 1$) corresponding to any coset $\ell \in \mathcal{C}_3$ are zero, the corresponding state $\rho \in \mathcal{M}_3$ is either separable or NPT entangled. To check the latter, we can use Thm. 5.1. Due to the symmetries of $\mathcal{M}_3$ that preserve the entanglement class of $\rho \in \mathcal{M}_3$ [16, 25], this reduces to six distinct cases. There is a total of $432 = |\text{AGL}(2,3)|$ such symmetry transformations that map cosets to cosets (cf. Sec. 3.2). In the following, each case beyond the first is obtained from the former by adding one additional $c_{k,l} \neq 0$ and accounting for these symmetries.

Case 1: If only one $c_{k,l} > 0$, it is necessarily equal to 1 (as $\sum_{k,l=0}^2 c_{k,l} = 1$). Utilizing the symmetries of $\mathcal{M}_3$, we can shift any one $c_{k,l}$ to $c_{0,0}$ by local operations. Hence, this case reduces to

$$C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (56)$$

for which (32) yields $0 < 1$, and thus $\rho \in \text{NPT}$.

Case 2: If two $c_{k,l} > 0$, we can locally shift one coefficient $c_{k,l}$ to $c_{0,0}$ and the other one to $c_{0,1}$, so this case reduces to

$$C_2 = \begin{pmatrix} c_{0,0} & c_{0,1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (57)$$

Equation (32) becomes $0 < c_{0,0}^3 + c_{0,1}^3$, which is satisfied. Hence, $\rho \in \text{NPT}$.

Case 3: Three coefficients $c_{k,l} > 0$ reduce to two distinct cases after utilizing the symmetries of $\mathcal{M}_3$. The first one is given by

$$C_{3a} = \begin{pmatrix} c_{0,0} & c_{0,1} & 0 \\ c_{1,0} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (58)$$

with (32) being $0 < c_{0,0}^3 + c_{0,1}^3 + c_{1,0}^3$, and thus $\rho \in \text{NPT}$. The second one is

$$C_{3b} = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (59)$$

for which (32) becomes (after some simplifications)

$$-|z|^2 \underbrace{(c_{0,0} + c_{0,1} + c_{0,2})}_{=1} < 0, \quad (60)$$

where $z := c_{0,0} + \omega c_{0,1} + \omega^* c_{0,2}$ and $\omega = e^{2\pi i/3}$. The inequality is violated if and only if $|z| = 0$, which is equivalent to $c_{0,0} = c_{0,1} = c_{0,2} = 1/3$ (because $\sum_{k,l=0}^{d-1} c_{k,l} = 1$), and corresponds to a separable subgroup state [16]. Otherwise, the inequality is satisfied and $\rho \in \text{NPT}$.

Case 4: With four coefficients $c_{k,l} > 0$, we again distinguish two cases. The first one is

$$C_{4a} = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} \\ c_{1,0} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (61)$$

Equation (32) is given by

$$\underbrace{3c_{0,0}c_{0,1}c_{0,2} - c_{0,0}^3 - c_{0,1}^3 - c_{0,2}^3}_{\leq 0 \text{ by (60)}} - c_{1,0}^3 < 0, \quad (62)$$

and thus $\rho \in \text{NPT}$. The second case is

$$C_{4b} = \begin{pmatrix} c_{0,0} & c_{0,1} & 0 \\ c_{1,0} & c_{1,1} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (63)$$

yielding $0 < c_{0,0}^3 + c_{0,1}^3 + c_{1,0}^3 + c_{1,1}^3$, and thus $\rho \in \text{NPT}$.

Case 5: The case with five $c_{k,l} > 0$ is obtained by adding one extra $c_{k,l}$ to either (61) or (63). After accounting for the symmetries in $\mathcal{M}_3$, the only possibility with one complete line of zeros is

$$C_5 = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} \\ c_{1,0} & c_{1,1} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (64)$$

This gives

$$\underbrace{3c_{0,0}c_{0,1}c_{0,2} - c_{0,0}^3 - c_{0,1}^3 - c_{0,2}^3}_{\leq 0 \text{ by (60)}} - c_{1,0}^3 - c_{1,1}^3 < 0, \quad (65)$$

and thus $\rho \in \text{NPT}$.

Case 6: Accounting for the symmetries of $\mathcal{M}_3$, the only possibility to have six $c_{k,l} > 0$ but still one full line equal to zero is given by

$$C_6 = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} \\ c_{1,0} & c_{1,1} & c_{1,2} \\ 0 & 0 & 0 \end{pmatrix}, \quad (66)$$

for which (32) is

$$\underbrace{3c_{0,0}c_{0,1}c_{0,2} - c_{0,0}^3 - c_{0,1}^3 - c_{0,2}^3}_{\leq 0 \text{ by (60)}} + \underbrace{3c_{1,0}c_{1,1}c_{1,2} - c_{1,0}^3 - c_{1,1}^3 - c_{1,2}^3}_{\leq 0 \text{ by (60)}} < 0. \quad (67)$$

Thus, the inequality is violated if and only if $c_{0,0} = c_{0,1} = c_{0,2}$ and $c_{1,0} = c_{1,1} = c_{1,2}$, which corresponds to a convex mixture of two separable subgroup states (15), and thus is also separable. Otherwise, it is satisfied and $\rho \in \text{NPT}$.

C Explicit form of $\kappa_{i,j}$ for W_{NPT}

Here, we give the explicit form of $\kappa_{i,j}$ in (40) for the entanglement witness (39):

$$\begin{aligned} \kappa_{0,0} &= -c_{0,0}^2 + c_{0,1}c_{0,2} + c_{1,0}c_{2,0} + c_{1,1}c_{2,2} + c_{1,2}c_{2,1}, \\ \kappa_{0,1} &= -c_{0,1}^2 + c_{0,0}c_{0,2} + c_{1,1}c_{2,1} + c_{1,2}c_{2,0} + c_{1,0}c_{2,2}, \\ \kappa_{0,2} &= -c_{0,2}^2 + c_{0,0}c_{0,1} + c_{1,2}c_{2,2} + c_{1,0}c_{2,1} + c_{1,1}c_{2,0}, \\ \kappa_{1,0} &= -c_{1,0}^2 + c_{1,1}c_{1,2} + c_{0,0}c_{2,0} + c_{0,2}c_{2,1} + c_{0,1}c_{2,2}, \\ \kappa_{1,1} &= -c_{1,1}^2 + c_{1,0}c_{1,2} + c_{0,1}c_{2,1} + c_{0,0}c_{2,2} + c_{0,2}c_{2,0}, \\ \kappa_{1,2} &= -c_{1,2}^2 + c_{1,0}c_{1,1} + c_{0,2}c_{2,2} + c_{0,1}c_{2,0} + c_{0,0}c_{2,1}, \\ \kappa_{2,0} &= -c_{2,0}^2 + c_{2,1}c_{2,2} + c_{0,0}c_{1,0} + c_{0,1}c_{1,2} + c_{0,2}c_{1,1}, \\ \kappa_{2,1} &= -c_{2,1}^2 + c_{2,0}c_{2,2} + c_{0,1}c_{1,1} + c_{0,2}c_{1,0} + c_{0,0}c_{1,2}, \\ \kappa_{2,2} &= -c_{2,2}^2 + c_{2,0}c_{2,1} + c_{0,2}c_{1,2} + c_{0,0}c_{1,1} + c_{0,1}c_{1,0}. \end{aligned}$$

Here, the second, third, fourth, and fifth term in each $\kappa_{i,j}$ corresponds to the horizontal, vertical, diagonal, and anti-diagonal cosets containing $c_{i,j}$, respectively.

References

- ¹J. Preskill, “Quantum Computing in the NISQ era and beyond”, en-GB, [Quantum](#) **2**, Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 79 (2018).
- ²V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, “The security of practical quantum key distribution”, [Reviews of Modern Physics](#) **81**, Publisher: American Physical Society, 1301–1350 (2009).
- ³C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”, [Physical Review Letters](#) **70**, Publisher: American Physical Society, 1895–1899 (1993).

- ⁴R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement”, *Reviews of Modern Physics* **81**, Publisher: American Physical Society, 865–942 (2009).
- ⁵L. Gurvits, “Classical deterministic complexity of Edmonds’ Problem and quantum entanglement”, en, in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing* (June 2003), pp. 10–19.
- ⁶S. Gharibian, “Strong NP-Hardness of the Quantum Separability Problem”, en, *Quantum Information & Computation* **10**, 343–360 (2010).
- ⁷C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, “Concentrating partial entanglement by local operations”, *Physical Review A* **53**, Publisher: American Physical Society, 2046–2052 (1996).
- ⁸A. Peres, “Separability Criterion for Density Matrices”, en, *Physical Review Letters* **77**, Publisher: American Physical Society (APS), 1413–1415 (1996).
- ⁹M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions”, en, *Physics Letters A* **223**, Publisher: Elsevier BV, 1–8 (1996).
- ¹⁰P. Horodecki, “Separability criterion and inseparable mixed states with positive partial transposition”, *Physics Letters A* **232**, 333–339 (1997).
- ¹¹M. Horodecki, P. Horodecki, and R. Horodecki, “Mixed-State Entanglement and Distillation: Is there a ‘Bound’ Entanglement in Nature?”, en, *Physical Review Letters* **80**, Publisher: American Physical Society (APS), 5239–5242 (1998).
- ¹²B. C. Hiesmayr, C. Popp, and T. C. Sutter, “Bipartite bound entanglement”, *International Journal of Quantum Information* **23**, Publisher: World Scientific Publishing Co., 2530003 (2025).
- ¹³K. Chen and L. A. Wu, “A matrix realignment method for recognizing entanglement”, *Quant. Inf. Comput.* **3**, 193–202 (2003).
- ¹⁴O. Rudolph, “Further Results on the Cross Norm Criterion for Separability”, en, *Quantum Information Processing* **4**, Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 3 Publisher: Kluwer Academic Publishers-Plenum Publishers, 219–239 (2005).
- ¹⁵K. S. Ranade and G. Alber, “Asymptotic correctability of Bell-diagonal qudit states and lower bounds on tolerable error probabilities in quantum cryptography”, en, *Journal of Physics A: Mathematical and Theoretical* **40**, 139 (2006).
- ¹⁶B. Baumgartner, B. Hiesmayr, and H. Narnhofer, “A special simplex in the state space for entangled qudits”, en, *Journal of Physics A: Mathematical and Theoretical* **40**, Publisher: IOP Publishing, 7919–7938 (2007).
- ¹⁷C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, “Mixed-state entanglement and quantum error correction”, *Physical Review A* **54**, Publisher: American Physical Society, 3824–3851 (1996).
- ¹⁸C. Popp, T. C. Sutter, and B. C. Hiesmayr, *A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits*, arXiv:2408.02383 [quant-ph], May 2025.
- ¹⁹C. Popp and B. C. Hiesmayr, “Almost complete solution for the NP-hard separability problem of Bell diagonal qutrits”, en, *Scientific Reports* **12**, 12472 (2022).
- ²⁰C. Popp and B. C. Hiesmayr, “Special features of the Weyl–Heisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states”, en, *New Journal of Physics* **26**, 013039 (2024).
- ²¹C. Popp and B. C. Hiesmayr, “Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts”, en, *Scientific Reports* **13**, Publisher: Springer Science and Business Media LLC, 10.1038/s41598-023-29211-w (2023).
- ²²C. Popp, “BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits”, en, *Journal of Open Source Software* **8**, 4924 (2023).
- ²³G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, 2013).
- ²⁴G. Sarbicki, G. Scala, and D. Chruściński, “Family of multipartite separability criteria based on a correlation tensor”, *Physical Review A* **101**, Publisher: American Physical Society, 012341 (2020).
- ²⁵B. Baumgartner, B. C. Hiesmayr, and H. Narnhofer, “State space for two qutrits has a phase space structure in its core”, en, *Physical Review A* **74**, Publisher: American Physical Society (APS), 10.1103/physreva.74.032327 (2006).
- ²⁶I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
- ²⁷K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, “Discrete phase space based on finite fields”, en, *Physical Review A* **70**, Publisher: American Physical Society (APS), 10.1103/physreva.70.062101 (2004).

- ²⁸H. Narnhofer, “Entanglement reflected in Wigner functions”, en, [Journal of Physics A: Mathematical and General](#) **39**, 7051 (2006).
- ²⁹W. Rudin, *Fourier analysis on groups*, eng (Interscience Publishers, New York, 1962).
- ³⁰T. Beth, D. Jungnickel, and H. Lenz, *Design Theory* (Cambridge University Press, 1999).
- ³¹R. Augusiak, M. Demianowicz, and P. Horodecki, “Universal observable detecting all two-qubit entanglement and determinant-based separability tests”, en, [Physical Review A](#) **77**, Publisher: American Physical Society (APS), 10.1103/physreva.77.030301 (2008).
- ³²R. Bhatia, *Matrix analysis*, en, GTM (Springer, New York, 1997).
- ³³S. Rana, “Negative eigenvalues of partial transposition of arbitrary bipartite states”, en, [Physical Review A](#) **87**, Publisher: American Physical Society (APS), 10.1103/physreva.87.054301 (2013).
- ³⁴N. Johnston, “Non-positive-partial-transpose subspaces can be as large as any entangled subspace”, en, [Physical Review A](#) **87**, Publisher: American Physical Society (APS), 10.1103/physreva.87.064302 (2013).
- ³⁵J. Watrous, *The Theory of Quantum Information*, en, 1st ed. (Cambridge University Press, Apr. 2018).
- ³⁶M.-D. Choi, “Completely positive linear maps on complex matrices”, en, [Linear Algebra and its Applications](#) **10**, Publisher: Elsevier BV, 285–290 (1975).
- ³⁷A. Jamiolkowski, “Linear transformations which preserve trace and positive semidefiniteness of operators”, [Reports on Mathematical Physics](#) **3**, 275–278 (1972).
- ³⁸P. Caban and B. C. Hiesmayr, “Bound entanglement is not Lorentz invariant”, en, [Scientific Reports](#) **13**, Publisher: Nature Publishing Group, 11189 (2023).
- ³⁹A. O. Pittenger and M. H. Rubin, “Convexity and the separability problem of quantum mechanical density matrices”, [Linear Algebra and its Applications](#) **346**, 47–71 (2002).
- ⁴⁰D. Chruściński and A. O. Pittenger, “Generalized circulant densities and a sufficient condition for separability”, en, [Journal of Physics A: Mathematical and Theoretical](#) **41**, 385301 (2008).

B.6. A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits

A Novel Stabilizer-based Entanglement Distillation Protocol for Qudits

Christopher Popp, Tobias C. Sutter, and Beatrix C. Hiesmayr

University of Vienna, Faculty of Physics, Währingerstrasse 17, 1090 Vienna.

Entanglement distillation, the process of converting weakly entangled states into maximally entangled ones using Local Operations and Classical Communication (LOCC), is pivotal for robust entanglement-assisted quantum information processing in error-prone environments. A construction based on stabilizer codes offers an effective method for designing such protocols. By analytically investigating the effective action of stabilizer protocols for systems of prime dimension d , we establish a standard form for the output states of recurrent stabilizer-based distillation. This links the properties of input states, stabilizers, and encodings to the properties of the protocol. Based on those insights, we present a novel two-copy distillation protocol, applicable to all bipartite states in prime dimension, that maximizes the fidelity increase per iteration for Bell-diagonal states. The power of this framework and the protocol is demonstrated through numerical investigations, which provide evidence for superior performance in terms of efficiency and distillability of low-fidelity states compared to other well-established recurrence protocols. By elucidating the interplay between states, errors, and protocols, our contribution advances the systematic development of highly effective distillation protocols, enhancing our understanding of distillability.

1 Introduction

Quantum information science offers exciting potentials to quantum technology like superior computational power, secure communication and improved sensing by leveraging non-classical resources. One of the most valuable and intriguing quantum resources is entanglement. Prominent applications using this resource include quantum teleportation [1], quantum dense coding [2, 3], or measurement-based quantum computing [4]. An important class of systems, especially in the context of quantum communication, are shared systems between two parties, named Alice and Bob. Assuming that each party possesses d -level systems, called qudits, entanglement is shared in the form of joint quantum states that cannot be locally described by the individual systems, but only if the combined global system is considered. The gold standard of this resource are two-qudit maximally entangled states, so-called Bell states. Due to interactions with the environment, this resource is generally not available in its pure form, but affected by noise. Depending on the level of noise, the entanglement of the shared pair is effectively reduced or even destroyed. One method to deal with this problem is entanglement distillation. The objective for Alice and Bob is to use local operations on their individual systems and

Christopher Popp: christopher.popp@univie.ac.at

Tobias C. Sutter: tobias.christoph.sutter@univie.ac.at

Beatrix C. Hiesmayr: beatrix.hiesmayr@univie.ac.at

classical communication (LOCC) to transform several pairs of noisy and therefore weakly entangled states to a smaller number of strongly entangled states. Not all entangled states can be distilled via LOCC, due to the existence of bound entanglement of states that are positive under partial transposition and the potential existence of undistillable states with negative partial transposition [5, 6]. Entanglement distillation (sometimes also called purification) was first introduced for bipartite two-level systems, i.e., qubits, and then developed to become more efficient [7–11] or allow for the distillation of entangled qudits [12–14]. Two main classes of distillation protocols can be identified. Recurrence protocols operate on a fixed set of input pairs iteratively, while hashing or breeding protocols operate on the whole ensemble of pairs. While the later class has in principle a higher efficiency, i.e., the inverse expected number of input states required to produce one highly entangled pair, they require low levels of noise. Recurrence protocols, on the other hand, suffer from lower efficiency but can operate on states with stronger noise. For both classes, many distillation protocols have been shown to be special cases of two generalizing schemes. Permutation-based schemes [15] use permutations of products of Bell states that can be realized by LOCC. Stabilizer-based protocols utilize stabilizer codes [16, 17], i.e., codes based on a commutative subgroup of the Pauli group of errors that take the inherent structure of error processes into account. In this contribution, we mainly consider recurrence protocols based on the stabilizer scheme.

Reflecting a general connection between error correction and entanglement distillation [18–20], it was shown that for any stabilizer code an entanglement distillation protocol can be defined [21]. In these recurrence-type protocols, which have also been extended to breeding-type protocols [22], the two parties carry out stabilizer measurements that project their local states of several qudits to subspaces called codespaces. A basis of eigenstates of these codespaces are the codewords, and a corresponding basis transformation is named encoding. It was shown that for $d = 2$ and a special class of encodings, the choice of encoding affects the performance of the corresponding protocol [23]. However, in a general setting, there is no systematic way to derive powerful protocols regarding their efficiency and minimal fidelity requirements yet. In Ref.[24], a recurrence protocol was introduced that includes information about the input state to choose between two distillation routines that can be related to certain stabilizer protocols. Showing a good efficiency for high-fidelity states, but failing to distill low-fidelity states, the information about the input state is not effectively leveraged to derive the optimal protocol.

In this contribution, we develop a general theory of stabilizer-based distillation in prime dimension that allows to relate information about the input states, error operators and used stabilizer codes to the efficacy and properties of the corresponding stabilizer distillation protocol. We demonstrate the power of this method by proposing a new distillation protocol that exhibits superior performance compared to other recurrence protocols for several state families. The paper is organized as follows. In Section 2, after introducing the notation of Bell states, error operators and the stabilizer formalism, we summarize the standard routine of stabilizer-based distillation. Section 3 analyzes the effect of errors in a given stabilizer encoding to derive a standard form, relating adjustable parameters of the protocols and information of the input states to properties of the output states. In Section 4, the developed theory is applied to the case of two-copy distillation in prime dimensions to propose the distillation protocol FIMAX that is shown to maximize the fidelity increase of Bell-diagonal states in each iteration for all stabilizer protocols. The efficacy of FIMAX is demonstrated by comparing it to other prominent recurrence protocols, where it shows notable results regarding efficiency and especially the distillation of low-fidelity states. Finally, the results are discussed in Section 5.

2 Preliminaries for stabilizer-based entanglement distillation

2.1 Bell states and stabilizer codes

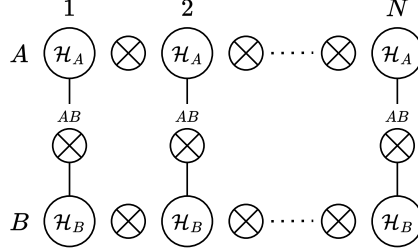


Figure 1: Hilbert space $\mathcal{H}^{\otimes N}$ of N -copies of a bipartite system $\mathcal{H} = \mathcal{H}_A \overset{AB}{\otimes} \mathcal{H}_B$.

Let $\mathcal{H} \equiv \mathcal{H}_A \overset{AB}{\otimes} \mathcal{H}_B \cong \mathbb{C}^d \otimes \mathbb{C}^d$ be the Hilbert space of bipartite quantum states with dimension of the subsystems d and $\mathcal{H}^{\otimes N} \equiv \bigotimes_{n=1}^N \mathcal{H} \cong \mathcal{H}_A^{\otimes N} \overset{AB}{\otimes} \mathcal{H}_B^{\otimes N}$ be the corresponding Hilbert space of N -copy quantum states. In this work, we restrict d to be a prime number, simplifying the analytical investigations. The tensor structure of this Hilbert space is depicted in Figure 1. Let $\mathcal{D}(\mathcal{H})$ be the set of density operators on $\mathcal{H}$. We write $\mathcal{Z}_d \equiv \mathbb{Z}/d\mathbb{Z}$ for the quotient ring of integers with addition and multiplication modulo d . Note that $\mathcal{Z}_d$ is a field for prime d as each nonzero element has a unique inverse. Let $\omega \equiv \exp\left(\frac{2\pi i}{d}\right)$, and denote equality up to a phase by $\propto$. Complex conjugation ($\star$) is defined in the computational basis.

In the following, we define the relevant objects for this work related to stabilizers of the group of Weyl-Heisenberg errors. Assuming that the reader is familiar with the basics of the stabilizer formalism, we do not provide a complete introduction. For more information about stabilizers in arbitrary dimension, see Refs. [25, 26]. In Appendix A, we provide an example for the introduced objects.

Definition 1 (Weyl(-Heisenberg) operators, Weyl(-Heisenberg) errors).

$$W_{k,l} := \sum_{j=0}^{d-1} \omega^{jk} |j\rangle \langle j+l|, \quad k, l \in \mathcal{Z}_d \quad (1)$$

$$\mathcal{E}_N := \left\{ W(e) \mid e = (\vec{k}, \vec{l}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \right\} := \left\{ \bigotimes_{n=1}^N W_{k_n, l_n} \mid k_n, l_n \in \mathcal{Z}_d \right\} \quad (2)$$

The Weyl-Heisenberg operators satisfy the Weyl relations, i.e.,

$$\begin{aligned} W_{k_1, l_1} W_{k_2, l_2} &= \omega^{l_1 k_2} W_{k_1+k_2, l_1+l_2}, \\ W_{k, l}^\dagger &= \omega^{kl} W_{-k, -l} = W_{k, l}^{-1}, \end{aligned} \quad (3)$$

implying that the set of Weyl-errors $\mathcal{E}_N$ forms a group under multiplication, if we identify errors that are equal up to a phase, i.e., $W_{k_1, l_1} W_{k_2, l_2} \equiv W_{k_1+k_2, l_1+l_2}$. $E \in \mathcal{E}_N$ are called *error operators*. For $E = W(e)$, $e \equiv (\vec{k}, \vec{l}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ are called *error elements*. The group structure of $\mathcal{E}_N$ induces a group structure via the Weyl relations (3) for error elements on $\mathcal{Z}_d^N \times \mathcal{Z}_d^N$ with addition modulo d .

Definition 2 (Bell states).

$$|\Omega_{k,l}\rangle := (W_{k,l} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle := (W_{k,l} \otimes \mathbb{1}_d) \frac{1}{\sqrt{d}} \sum_i |i\rangle \otimes |i\rangle, \quad i, k, l \in \mathcal{Z}_d \quad (4)$$

$$|\Omega(e)\rangle := (W(e) \otimes \mathbb{1}_d^{\otimes N}) |\Omega_{0,0}\rangle^{\otimes N} = \bigotimes_{n=1}^N |\Omega_{k_n, l_n}\rangle, \quad e \equiv (\vec{k}, \vec{l}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \quad (5)$$

Definition 3 (Stabilizer group, generating operators, generating elements).

A stabilizer group or stabilizer S is an abelian subgroup of $\mathcal{E}_N$. If $\{W(g_1), \dots, W(g_p)\}$ forms a minimal generating set, each $W(g_j)$ is called a generating operator and we write $S = \langle W(g_1), \dots, W(g_p) \rangle$. Each $g_j \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ is called a generating element. Given one choice of generating elements, the corresponding subgroup of $\mathcal{Z}_d^N \times \mathcal{Z}_d^N$, i.e., $G_S := \{g \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \mid W(g) \in S\}$ is also denoted as $G_S = \langle g_1, \dots, g_p \rangle$.

Let S be a stabilizer and $\{g_1, \dots, g_p\}$ be the generating elements. The spectra of the generators have the following property:

Lemma 1. For prime dimension d , each generator $W(g) \neq \mathbb{1}_d^{\otimes N}$ has d distinct eigenvalues with equal multiplicity.

Proof. First consider the special case of $d = 2$. The eigenvalues of a Weyl operator $W_{k,l}$ are $\{\omega_y = \omega^{y - \frac{1}{2}kl} \mid y \in \mathcal{Z}_2\}$ (c.f. Lemma 12.1), with $y \in \mathcal{Z}_2$. Note that there exists exactly one eigenvalue for each y , so y is uniformly distributed in $\mathcal{Z}_2$. The eigenvalues of a generator $W(g) = \bigotimes_{n=1}^N W_{k_n, l_n}$ are, consequently, $\{\prod_{n=1}^N \omega_{y_n} = \omega^{\sum_n (y_n - \frac{1}{2}k_n l_n)} \mid y_n \in \mathcal{Z}_2\}$. Since the additional phase $\sum_n \frac{1}{2}k_n l_n$ is fixed for a given generator, each eigenvalue corresponds to an $x := \sum_n y_n \in \mathcal{Z}_2$, which is again uniformly distributed in $\mathcal{Z}_2$ because each y_n is uniformly distributed and addition in $\mathcal{Z}_2$ is a bijection. This implies equal multiplicity of the two eigenvalues. A similar argument holds for prime dimensions $d \geq 3$. In this case, each $W_{k,l} \neq W_{0,0}$ has d distinct eigenvalues $\{\omega_y = \omega^y \mid y \in \mathcal{Z}_d\}$ (c.f. Lemma 12.1). Again, the non-degeneracy of each eigenvalue of $W_{k,l}$ implies that the corresponding values y are uniformly distributed over $\mathcal{Z}_d$. The eigenvalues of $W(g) = \bigotimes_{n=1}^N W_{k_n, l_n}$ are $\{\prod_{n=1}^N \omega_{y_n} = \omega^{\sum_n y_n} \mid y_n \in \mathcal{Z}_d\}$, for which each eigenvalue can again be associated with a unique element in $\mathcal{Z}_d$: $x := \sum_n y_n \in \mathcal{Z}_d$. The distribution of x in $\mathcal{Z}_d$, given the uniform distribution of y_n , is again uniform due to the bijective addition in $\mathcal{Z}_d$, so each of the d eigenvalues has the same multiplicity. Note that any trivial Weyl operator $W_{0,0}$ contained in $W(g)$ does not introduce new eigenvalues to the spectrum, but the multiplicity of each eigenvalue is increased by a factor of d . Consequently, each non-trivial generator $W(g)$ for prime dimension d has precisely d eigenvalues with equal multiplicity. $\square$

While this proof only depends on properties of the Weyl operators, we note that Lemma 1 also follows directly as a special case from the more general Theorem 2 in Ref. [17].

Using this property, we can associate each of the d eigenspaces of a generator $W(g_j)$ with a number $x_j \in \mathcal{Z}_d$. Let $x = (x_1, \dots, x_p) \in \mathcal{Z}_d^p$ and $\mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes N}$ be the joint eigenspace of $\{W(g_j)\}_{j=1}^p$, such that for all $|\phi\rangle \in \mathcal{Q}(x)$ and $j \in \{1, \dots, p\}$, we have $W(g_j)|\phi\rangle = \omega_{x_j}|\phi\rangle$ with $x_j \in \mathcal{Z}_d$ as in Lemma 1. As an example, consider $d = 2$, $N = 2$ and the generators $W(g_1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes \mathbb{1}_2$, $W(g_2) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \mathbb{1}_2$. The eigenvalues of $W(g_1)$ are $\{-i, i\}$. We associate $x_1 \leftrightarrow \omega^{x_1 - 1/2}$, so that $x_1 = 0 \leftrightarrow \omega^{-1/2} = -i$ and $x_1 = 1 \leftrightarrow \omega^{1/2} = i$. The eigenvalues of $W(g_2)$ are $\{-1, 1\}$ and we associate $x_2 \leftrightarrow \omega^{x_2}$, implying $x_2 = 0 \leftrightarrow \omega^0 = 1$ and $x_2 = 1 \leftrightarrow \omega^1 = -1$. Therefore, the eigenspace $\mathcal{Q}((0, 1))$ corresponds to the pair of eigenvalues $(-i, -1)$.

Decomposing the Hilbert space as $\mathcal{H}_A^{\otimes N} \cong \bigoplus_{x \in \mathcal{Z}_d^p} \mathcal{Q}(x)$ defines the so-called *codespaces* $\mathcal{Q}(x)$ in $\mathcal{H}_A^{\otimes N}$. Given $p \leq N$ generators of a stabilizer in $\mathcal{E}_N$ for prime dimension d , the codespaces have dimension $\dim(\mathcal{Q}(x)) = d^{N-p} \forall x \in \mathcal{Z}_d^p$. This has been shown in Theorem 2 of Ref. [17] and can also be seen from the following argument. Each codespace $\mathcal{Q}(x)$ corresponds to a p -tuple $x = (x_1, \dots, x_p) \in \mathcal{Z}_d^p$, where each entry x_j corresponds to a specific eigenvalue of $W(g_j)$. Lemma 1 above states that each generator has exactly d eigenvalues, so there are d^p distinct p -tuples and corresponding codespaces. Since also the multiplicity of each eigenvalue is equal, all codespaces must have equal dimension, which implies $\dim(\mathcal{Q}(x)) = \dim(\mathcal{H}_A^{\otimes N})/d^p = d^{N-p}$.

Let S^* be the stabilizer with complex conjugated elements of S and generating operators $W^*(g_j)$. $\mathcal{Q}^*(x) \subset \mathcal{H}_B^{\otimes N}$ is the joint eigenspace of $\{W^*(g_j)\}_{j=1}^p$, such that for all $|\phi\rangle \in \mathcal{Q}^*(x)$ and $j \in \{1, \dots, p\}$, we have $W^*(g_j)|\phi\rangle = \omega_{x_j}^*|\phi\rangle$.

Definition 4 (Error coset).

Let $S \subset \mathcal{E}_N$ be a stabilizer and let $G_S \subset \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ be the corresponding subgroup of error elements, such that $S = \{W(g) \mid g \in G_S\}$. Given an error element $e \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$, the error coset is defined as $C(e) := e + G_S = \{e + h \mid h \in G_S\}$.

Note that a coset $C(e)$ only depends on the error element e and on the stabilizer S via the corresponding subgroup G_S and not on a particular choice of stabilizer generators/generating elements.

Definition 5 (Codeword).

Let $\mathcal{H}_A^{\otimes N} \cong \bigoplus_{x \in \mathcal{Z}_d^p} \mathcal{Q}(x)$ be decomposed into d^{N-p} -dimensional codespaces of a stabilizer S . Let $\{|u_{x,k}\rangle\}_{x \in \mathcal{Z}_d^p, k \in \mathcal{Z}_d^{N-p}}$ be an orthonormal basis of $\mathcal{H}_A^{\otimes N}$ with $|u_{x,k}\rangle \in \mathcal{Q}(x)$, i.e., $W(g_j)|u_{x,k}\rangle = \omega_{x_j}|u_{x,k}\rangle, \forall k, j$. The vectors $|u_{x,k}\rangle \in \mathcal{H}_A^{\otimes N}$ are called *codewords* of S in $\mathcal{H}_A^{\otimes N}$. The codewords of S^* in $\mathcal{H}_B^{\otimes N}$ are denoted by $|u_{x,k}^*\rangle \in \mathcal{H}_B^{\otimes N}$.

Since all codespaces are of equal dimension d^{N-p} for prime d , we can define a simple mapping from the computational basis to the basis of codewords. This mapping is called encoding.

Definition 6 (Encoding).

Let $\{|x\rangle\}_{x \in \mathcal{Z}_d^p}$ be the computational basis of $\mathcal{H}_A^{\otimes p}$ and $\{|k\rangle\}_{k \in \mathcal{Z}_d^{N-p}}$ be the computational basis of $\mathcal{H}_A^{\otimes N-p}$. A unitary operator U on $\mathcal{H}_A^{\otimes N} \cong \mathcal{H}_A^{\otimes p} \otimes \mathcal{H}_A^{\otimes N-p}$ is an encoding for a stabilizer S in $\mathcal{H}_A^{\otimes N}$ if $\forall x \in \mathcal{Z}_d^p, k \in \mathcal{Z}_d^{N-p} : U(|x\rangle \otimes |k\rangle) =: |u_{x,k}\rangle$ is a codeword of S in $\mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes N}$.

If $U : |x\rangle \otimes |k\rangle \mapsto |u_{x,k}\rangle \in \mathcal{H}_A^{\otimes N}$ is an encoding for S in $\mathcal{H}_A^{\otimes N}$, then $U^* : |x\rangle \otimes |k\rangle \mapsto |u_{x,k}^*\rangle \in \mathcal{H}_B^{\otimes N}$ is an encoding for S^* in $\mathcal{H}_B^{\otimes N}$. Let $\mathcal{P}(x) : \mathcal{H}_A^{\otimes N} \rightarrow \mathcal{Q}(x)$ be the projection to the codespace $\mathcal{Q}(x)$. One has $\mathcal{P}(x) = \sum_k |u_{x,k}\rangle\langle u_{x,k}|$ for any basis of codewords $|u_{x,k}\rangle$. The same holds for projections corresponding to S^* : $\mathcal{P}^*(x) = \sum_k |u_{x,k}^*\rangle\langle u_{x,k}^*|$. For any encoding $U : |x\rangle \otimes |k\rangle \mapsto |u_{x,k}\rangle$, we have $|\Omega(\vec{0}, \vec{0})\rangle = \frac{1}{d^{N/2}} \sum_{x \in \mathcal{Z}_d^p} \sum_{k \in \mathcal{Z}_d^{N-p}} |u_{x,k}\rangle \otimes^{AB} |u_{x,k}^*\rangle$ [21].

Definition 7 (Symplectic product decomposition).

The symplectic product of two error elements, $e = (\vec{k}, \vec{l})$ and $f = (\vec{m}, \vec{n})$ is defined as $\langle e, f \rangle := \sum_{i=0}^{N-1} l_i m_i - k_i n_i$. It induces a decomposition of the set of errors according to its values $s = (s_1, \dots, s_p)$ with respect to a stabilizer with generating elements $\{g_1, \dots, g_p\} : \mathcal{E}_N \cong \bigoplus_s \mathcal{E}(s)$, with

$$\mathcal{E}(s) := \{e \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N \mid \langle g_j, e \rangle = s_j \forall j = 1, \dots, p\}. \quad (6)$$

2.2 The standard stabilizer distillation protocol

A bipartite $[N, K]$ distillation protocol transforms N copies of a bipartite *input state* $\rho_{in} \in \mathcal{D}(\mathcal{H}^{\otimes N})$ into K copies of a highly entangled bipartite *output state* $\rho_{out} \in \mathcal{D}(\mathcal{H}^{\otimes K})$ via LOCC. Recurrence protocols are iteratively applied to several copies of the input state until the output state is considered close to a specified *target state*. In this work, we consider the maximally entangled state $|\Omega_{0,0}\rangle$ (4) as the target state for recurrence protocols based on stabilizer measurements.

Let $N \geq 2$ and S be a stabilizer of $\mathcal{E}_N$ with generating elements g_j , $j \in \{1, \dots, p\}$. Assume that one party, Alice, acting locally on $\mathcal{H}_A$, can perform “stabilizer measurements” of local observables with the same eigenspaces as $W(g_j)$. Measurement of those observables corresponds to the projection onto a codespace $\mathcal{Q}(a)$. Further, assume that the second party, Bob, acting locally on $\mathcal{H}_B$, can perform stabilizer measurements corresponding to projection onto the eigenspace $\mathcal{Q}^*(b)$ of S^* . $a, b \in \mathcal{Z}_d^p$ are called “measurement outcomes”. Consider a stabilizer with p distinct generating elements. The main steps of a stabilizer-based $[N, N - p]$ distillation protocol are (cf. [21] for details):

Standard stabilizer distillation protocol:

1. Alice and Bob perform local stabilizer measurements with outcomes a, b .
2. Bob sends Alice his measurement outcome b . Alice may declare failure of the protocol depending on a and b .
3. Alice and Bob apply the inverse of stabilizer encodings U (Alice) and U^* (Bob) on $\mathcal{H}_A^{\otimes N}$ and $\mathcal{H}_B^{\otimes N}$.
4. Alice and Bob discard p qudits that are determined by the measurement, and Alice identifies and applies a local correction operation to the remaining $N - p$ qudits.

Note that by choosing a unitarily equivalent correction operation, the last two steps can in principle be carried out in changed sequence. Iteratively applying this protocol defines a recurrence distillation scheme [21].

3 Generalized stabilizer distillation

In this section, we analyze the impact of a stabilizer protocol on a state. Exploiting the algebraic properties of the Weyl operators and related Bell states, errors, and stabilizers, the action of such a protocol can be written in a form, in which the effect of different choices regarding stabilizer, encoding, measurement, and correction operation become clearly visible. These insights allow for proving certain properties and optimizations of stabilizer-based protocols.

3.1 The action of Weyl errors on codewords

A general input state can be written in the Bell basis as $\rho_{in} = \sum_{e,f \in \mathcal{E}_N} \rho(e, f) |\Omega(e)\rangle\langle\Omega(f)|$, with two error elements $e, f \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ and the density matrix elements $\rho(e, f)$. Following the same arguments as in Ref. [21], the combined effect of the measurements with outcomes a and b and the application

of the inverse encoding operations is of the following form after the third step of the protocol:

$$\begin{aligned} \rho_{in} &\mapsto \frac{1}{\text{Prob}(a=b+s)} (U^{-1} \otimes^{AB} (U^\star)^{-1}) (\mathcal{P}(a) \otimes^{AB} \mathcal{P}^\star(b)) \rho_{in} (\mathcal{P}(a) \otimes^{AB} \mathcal{P}^\star(b)) (U \otimes^{AB} U^\star) \\ &= \frac{d^{-(N-p)}}{\text{Prob}(a=b+s)} \sum_{e,f \in \mathcal{E}(s)} \sum_{j,l \in \mathcal{Z}_d^{N-p}} \rho(e,f) (U^\dagger W(e)U (|b\rangle\langle b| \otimes |j\rangle\langle l|) U^\dagger W(f)^\dagger U) \otimes^{AB} (|b\rangle\langle b| \otimes |j\rangle\langle l|). \end{aligned} \quad (7)$$

Here, $\text{Prob}(a=b+s)$ denotes the probability for obtaining the measurement outcomes a and b with $s \equiv a-b$.

An important class of input states are Bell-diagonal states (BDS) (cf., e.g., [27, 28]), arising naturally when local errors affect the maximally entangled state $|\Omega_{0,0}\rangle$. Let $\mathbb{p} : \mathcal{E}_N \rightarrow \mathbb{R}$ be a discrete probability distribution on the set of Weyl-Heisenberg errors $\mathcal{E}_N$ and $\mathbb{P}$ be the corresponding probability measure. For an error element $e \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$, $\mathbb{p}(e)$ is called *error probability*. If N -copies of maximally entangled states are affected by e with probability $\mathbb{p}(e)$, we can write the state as $\rho_{in} = \sum_{e \in \mathcal{E}} \mathbb{p}(e) |\Omega(e)\rangle\langle\Omega(e)|$. Assuming such a multi-copy Bell-diagonal input state, the state is transformed by the protocol as follows:

$$\rho_{in} \mapsto \frac{d^{-(N-p)}}{\mathbb{P}(\mathcal{E}(s))} \sum_{e \in \mathcal{E}(s)} \sum_{j,l \in \mathcal{Z}_d^{N-p}} \mathbb{p}(e) (U^\dagger W(e)U (|b\rangle\langle b| \otimes |j\rangle\langle l|) U^\dagger W(e)^\dagger U) \otimes^{AB} (|b\rangle\langle b| \otimes |j\rangle\langle l|). \quad (8)$$

From (7) and (8) it is clear that the effect of errors e in the encoding, i.e., $U^\dagger W(e)U$, is relevant for the performance of the protocol. In Ref. [23], a special subset of encodings for the case $d=2$ was analyzed regarding their impact on the performance of stabilizer-based distillation. Here, we investigate all encodings for prime dimension d .

Starting with Lemma 2, we show that an error $W(e)$ maps a codeword of the codespace $\mathcal{Q}(x)$ to a generally different codespace $\mathcal{Q}(x+s)$. s depends solely on the generating elements of the chosen stabilizer and the acting error.

Lemma 2. *Let S be a stabilizer, $x \in \mathcal{Z}_d^p$ and $|\phi\rangle \in \mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes N}$ be an eigenvector in the common eigenspace of generating operators $W(g_j)$ with corresponding eigenvalue ω_{x_j} for $j \in \{1, \dots, p\}$. Let $W(e)$ be an error operator, $s_j = \langle g_j, e \rangle$, and $s = (s_1, \dots, s_p) \in \mathcal{Z}_d^p$. Then $W(e)|\phi\rangle \in \mathcal{Q}(x+s)$.*

Proof. Using the Weyl relations (3), one has $\forall j$:

$$W(g_j)W(e)|\phi\rangle = \omega^{(g_j,e)} W(e)W(g_j)|\phi\rangle = \omega^{s_j} \omega_{x_j} W(e)|\phi\rangle = \omega_{x_j+s_j} W(e)|\phi\rangle \quad \forall j \implies W(e)|\phi\rangle \in \mathcal{Q}(x+s)$$

□

Lemma 3 demonstrates that for each error e the precise *action of the error e* within the codespace is determined by a d^{N-p} -dimensional unitary transformation $T_x^{U,e}$ that depends on the codespace, the encoding and the error. In essence, these operators reflect which effect an error has on the codewords of a stabilizer code. In Section 3.2, it is demonstrated that, together with the stabilizer measurements, these action operators determine the output state of the stabilizer protocol.

Lemma 3. *Let U be an encoding of a stabilizer S with generating elements $\{g_1, \dots, g_p\}$. For each codespace $\mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes N}$ and for each $W(e) \in \mathcal{E}_N$ with $(\langle g_1, e \rangle, \dots, \langle g_p, e \rangle) = (s_1, \dots, s_p)$, there exist unitary “action” operators $T_x^{U,e} : \mathcal{H}_A^{\otimes N-p} \rightarrow \mathcal{H}_A^{\otimes N-p}$ satisfying $U^\dagger W(e)U = \sum_{x \in \mathcal{Z}_d^p} |x+s\rangle\langle x| \otimes T_{x+s}^{U,e}$.*

Proof. Lemma 2 implies $W(e)|u_{b,j}\rangle \in \mathcal{Q}(b+s)$, which is spanned by $\{|u_{b+s,k}\rangle\}_{k \in \mathcal{Z}_d^{N-p}}$. Consequently,

$$U^\dagger W(e)U(|b\rangle \otimes |j\rangle) = U^\dagger W(e)|u_{b,j}\rangle = U^\dagger \sum_{k \in \mathcal{Z}_d^{N-p}} t_{b+s,k}^{e,j} |u_{b+s,k}\rangle = |b+s\rangle \otimes \sum_{k \in \mathcal{Z}_d^{N-p}} t_{b+s,k}^{e,j} |k\rangle.$$

The elements $\{t_{b+s,k}^{e,j}\}_{k,j \in \mathcal{Z}_d^{N-p}}$ define the action operator $T_{b+s}^{U,e}$ of the error e in the computational basis via

$$T_{b+s}^{U,e} := \sum_{k,j \in \mathcal{Z}_d^{N-p}} t_{b+s,k}^{e,j} |k\rangle \langle j|. \quad (9)$$

One then has $U^\dagger W(e)U(|b\rangle \otimes |j\rangle) = |b+s\rangle \otimes T_{b+s}^{U,e} |j\rangle$, implying $U^\dagger W(e)U = \sum_{x \in \mathcal{Z}_d^p} |x+s\rangle \langle x| \otimes T_{x+s}^{U,e}$. $\square$

The following two lemmas show that the group properties of errors and stabilizers are also reflected by the action operators T and therefore by the effective action of errors in a given encoding. This will be leveraged to derive a simple form of the output state of a stabilizer protocol in Section 3.2.

Lemma 4 illustrates that the linear structure of $\mathcal{E}_N$ naturally extends to the action of errors.

Lemma 4. *Let e, f be two error elements and $T_{x+s_e}^{U,e}$ and $T_{x+s_f}^{U,f}$ be the corresponding actions in the codespaces. We then have:*

$$U^\dagger W(e+f)U \propto \sum_x |x+s_e+s_f\rangle \langle x| \otimes T_{x+s_e}^{U,e} T_{x+s_f}^{U,f} \quad (10)$$

$$U^\dagger W(-e)U \propto \sum_x |x-s_e\rangle \langle x| \otimes (T_x^{U,e})^\dagger \quad (11)$$

Proof. Follows directly from $W(e+f) \propto W(e)W(f)$ and $W(-e) \propto W(e)^\dagger$ and Lemma 3. $\square$

The following lemma shows that the actions of two errors that are related by a generating element are equivalent up to a phase. This directly implies that errors of the same coset (Definition 4) are also equivalent up to a phase.

Lemma 5. *Let U be an encoding for a stabilizer S as in Lemma 3 and let $T_{x+s}^{U,e}$ be the corresponding action for an error element e . The following equality holds:*

$$T_{x+s}^{U,e+g_j} = \omega_{x_j} T_{x+s}^{U,e} \quad \forall j \in \{1, \dots, p\}. \quad (12)$$

Proof. Noting $U^\dagger W(g_j)U = \sum_{x \in \mathcal{Z}_d^p} |x\rangle \langle x| \otimes \omega_{x_j} \mathbb{1}_{d^{N-p}}$, the claim follows from Lemma 4 (10) with $f = g_j$. $\square$

Proposition 6 establishes a connection between all possible encodings. Given an encoding, all other encodings are related by concatenation via a block-diagonal unitary transformation. Conversely, any such concatenation provides another encoding. Moreover, the action of errors in a concatenated encoding can be directly derived.

Proposition 6. *Let U be an encoding for a stabilizer S with generating elements $\{g_1, \dots, g_p\}$.*

- (i) *A unitary V is another encoding for S if and only if for each codespace $\mathcal{Q}(x)$, $x \in \mathcal{Z}_d^p$, there exists unitary $(\dim(\mathcal{Q}(x)) \times \dim(\mathcal{Q}(x)))$ -matrices Y_x so that $V = U (\sum_x |x\rangle \langle x| \otimes Y_x)$.*

(ii) Let $V = U (\sum_x |x\rangle\langle x| \otimes Y_x)$. If $U^\dagger W(e)U = \sum_x |x+s\rangle\langle x| \otimes T_{x+s}^{U,e}$ as in Lemma 3. Then $V^\dagger W(e)V = \sum_x |x+s\rangle\langle x| \otimes Y_{x+s}^\dagger T_{x+s}^{U,e} Y_x$.

Proof.

(i) Let U and V be encodings such that $U(|y\rangle \otimes |l\rangle) \equiv |u_{y,l}\rangle$ and $V(|x\rangle \otimes |k\rangle) \equiv |v_{x,k}\rangle$. Then

$$(\langle y| \otimes \langle l|) U^\dagger V (|x\rangle \otimes |k\rangle) = \langle u_{y,l} | v_{x,k} \rangle = \delta_{x,y} \langle u_{y,l} | v_{x,k} \rangle$$

and consequently

$$U^\dagger V = \sum_{x \in \mathbb{Z}_d^p} |x\rangle\langle x| \otimes \sum_{l,k \in \mathbb{Z}_d^{N-p}} \langle u_{x,l} | v_{x,k} \rangle |l\rangle\langle k| =: \sum_x |x\rangle\langle x| \otimes Y_x.$$

Conversely, assume $V = U (\sum_x |x\rangle\langle x| \otimes Y_x)$. We need to show that $V(|x\rangle \otimes |k\rangle)$ defines a codeword, i.e., that $W(g_j)V(|x\rangle \otimes |k\rangle) = \omega_{x_j} V(|x\rangle \otimes |k\rangle) \quad \forall j \in \{1, \dots, p\}, x \in \mathbb{Z}_d^p, k \in \mathbb{Z}_d^{N-p}$. This can be shown by direct calculation:

$$\begin{aligned} W(g_j)V(|x\rangle \otimes |k\rangle) &= W(g_j) U(|x\rangle \otimes Y_x |k\rangle) = W(g_j) U(|x\rangle \otimes \sum_l y_{x,l} |l\rangle) \\ &= \omega_{x_j} \sum_l y_{x,l} U(|x\rangle \otimes |l\rangle) = \omega_{x_j} U(|x\rangle \otimes Y_x |k\rangle) \\ &= \omega_{x_j} V(|x\rangle \otimes |k\rangle). \end{aligned}$$

(ii) Using Lemma 3 for both U and V yields

$$\begin{aligned} V^\dagger W(e)V &= (\sum_z |z\rangle\langle z| \otimes Y_z^\dagger) U^\dagger W(e) U (\sum_y |y\rangle\langle y| \otimes Y_y) \\ &= (\sum_z |z\rangle\langle z| \otimes Y_z^\dagger) (\sum_x |x+s\rangle\langle x| \otimes T_{x+s}^{U,e}) (\sum_y |y\rangle\langle y| \otimes Y_y) \\ &= \sum_x |x+s\rangle\langle x| \otimes Y_{x+s}^\dagger T_{x+s}^{U,e} Y_x. \end{aligned}$$

□

3.2 Standard form of stabilizer distillation protocols

The results of the previous section suggest that the effect of a stabilizer protocol (cf. (7) (8)) can be made more concise by considering the action of errors in a given encoding. First, Proposition 7 provides a simplified form of the effect of the stabilizer measurements and decoding operations. Theorem 8 introduces a “standard form” of the stabilizer distillation protocol for Bell-diagonal input states, incorporating the group properties of the Weyl errors.

Proposition 7. Let S be a stabilizer with p generating elements defining the symplectic partition of errors in $\mathcal{E}_N \cong \bigoplus_s \mathcal{E}(s)$ and U be an encoding. Let $\rho_{in} = \sum_{e,f \in \mathcal{E}_N} \rho(e,f) |\Omega(e)\rangle\langle\Omega(f)|$ be a general input state in the Bell basis. Let $\text{Prob}(a = b + s)$ be the probability of obtaining such outcomes for the stabilizer measurements. After projection of the state onto the codespace $\mathcal{Q}(a) \otimes \mathcal{Q}^*(b)$ with $s = a - b$,

applying $U^{-1} \otimes^{AB} (U^\star)^{-1}$ and discarding the first p copies, the output state $\rho_{out} \in \bigotimes_{n=p+1}^N \mathcal{H}_A \otimes^{AB} \mathcal{H}_B$ is

$$\rho_{out} = \frac{1}{\text{Prob}(a = b + s)} \sum_{e,f \in \mathcal{E}(s)} \rho(e, f) (T_{b+s}^{U,e} \otimes^{AB} \mathbb{1}) |\Omega_{0,0}\rangle \langle \Omega_{0,0}|^{\otimes N-p} (T_{b+s}^{U,f} \otimes^{AB} \mathbb{1})^\dagger. \quad (13)$$

Proof. Starting with the state in the form of Eq.(7), Lemma 2 implies that after the projective measurements with outcomes a and b , only terms relating to errors contained in $\mathcal{E}(s)$ have nonzero components. Lemma 3 determines the form of those copies that are not trivially determined by the measurement outcomes a, b via the action of errors in the applied encoding. $\square$

For Bell-diagonal input states, further simplification can be achieved by considering error cosets because, according to Lemma 5, errors from the same coset only differ by a phase that cancels for Bell-diagonal input states. We can therefore represent all action operators $T_x^{U,e}$ for errors of the same coset C by a single coset action operator $T_x^{U,C}$. As a N -copy Bell-diagonal state corresponds to a probability distribution $\mathbb{p}$ on $\mathcal{E}_N$, combining errors of the same coset induces a distribution for the error coset probabilities.

Definition 8.

$$\mathcal{C} := \{C \mid C \text{ is an error coset}\} \quad (14)$$

$$\mathcal{C}(s) := \{C \in \mathcal{C} \mid C \subset \mathcal{E}(s)\} \quad (15)$$

$$\mathbb{P}(C) = \sum_{e \in C} \mathbb{P}(e) \quad (16)$$

$$T_{b+s}^{U,C} := T_{b+s}^{U,e} \text{ for an arbitrary } e \in C \quad (17)$$

$$T_{b+s}^{U,(C_1-C_2)} := T_{b+s}^{U,(e_1-e_2)} \text{ for arbitrary } e_1 \in C_1, e_2 \in C_2 \quad (18)$$

Given Bell-diagonal input states and combining these definitions with Proposition 7 allows deriving a form of the output state that solely depends on objects relating to cosets instead of individual errors.

Theorem 8. Let $\rho_{in} = \sum_{e \in \mathcal{E}} \mathbb{P}(e) |\Omega(e)\rangle \langle \Omega(e)|$ be a bipartite N -copy Bell-diagonal state, inducing a probability measure $\mathbb{P}$ on $\mathcal{E}_N$. Let S and U be as in Proposition 7. After projection of the state onto the codespace $\mathcal{Q}(a) \otimes^{AB} \mathcal{Q}^\star(b)$ with $s = a - b$, applying $U^{-1} \otimes^{AB} (U^\star)^{-1}$ and discarding the first p copies, ρ_{in} is transformed to $\rho_{out} \in \mathcal{D}(\mathcal{H}_A \otimes^{AB} \mathcal{H}_B)^{\otimes N-p}$ in the so-called “standard form”:

$$\rho_{out} = \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) (T_{b+s}^{U,C} \otimes^{AB} \mathbb{1}) |\Omega_{0,0}\rangle \langle \Omega_{0,0}|^{\otimes N-p} (T_{b+s}^{U,C} \otimes^{AB} \mathbb{1})^\dagger. \quad (19)$$

Proof. Assuming Bell-diagonal form, we have $\text{Prob}(a = b + s) = \mathbb{P}(\mathcal{E}(s))$ and $\rho(e, e) \equiv \mathbb{p}(e)$. Proposition 7 implies the claimed form, by noting that Lemma 5 allows identifying errors of the same coset C (cf. Definition 8), as their action operators only differ by a phase that cancels for diagonal elements. $\square$

In this standard form, the effect of the adjustable parameters of the protocol become clearly visible:

- *Input state ρ_{in} :* The input state reflects the probability measure $\mathbb{P}$ on $\mathcal{E}_N$ and thus the probability of errors.
- *Stabilizer S :* The number of generators p of S determines how many copies are used to gain information from measurements and are discarded afterward. The remaining $N - p$ copies

form a maximally entangled state that is affected locally by an error with some probability. S determines the decomposition of the Hilbert space in codespaces and of $\mathcal{E}_N$ in the sets $\mathcal{E}(s)$ as well as in cosets C .

- *Measurement and classical communication:* The measurement outcomes a and b effectively limit the set of possible errors to $\mathcal{E}(s)$. If only Alice needs to know s for further operations, one-way communication from Bob to Alice is enough. Otherwise, two-way communication is required.
- *Encoding U :* The encoding fixes a basis of codewords for the codespaces and therefore determines the effect of local errors according to the error action operators $T_{b+s}^{U,e}$. The output state ρ_{out} can be written as mixed Bell states that are locally affected by these error actions on Alice's system.

For the sake of clarity, we omit the indices representing the dependence of the error action operators on the chosen encoding and the subspace they act on. Hence, we write $T_{b+s}^{U,C} \equiv T_C$ if there is no risk of confusion.

3.3 Fidelities and local error correction in the standard form

The standard form of Theorem 8 allows finding local correction operations and calculating fidelities for a target state $|\Omega(\vec{k}, \vec{l})\rangle$, $(\vec{k}, \vec{l}) \in \mathcal{Z}_d^{N-p} \times \mathcal{Z}_d^{N-p}$. Note that these are $(N-p)$ -copy states, as the p copies containing only information about the measurement outcomes are discarded in the protocol. Let $V \otimes \mathbb{1}_{d^{N-p}}^{AB}$ be a local unitary, applied by Alice (w.l.o.g.) depending on the measurement outcomes and transforming the output state to

$$\rho_{out} \mapsto \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) (VT_C \otimes \mathbb{1}^{AB} |\Omega_{0,0}\rangle \langle \Omega_{0,0}|^{\otimes N-p} (VT_C \otimes \mathbb{1}^{AB})^\dagger. \quad (20)$$

Choosing $V = T_{\hat{C}}^\dagger$ with some coset $\hat{C}$ shows that a fidelity of $\frac{\mathbb{P}(\hat{C})}{\mathbb{P}(\mathcal{E}(s))}$ can be achieved:

$$\rho_{out} \mapsto \frac{\mathbb{P}(\hat{C})}{\mathbb{P}(\mathcal{E}(s))} (|\Omega_{0,0}\rangle \langle \Omega_{0,0}|^{\otimes N-p} + \sum_{C \in \mathcal{C}(s) \setminus \hat{C}} \mathbb{P}(C) (T_{C-\hat{C}} \otimes \mathbb{1}^{AB} |\Omega_{0,0}\rangle \langle \Omega_{0,0}|^{\otimes N-p} (T_{C-\hat{C}} \otimes \mathbb{1}^{AB})^\dagger. \quad (21)$$

In Ref. [21] it was shown that this fidelity with $|\Omega(\vec{0}, \vec{0})\rangle = |\Omega_{0,0}\rangle^{\otimes N-p}$ can also be obtained if Alice applies $W(e)^{-1}$, $e \in \hat{C}$ before the inverse encoding U^{-1} . The standard form (19) has the advantage that the fidelities for all Bell states for any encoding can be directly calculated from its error actions T_C . This makes a quantitative comparison of different encodings possible.

The fidelity between any multi-copy Bell state and the output state in standard form (19) is

$$\begin{aligned} \mathcal{F}(\vec{k}, \vec{l}) &:= \langle \Omega(\vec{k}, \vec{l}) | \rho_{out} | \Omega(\vec{k}, \vec{l}) \rangle = \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) |\langle \Omega(\vec{k}, \vec{l}) | (T_C \otimes \mathbb{1}^{AB}) | \Omega(\vec{0}, \vec{0}) \rangle|^2 \\ &= \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) |\text{Tr}(\frac{1}{d^{N-p}} W^\dagger(\vec{k}, \vec{l}) T_C)|^2. \end{aligned} \quad (22)$$

For the last equality of (22), the following identity for $|\Omega(\vec{0}, \vec{0})\rangle$ of dimension D and all $(D \times D)$ matrices M is used:

$$\langle \Omega(\vec{0}, \vec{0}) | M \otimes \mathbb{1}^{AB} | \Omega(\vec{0}, \vec{0}) \rangle = \frac{1}{D} \text{Tr}(M). \quad (23)$$

With the *Weyl representation* of T_C ,

$$T_C = \sum_{(\vec{i}, \vec{j}) \in \mathbb{Z}_d^{2(N-p)}} \beta_C(\vec{i}, \vec{j}) W(\vec{i}, \vec{j}), \quad \beta_C(\vec{i}, \vec{j}) := \frac{1}{d^{N-p}} \text{Tr}(W^\dagger(\vec{i}, \vec{j}) T_C) \in \mathbb{C}, \quad (24)$$

the fidelities are directly related to the coefficients of β_C in this representation. More precisely, Corollary 9 demonstrates that the total fidelity is a weighted sum of so-called *coset fidelities* $f_C(\vec{k}, \vec{l}) := |\beta_C(\vec{k}, \vec{l})|^2$.

Corollary 9. *Let ρ_{out} be as in Theorem 8. Then the fidelity $\mathcal{F}(\vec{k}, \vec{l})$ of ρ_{out} can be expressed as*

$$\mathcal{F}(\vec{k}, \vec{l}) = \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) |\beta_C(\vec{k}, \vec{l})|^2 =: \frac{1}{\mathbb{P}(\mathcal{E}(s))} \sum_{C \in \mathcal{C}(s)} \mathbb{P}(C) f_C(\vec{k}, \vec{l}). \quad (25)$$

Proof. Follows directly from (22) with (24). $\square$

Proposition 10 states properties of the coset fidelities and of the states $\{T_C \otimes \mathbb{1} |\Omega(\vec{k}, \vec{l})\rangle\}$, emerging in the standard form (19). These properties will allow showing that the protocol proposed in section 4 maximizes the increase in fidelity in each iteration among all stabilizer-based protocols. We introduce the following notation for cosets C_1, C_2 :

$$\delta_{C_1, C_2} := \begin{cases} 1 & \text{if } C_1 = C_2 \\ 0 & \text{else.} \end{cases} \quad (26)$$

Proposition 10. *Let S be a stabilizer, G_S the set of corresponding error elements as in Definition 4, T_C be as in Theorem 8 and $f_C(\vec{k}, \vec{l})$ be as in Corollary 9. Then,*

- (i) $\forall C : \sum_{(\vec{k}, \vec{l})} f_C(\vec{k}, \vec{l}) = 1$.
- (ii) $\forall(\vec{k}, \vec{l})$, states in $\{T_C^\dagger \otimes \mathbb{1} |\Omega(\vec{k}, \vec{l})\rangle \mid C \in \mathcal{C}(s)\}$ are orthonormal up to a phase $\iff \frac{1}{d^{N-p}} \text{Tr}(T_C^\dagger) \propto \delta_{C, G_S}$.
- (iii) $\{T_C^\dagger \otimes \mathbb{1} |\Omega(\vec{k}, \vec{l})\rangle \mid C \in \mathcal{C}(s)\}$ is an ONB of $(\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes N-p} \implies \sum_{C \in \mathcal{C}} f_C(\vec{k}, \vec{l}) = 1$.

Proof.

(i) $\forall C$, T_C is unitary, implying $1 = \sum_{(\vec{k}, \vec{l})} \beta^*(\vec{k}, \vec{l}) \beta(\vec{k}, \vec{l}) = \sum_{(\vec{k}, \vec{l})} f(\vec{k}, \vec{l})$ by (24).

(ii) Consider the identity (23). For $C_1, C_2 \in \mathcal{C}(s)$ this implies

$$\langle \Omega(\vec{k}, \vec{l}) | (T_{C_1}^\dagger \otimes \mathbb{1}) (T_{C_2}^\dagger \otimes \mathbb{1}) | \Omega(\vec{k}, \vec{l}) \rangle = \frac{1}{d^{N-p}} \text{Tr}(T_{C_1} T_{C_2}^\dagger) \propto \frac{1}{d^{N-p}} \text{Tr}(T_{C_1 - C_2}).$$

The last equation follows from Lemma 5 and Lemma 4 together with Definition 8. Assume the right-hand side of the equivalence. The equation above implies orthonormality up to a phase for the states $T_C^\dagger \otimes \mathbb{1} |\Omega(\vec{k}, \vec{l})\rangle$ if $C_1 - C_2 = G_S \Leftrightarrow C_1 = C_2$. Conversely, orthonormality up to a phase of elements on the left side of the equivalence implies $\frac{1}{d^{N-p}} \text{Tr}(T_C) \propto \frac{1}{d^{N-p}} \text{Tr}(T_{C - G_S}) \propto \langle \Omega(\vec{k}, \vec{l}) | (T_C^\dagger \otimes \mathbb{1}) (T_{G_S}^\dagger \otimes \mathbb{1}) | \Omega(\vec{k}, \vec{l}) \rangle \propto \delta_{C, G_S}$.

(iii) Comparing (22) with (25) and assuming the ONB property, one has

$$\begin{aligned} \sum_{C \in \mathcal{C}} f_C(\vec{k}, \vec{l}) &= \sum_{C \in \mathcal{C}} |\langle \Omega(\vec{k}, \vec{l}) | (T_C^{AB} \otimes \mathbb{1}) | \Omega(\vec{0}, \vec{0}) \rangle|^2 \\ &= \sum_{C \in \mathcal{C}} \langle \Omega(\vec{k}, \vec{l}) | (T_C^{AB} \otimes \mathbb{1}) | \Omega(\vec{0}, \vec{0}) \rangle \langle \Omega(\vec{0}, \vec{0}) | (T_C^{\dagger AB} \otimes \mathbb{1}) | \Omega(\vec{k}, \vec{l}) \rangle = \text{Tr}(|\Omega(\vec{0}, \vec{0})\rangle \langle \Omega(\vec{0}, \vec{0})|) = 1. \end{aligned}$$

□

4 Two-copy distillation in prime dimension

In this section, we consider the case $N = 2$ for prime dimension d . Introducing the canonical encoding, we apply the results of the previous section to characterize all other encodings. In section 4.3, we propose a protocol that maximizes the fidelity increase in each iteration for Bell-diagonal input states and compare it numerically to other protocols in section 4.4. Note that the results of the sections 4.1 and 4.2 do not depend on the input state and are therefore applicable to non-Bell-diagonal states as well.

4.1 Error sets and stabilizers for prime dimension

From the standard form (cf. Theorem 8) it follows that the output state of the generalized stabilizer protocol is a mixed $(N-p)$ -copy bipartite state. Consequently, the only number of generating elements p of a stabilizer S , which results in a non-trivial transformation of the input state for $N = 2$, is $p = 1$. Therefore, all relevant stabilizer groups S have exactly one generating element $g \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2$. Since d is prime, the order of each S is d and we can explicitly write $S = \{\mathbb{1}, W(g), W(2g), \dots, W((d-1)g)\}$. Every cyclic group is abelian, so any error operator of $W(e) \neq \mathbb{1}$ generates a stabilizer group and two stabilizer groups sharing one element are identical. The following Lemma 11 shows that the partitions of errors in $\mathcal{E}_N \cong \bigoplus_s \mathcal{E}(s)$ (cf. Definition 7) are of equal size for prime dimensions.

Lemma 11. *Let d be prime, $g \neq (\vec{0}, \vec{0}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$ inducing $\mathcal{E}_N \cong \bigoplus_s \mathcal{E}(s)$. Then $\forall s \in \mathcal{Z}_d$ we have $|\mathcal{E}(s)| = d^{2N-1}$.*

Proof. We define a bijective map $M : \mathcal{E}(0) \rightarrow \mathcal{E}(s)$, implying equal cardinality of the sets. Let $g = (\vec{k}, \vec{l}) \neq (\vec{0}, \vec{0}) \in \mathcal{Z}_d^N \times \mathcal{Z}_d^N$. Assume that there exists a component (k_i, l_i) with $k_i \neq 0$ (w.l.o.g.). Let $e_0 \in \mathcal{E}(0)$. Such an element always exists, since $g \in \mathcal{E}(0)$. d is prime, so $\exists k_i^{-1} \in \mathcal{Z}_d$. Define for $e = (\vec{m}, \vec{n}) = ((m_1, \dots, m_N), (n_1, \dots, n_N)) \in \mathcal{E}(0)$ the map $M : e_0 \mapsto e_s = ((m_1, \dots, m_i + sk_i^{-1}, \dots, m_N), (n_1, \dots, n_N))$. k_i^{-1} is unique and $sk_i^{-1} \neq 0$ for $s \neq 0$. M is a bijection and $\langle g, e_s \rangle = s$, so the d partitions $\mathcal{E}(s)$ must be of equal cardinality d^{2N-1} . □

4.2 A canonical stabilizer encoding

In this section, a specific encoding based on the eigenvectors of the Weyl-Heisenberg operators is defined. We show that this *canonical encoding* has special properties and implications for corresponding stabilizer distillation protocols.

First, we analyze how Weyl errors affect eigenstates of the Weyl-Heisenberg operators, and thus code-words in the canonical encoding, in Lemma 12. The proof relies on the technical lemmas 12.1 and 12.2 following below.

Lemma 12. Let d be prime and $|\omega_\lambda\rangle$ be an eigenvector of $W_{a,b}$ with eigenvalues depending on λ as in Lemma 12.1. For $x, y \in \mathbb{Z}_d$, it holds that $W_{x,y} |\omega_\lambda\rangle = \omega^\Phi |\omega_{\lambda+s}\rangle$ with $\Phi = \Phi(\lambda, a, b, x, y) = t(a, b, x, y)\lambda + c(a, b, x, y)$ and $s = \langle g, e \rangle$.

Proof. Follows directly from the Lemmas 12.1 and 12.2. $\square$

Lemma 12.1. For prime $d > 2$, the eigenvalues and eigenvectors of $W_{a,b}$, $a, b \in \mathbb{Z}_d$ are as follows.

- $a = b = 0$:

Eigenvalues: $\omega_\lambda := \omega^0 = 1$, $\lambda \in \mathbb{Z}_d$.

Eigenvectors: $|\omega_\lambda\rangle := |\lambda\rangle$

- $a \neq 0, b = 0$: d is prime, so $a^{-1} \in \mathbb{Z}_d$ exists and is unique.

Eigenvalues: $\omega_\lambda := \omega^\lambda$, $\lambda \in \mathbb{Z}_d$.

Eigenvectors: $|\omega_\lambda\rangle := |\lambda a^{-1}\rangle$.

- $a \in \mathbb{Z}_d, b \neq 0$: d is prime, so $b^{-1} \in \mathbb{Z}_d$ exists and is unique.

Eigenvalues: $\omega_\lambda := \omega^\lambda$, $\lambda \in \mathbb{Z}_d$.

Eigenvectors: $|\omega_\lambda\rangle := \frac{1}{\sqrt{d}} \sum_{j \in \mathbb{Z}_d} \omega^{\Gamma_{\lambda,j}} |j\rangle$, $\Gamma_{\lambda,j} := jb^{-1}\lambda - \frac{jb^{-1}(jb^{-1}-1)}{2}ab$.

In the unique special case of $d = 2$ and $a \in \mathbb{Z}_2, b \neq 0$, one has:

- $a \in \mathbb{Z}_2, b \neq 0$:

Eigenvalues: $\omega_\lambda := \omega^{\lambda - \frac{1}{2}ab}$, $\lambda \in \mathbb{Z}_2$.

Eigenvectors: $|\omega_\lambda\rangle = \frac{1}{\sqrt{2}}(|0\rangle + w_\lambda|1\rangle)$

Proof. That the states $|\omega_\lambda\rangle$ are eigenvectors with eigenvalues as stated can be seen by direct calculation. Since the spectrum is non-degenerate for the nontrivial case $(a, b) \neq (0, 0)$, these states form a basis. $\square$

Lemma 12.2. Let $|\omega_\lambda\rangle$ be an eigenvector of $W_{a,b}$. Given $W_{x,y}$ for prime d , the following holds:

- $a = b = 0$: $W_{x,y} |\omega_\lambda\rangle = \omega^{\Phi(\lambda, 0, 0, x, y)} |\omega_{\lambda-y}\rangle := \omega^{x(\lambda-y)} |\omega_{\lambda-y}\rangle$.
- $a \neq 0, b = 0$: $W_{x,y} |\omega_\lambda\rangle = \omega^{\Phi(\lambda, a, 0, x, y)} |\omega_{\lambda-ay}\rangle := \omega^{xa^{-1}(\lambda-ay)} |\omega_{\lambda-ay}\rangle$.

For $d > 2$:

- $a \in \mathbb{Z}_d, b \neq 0$: $W_{x,y} |\omega_\lambda\rangle = \omega^{\Phi(\lambda, a, b, x, y)} |\omega_{\lambda+bx-ay}\rangle := \omega^{yb^{-1}(\lambda - \frac{1}{2}a(y-b))} |\omega_{\lambda+bx-ay}\rangle$.

For $d = 2$:

- $a \in \mathbb{Z}_2, b \neq 0$: $W_{x,y} |\omega_\lambda\rangle = \omega^{\Phi(\lambda, a, b, x, y)} |\omega_{\lambda+bx-ay}\rangle := \omega^{y(\lambda - \frac{1}{2}ab)} |\omega_{\lambda+bx-ay}\rangle$

Proof. Direct calculation with Lemma 12.1. $\square$

Given a stabilizer, the product basis of eigenstates of the Weyl-Heisenberg operators contained in the generating operator $W(g) = W_{a_1, b_1} \otimes W_{a_2, b_2}$ defines a valid encoding:

Definition 9. (Canonical encoding)

Given a stabilizer S with the generating element g and $W(g) = W_{a_1, b_1} \otimes W_{a_2, b_2} \in \mathcal{E}_2$ for prime d . Let $\{|\omega_\lambda^1\rangle\}_{\lambda \in \mathbb{Z}_d}$ be a basis of eigenvectors of W_{a_1, b_1} and $\{|\omega_\lambda^2\rangle\}_{\lambda \in \mathbb{Z}_d}$ be a basis of eigenvectors of W_{a_2, b_2} as in Lemma 12.1.

The canonical encoding is defined as

- (i) $U_c : |\lambda\rangle \otimes |k\rangle \mapsto |u_{\lambda,k}\rangle = |\omega_k^1\rangle \otimes |\omega_{\lambda-k}^2\rangle$ (if $(a_1, b_1), (a_2, b_2) \neq (0, 0)$).
- (ii) $U_c : |\lambda\rangle \otimes |k\rangle \mapsto |u_{\lambda,k}\rangle = |\omega_\lambda^1\rangle \otimes |\omega_k^2\rangle$ (if (w.l.o.g.) $(a_2, b_2) = (0, 0)$).

The canonical encoding has the interesting property that its actions of errors are Weyl-Heisenberg operators. This implies that the canonical encoding maps the Pauli group of Weyl-Heisenberg errors (1) onto itself, i.e., the encoding operator U_c are elements of the Clifford group. Lemma 13 identifies the precise Weyl-Heisenberg operator that represents the action of any error (coset). An example for a specific stabilizer can be found in the appendix A.

Lemma 13. *Let $g = ((a_1, b_1), (a_2, b_2))$ be the generator of the stabilizer and U_c be the canonical encoding. Let T_e be the corresponding action of errors $e = ((x_1, y_1), (x_2, y_2))$, and T_C the corresponding coset action. Then $T_e \propto T_C = W_{t, -s_1}$ with $t = t(a_1, a_2, b_1, b_2, x_1, x_2, y_1, y_2)$ and $s_1 = b_1 x_1 - a_1 y_1$.*

Proof. Definition 9 and Lemma 12 imply $U_c^{-1}(W_{x_1, y_1} \otimes W_{x_2, y_2})U_c |\lambda\rangle \otimes |k\rangle = |\lambda + s_1 + s_2\rangle \otimes \omega^{\Phi_1 + \Phi_2} |k + s_1\rangle$. In the case of assumption (i) in Definition 9, we have $\Phi_1 = \Phi_1(k, a_1, b_1, x_1, y_1)$ and $\Phi_2 = \Phi_2(\lambda - k, a_2, b_2, x_2, y_2)$, while in the case of assumption (ii) $\Phi_1 = \Phi_1(b, a_1, b_1, x_1, y_1)$ and $\Phi_2 = \Phi_2(k, a_2, b_2, x_2, y_2)$. Using Lemma 12.2 one has:

- (i) $\Phi_1 + \Phi_2 = \Phi_1(k, a_1, b_1, x_1, y_1) + \Phi_2(b - k, a_2, b_2, x_2, y_2) = (t_1 - t_2)k + c_1 + t_2 b =: tk + c$.
- (ii) $\Phi_1 + \Phi_2 = \Phi_1(b, a_1, b_1, x_1, y_1) + \Phi_2(k, a_2, b_2, x_2, y_2) = t_1 b + c_1 + t_2 k + c_2 =: tk + c$.

This implies

$$U_c^{-1}(W_{x_1, y_1} \otimes W_{x_2, y_2})U_c |x\rangle \otimes |k\rangle = |x + s\rangle \otimes \omega^c \omega^{tk} |k + s_1\rangle = |x + s\rangle \otimes \omega^{c - ts_1} W_{t, -s_1} |k\rangle.$$

and comparison to Lemma 3 and Definition 8 shows the claimed property. $\square$

By definition, any stabilizer contains the unity and consequently any error element in the stabilizer coset has trivial action. The following Lemma 14 shows that the stabilizer is the only coset for which that holds.

Lemma 14. *Let S be a stabilizer with error elements G_S as in Definition 4, U_c be the canonical encoding with $T_{x+s}^{U_c, e} \propto T_C = W_{t, -s_1}$. We then have the following equivalence: $C = G_S \Leftrightarrow T_C \propto W_{0,0} = \mathbb{1}_d$.*

Proof. Assume $W(e) \in S$. This implies $T_e \propto T_{G_S} \propto \mathbb{1}_d$ by Lemma 5. Conversely, for $e \in C$ assume $T_e \propto W_{0,0}$, implying $s_1 = 0$ by Lemma 13 and $\langle g, e \rangle \equiv s = s_1 + s_2 = s_2$. If $s_2 = 0 \Rightarrow s = 0$. Lemmas 2–3 imply

$$W(e)U |b\rangle \otimes |k\rangle = W(e)|u_{b,k}\rangle \propto |u_{b+s,k}\rangle = |u_{b,k}\rangle \implies W(e) \in S \implies C = G_S.$$

If $s_2 \neq 0$, $\exists s_2^{-1} \neq 0$ since d is prime. We then have by Lemmas 2–4:

$$W(e)^{(s_2^{-1})} |u_{b,k}\rangle \propto |u_{b+s_2^{-1}s_2,k}\rangle = |u_{b,k}\rangle \implies W(e)^{(s_2^{-1})} \in S \implies W(e) \in S \implies C = G_S.$$

$\square$

Combining these results with the standard form of the output state (19), Lemma 15 demonstrates that for the canonical encoding and Bell-diagonal input state, the output state is again a mixture of pure basis states. This implies, in particular, that BDS are mapped to BDS.

Lemma 15. *In the canonical encoding for $N = 2$ and d prime, $\forall k, l \{T_C^{AB} \otimes \mathbb{1} |\Omega_{k,l}\rangle \mid C \in \mathcal{C}(s)\}$ is a basis of $\mathcal{H}_A \otimes \mathcal{H}_B$.*

Proof. By Lemma 11, $|\mathcal{C}(s)| = d^2$. Orthonormality is shown by Proposition 10 (ii), by noting that Lemma 13 and 14 imply $\frac{1}{d} \text{Tr}(T_C) \propto \frac{1}{d} \text{Tr}(W_{t, -s_1}) = \delta_{(t, s_1), (0, 0)} W_{t, -s_1} = \delta_{C, G_S}$. $\square$

4.3 The fidelity increase maximizing distillation protocol “FIMAX”

Based on the standard form for stabilizer distillation protocols with Bell-diagonal input states (Theorem 8) and the properties of the canonical encoding (Definition 9), a distillation protocol is proposed that maximizes the increase in fidelity for each iteration.

Protocol: Fidelity Increase Maximizing Distillation Protocol (**FIMAX**)

Let ρ_{in} be a two-copy Bell-diagonal state ($N = 2$) for prime dimension d .

1. For each stabilizer S with generator $W(g)$, $g \in \mathbb{Z}_d^2 \times \mathbb{Z}_d^2$:
 - (i) Partition the error elements e according to their symplectic product $s = \langle g, e \rangle$ and calculate $\mathbb{P}(\mathcal{E}(s))$.
 - (ii) For each s and for each error coset $C \in \mathcal{C}(s)$ determine $(C_{max}, s_{max}) := \arg \max \frac{\mathbb{P}(C)}{\mathbb{P}(\mathcal{E}(s))}$.
2. Choose the stabilizer S_{max} , maximizing $\frac{\mathbb{P}(C_{max})}{\mathbb{P}(\mathcal{E}(s_{max}))}$ among all stabilizers.
3. Alice and Bob perform stabilizer measurements for S_{max} with measurement outcomes a, b .
4. Bob sends b to Alice. Alice declares failure of the protocol if $s_{max} \neq a - b$.
5. Alice and Bob apply the inverse of the canonical encoding U_c and U_c^* for S_{max} and S_{max}^* , respectively.
6. Alice and Bob discard the first qudit and Alice applies the unique $W_{k_{max}, l_{max}}^\dagger \propto T_{C_{max}}^\dagger$ to the remaining qudit.

One successful iteration of the protocol requires *one-way* classical communication, but for further iterations, *two-way* communication is required for both parties to independently determine S_{max} and C_{max} from the input state. Also note that the protocol is applicable to non-BDS states in two ways. First, by using the diagonal elements of the density matrix in the Bell basis as probabilities, $\mathbb{p}(e) := \rho(e, e)$, to choose S_{max} and C_{max} to apply the remaining protocol. Second, any non-BDS can be transformed to a BDS by twirling, e.g., by a “Weyl twirl” [29], leaving all diagonal elements invariant, or a depolarizing unitary twirl [13]. By performing such a twirl, the remaining protocol can be applied as described above. In those cases, it is not guaranteed that FIMAX always achieves the maximal increase in the fidelity for each iteration.

It remains to prove the eponymous property (cf. Theorem 18) of FIMAX. The relation of all stabilizer encodings established in Section 3.1 is used in Lemma 16 to show that the coset fidelities are probabilities. Then, the canonical encoding is shown to imply an optimal distribution of these probabilities that allows to obtain a maximal increase in fidelity (Proposition 17). These results are combined to show the maximal fidelity increase in Theorem 18.

Lemma 16. *Let U be any encoding of a stabilizer S for $N = 2$ and prime dimension d . Let R_C be the coset actions for that encoding with $C \in \mathcal{C}(s)$ and $f_C(k, l) \equiv |\text{Tr}(\frac{1}{d} W_{k,l}^\dagger R_C)|^2$. $\forall (k, l)$, we have $\sum_{C \in \mathcal{C}(s)} f_C(k, l) = 1$.*

Proof. Denoting the coset action of the canonical encoding by T_C , Proposition 6 (ii) implies $R_C = Y_1^\dagger T_C Y_2$ for some unitaries Y_1, Y_2 . Using the identity (23) and the cyclic property of the trace, we can

write

$$f_C(k, l) = |\text{Tr}(\frac{1}{d} Y_2 W_{k,l}^\dagger Y_1^\dagger T_C)|^2 = \langle \Omega_{0,0} | (T_C^\dagger \otimes \mathbb{1}) \sigma (T_C \otimes \mathbb{1}) | \Omega_{0,0} \rangle$$

with the quantum state $\sigma \equiv (Y_2 W_{k,l} Y_1^\dagger \otimes \mathbb{1}) | \Omega_{0,0} \rangle \langle \Omega_{0,0} | (Y_1 W_{k,l}^\dagger Y_2^\dagger \otimes \mathbb{1})$. Lemma 15 implies that $T_C \otimes \mathbb{1} | \Omega_{0,0} \rangle$ are basis states and thus $\sum_{C \in \mathcal{C}(s)} \langle \Omega_{0,0} | (T_C^\dagger \otimes \mathbb{1}) \sigma (T_C \otimes \mathbb{1}) | \Omega_{0,0} \rangle = \text{Tr}(\sigma) = 1$. $\square$

Proposition 17. *Let S be a stabilizer for $N = 2$ and prime d . Let U_c be the canonical encoding and V be another arbitrary encoding. Denote the output state fidelities as $\mathcal{F}^{U_c}(k, l)$ and $\mathcal{F}^V(k, l)$, respectively. Then $\exists(k_{\max}, l_{\max})$ such that $\mathcal{F}^{U_c}(k_{\max}, l_{\max}) \geq \mathcal{F}^V(k, l) \forall(k, l)$.*

Proof. With Corollary 9 and Lemma 16, we have $f_C(k, l) \geq 0$, $\sum_C f_C(k, l) = 1$ and $\mathbb{P}(C) \geq 0$ for all C . Therefore, the Karush-Kuhn-Tucker conditions [30, 31] are satisfied. In consequence, $\mathcal{F}^V(k, l) = \sum_{C \in \mathcal{C}(s)} \frac{\mathbb{P}(C)}{\mathbb{P}(\mathcal{E}(s))} f_C(k, l) \leq \frac{\mathbb{P}(C_{\max})}{\mathbb{P}(\mathcal{E}(s))} \forall V \forall(k, l)$, where $C_{\max} = \arg \max \mathbb{P}(C)$ is the error coset with maximum probability. Consider the canonical encoding U_c . By Lemma 13, the corresponding coset actions are of the form $T_C \propto W_{t(C), -s_1(C)}$. Define $(k_{\max}, l_{\max}) := (t(C_{\max}), -s_1(C_{\max}))$ and thus $T_{C_{\max}} \propto W_{k_{\max}, l_{\max}}$. The definition of f_C in Corollary 9 and Proposition 10 (i) imply $f_{C_{\max}}(k, l) = \delta_{(k,l), (k_{\max}, l_{\max})}$ and thus $\mathcal{F}^{U_c}(k_{\max}, l_{\max}) = \frac{\mathbb{P}(C_{\max})}{\mathbb{P}(\mathcal{E}(s))} \geq \mathcal{F}^V(k, l) \forall V \forall(k, l)$. $\square$

Theorem 18. *Let ρ_{in} be a Bell-diagonal state of prime dimension d . Among all two-copy stabilizer-based distillation protocols, the FIMAX protocol maximizes the increase in the fidelity for a single iteration.*

Proof. Using the standard form 8, Proposition 17 proves that the canonical encoding for a given stabilizer maximizes the fidelity gain for a single iteration of a stabilizer protocol among all encodings and that the achievable fidelity is precisely $\frac{\mathbb{P}(C)}{\mathbb{P}(\mathcal{E}(s))}$. Therefore, using the stabilizer that maximizes this quantity with the canonical encoding for specific measurement outcomes implies the maximal fidelity between the output state and some Bell state $|\Omega_{k,l}\rangle$. Applying the inverse of the corresponding Weyl operator $W_{k_{\max}, l_{\max}}^\dagger$ therefore maximizes the fidelity between the output state and $|\Omega_{0,0}\rangle$. $\square$

We close this section with a brief comment on the computational complexity on the application of FIMAX. Given the probability distribution on all error elements for $N = 2$ via the Bell-diagonal input state, the protocol requires determining all nontrivial stabilizers $S \neq \{\mathbb{1}\}, S \neq \mathcal{E}_2$, which is equivalent to finding all subgroups of $\mathcal{Z}_d \times \mathcal{Z}_d$. Since d is prime, all subgroups are cyclic, each $e \in \mathcal{Z}_d \times \mathcal{Z}_d, e \neq (\vec{0}, \vec{0})$ generates a subgroup of d elements and each element is part of only one subgroup. Each subgroup contains $d - 1$ elements in addition to the neutral element $(\vec{0}, \vec{0})$. Since the subgroups are disjoint up to the neutral element, the combined number of elements contained in n_S subgroups is $n_S(d - 1) + 1$. Equating this to the total number of d^4 elements in $\mathcal{E}_2$, we conclude that there are $n_S = (d^2 + 1)(d + 1)$ stabilizers. For each stabilizer, there are d^3 distinct cosets. Consequently, to find the stabilizer and coset maximizing $\frac{\mathbb{P}(C)}{\mathbb{P}(\mathcal{E}(s))}$ in the second step of FIMAX, $d^3(d^2 + 1)(d + 1)$ probabilities have to be calculated in each iteration. This implies that the complexity for FIMAX is of polynomial order in d . Note that this only holds for a fixed number of copies $N = 2$. For more general stabilizer distillation protocols, the number of stabilizers to consider generally grows exponentially with the number of used copies N .

4.4 Efficacy of the FIMAX protocol and comparison to other protocols

In this section, the efficacy of the FIMAX protocol is demonstrated. The proposed protocol is compared to well-established two-copy recurrence protocols regarding the minimal required fidelity and the protocol efficiency. We compare it to the generalization of the BBPSSW protocol [7, 13], the DEJMPS protocol [9] ($d = 2$ only) and the ADGJ protocol [12] to d -level systems. In addition, we compare to the so-called “P1-or-P2” protocol (here named “P12” protocol) [24], known to outperform the BBPSSW and ADGJ protocol in efficiency and minimal required fidelity for $d = 3$. Note, that these analyses are intended to show the potential of stabilizer-based distillation without providing a complete evaluation regarding its performance in general settings. Additional numerical analyses regarding the performance of FIMAX can be found in [32]. All applied methods are implemented as open source software [33].

First, we compare the distillation efficiency in dimensions $d = 2$ and $d = 3$. Given a target fidelity, the efficiency is defined as the inverse of the expected number of input states required to produce one output pair with fidelity larger than the target fidelity. Using two copies for each iteration with success probability p_i and requiring N_{it} iterations to reach the target fidelity, the efficiency is $2^{-N_{it}} \prod_i p_i$. Here, we choose a target fidelity of 0.999. If the target fidelity cannot be reached, the efficiency of the protocol is zero. In Figure 2 we analyze isotropic states. This family is defined as mixtures of the target state with the maximally mixed state π_{mm} , i.e., $\rho_{iso}(p) := p |\Omega_{0,0}\rangle\langle\Omega_{0,0}| + (1 - p) \pi_{mm}$. The proposed FIMAX protocol can distill all states with fidelity $> 1/d$ and is more efficient in wide ranges of initial fidelity than the other protocols (except for the DEJMPS protocol in $d = 2$, which has the same efficiency). Especially in the low fidelity regime, the efficiency of FIMAX protocol can be more than a magnitude higher than for each of the other protocols in $d = 3$. No fidelities are observed for which the efficiency of FIMAX is lower than any of the other protocols.

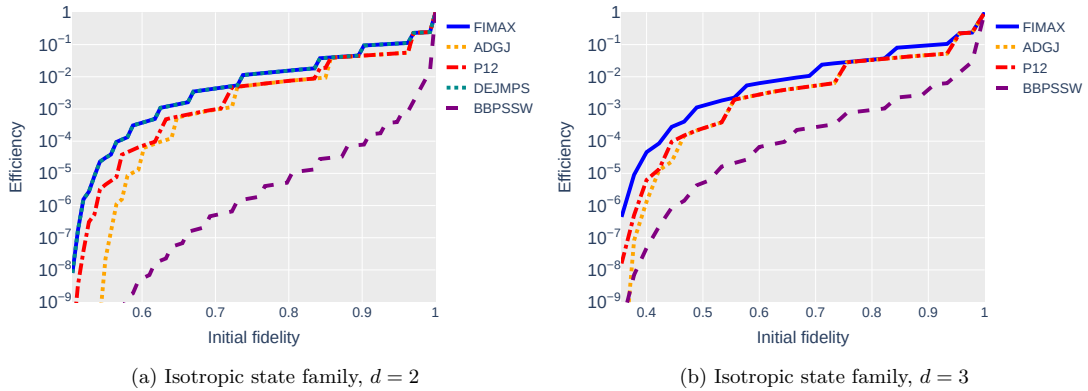


Figure 2: Protocol comparison of distillation efficiencies depending on the fidelity of isotropic input states.

The mean efficiency of randomly generated pure states that are grouped by their fidelity is depicted in Figure 3. We show the efficiency in the fidelity range that allows sampling random pure states with appropriate computational effort. Note, that it is very unlikely to sample uniformly distributed states with fidelity higher than a certain value, depending on the dimension. We choose the fidelity ranges to be $[0, 0.9]$ for $d = 2$ and $[0, 0.6]$ for $d = 3$. Within these ranges, we sample bins of 1000 states, where each bin corresponds to a unique fidelity value, rounded to two digits, and calculate the mean efficiency. To estimate the numerical error given the limited number of samples, we calculate the

standard deviation for the efficiency σ_{bin} in each bin. The error for the mean can then be estimated as $\sigma_{bin}/\sqrt{1000}$. Relative to the mean, the maximum error for all protocols and all fidelity bins is $< 1\%$ for $d = 2$ and $< 25\%$ for $d = 3$. Notably, this error is significantly higher for $d = 3$. FIMAX is applied with prior twirl to obtain Bell-diagonal states. Interestingly, FIMAX again performs best despite the additional twirling operation. Fidelity regions below $1/d$ are visible, in which the proposed protocol is the only one capable of distillation for the limited set of 1000 analyzed states per bin.

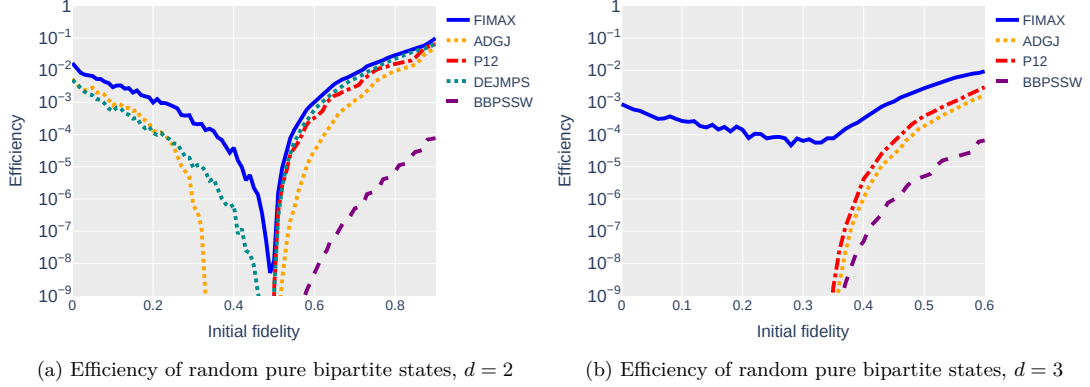


Figure 3: Efficiency comparison depending on the fidelity. States are grouped in bins of 1000 states by their fidelity, rounded to two digits, and the mean efficiency for all protocols is determined.

In Figure 4, the effect of the protocols for specific low-fidelity isotropic states in $d = 2$ and $d = 3$ is visualized. For $d = 2$, FIMAX and DEJMPS have equal fidelity increase for this state. ADGJ fails to distill, while BBPSSW generally increases the fidelity, but less than FIMAX. P12 has iterations, which do not significantly increase the fidelity. FIMAX reaches the target fidelity with the least number of iterations.

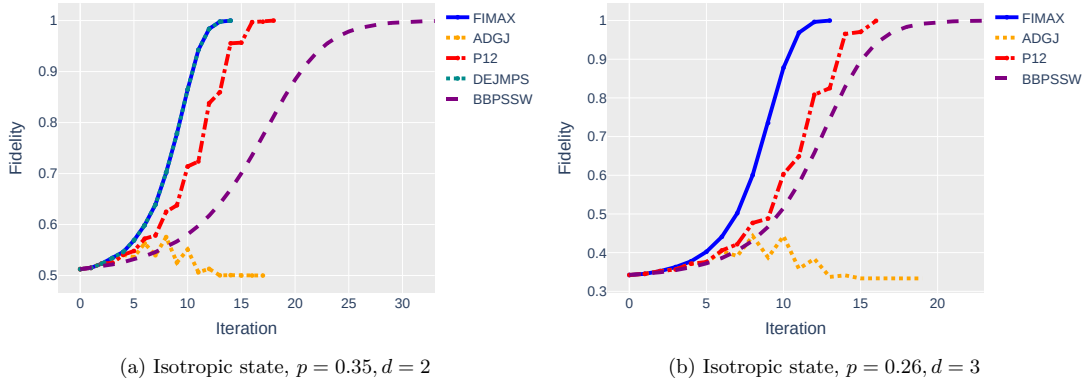


Figure 4: Protocol comparison of the iterative fidelity increase for a low fidelity isotropic input state.

Finally, Figure 5 demonstrates that FIMAX can distill states with fidelity $< 1/d$ with high efficiency. For $d = 3$, we define the states $\rho_{ol}(p) := p \sigma + (1 - p) \pi_{mm}$ with $\sigma := 1/3(|\Omega_{0,0}\rangle\langle\Omega_{0,0}| + |\Omega_{1,0}\rangle\langle\Omega_{1,0}| + |\Omega_{0,1}\rangle\langle\Omega_{0,1}|)$, the so-called “off-line states”. All of these states have fidelity $\leq 1/d$ and none of the other protocols can distill any state of this family. Figure 5(a) demonstrates that the FIMAX protocol can distill all off-line states with initial fidelity > 0.25 . Figure 5(b) shows the iterative increase in

fidelity together with the probability of success for each iteration. Interestingly, the protocol increases the fidelity in the first iteration to a value $> 1/d$.

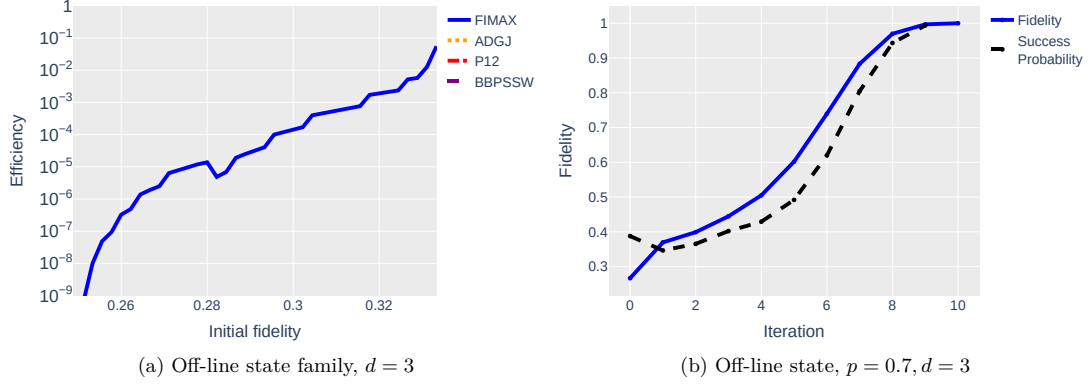


Figure 5: Distillation efficiency for off-line states in dependence on the fidelity (a). Fidelities and success probabilities for each FIMAX iteration for a low-fidelity off-line input state (b).

5 Discussion and Conclusion

In this work, we analyzed the action of the stabilizer-based distillation procedure in prime dimension to derive a standard form of the output state, making the effect of the adjustable parameters of the protocol transparent. We leveraged this standard form to propose FIMAX, a fidelity increase maximizing distillation protocol, that demonstrates superior efficacy compared to other well-established protocols regarding efficiency and minimal fidelity requirements.

It was shown how the effective action of Weyl-Heisenberg errors depends on the chosen codewords and how all encodings are analytically related. The group properties of the Weyl-Heisenberg errors were extended to the so-called error action operators, representing the effective action of a Weyl-Heisenberg error in the stabilizer protocol. It was further demonstrated how the stabilizer implies a decomposition of the set of errors given by its cosets and the outcomes of stabilizer measurements. Combining those general insights, we derived a standard form of the output state, rendering the role of input state, stabilizer, encoding, and measurement evident and making the calculation of all protocol output fidelities possible.

Focusing on two-copy stabilizer distillation in prime dimension, we introduced a canonical encoding, for which we find that the effective actions of errors are again of Weyl-Heisenberg form. Leveraging the standard form and the properties of the canonical encoding, we proposed the distillation protocol FIMAX and proved that among all two-copy stabilizer protocols in prime dimension, it implies the maximal increase in fidelity for Bell-diagonal states in each iteration.

Finally, we compared the new protocol to prominent recurrence protocols, namely BBPSSW [7, 13], DEJMPS [9], ADGJ [12] and “P1-or-P2/P12” [24], that have been shown to have good efficiency and also allow for the distillation of low-fidelity states. FIMAX demonstrated the best results regarding distillation efficiency and disability of both Bell-diagonal and, curiously, also non-Bell-diagonal states in all numerical investigations. Due to the limited number of samples and state families, these results do not prove general superiority, requiring more analyses with higher sample sizes, especially for $d \geq 3$. However, the reported results clearly indicate the potential of the developed formalism and the FIMAX protocol. Further results confirming this potential with focus on entanglement distillation of low-fidelity states can be found in [32].

The developed theory of stabilizer-based entanglement distillation aims to enable future research in the construction of new distillation protocols and in the general problem of distillability of mixed states. The successful application to the two-copy case in prime dimension illustrates how the derived standard form helps to develop stabilizer protocols and analyze their properties. Many existing protocols, including BBPSSW, DEJMPS and P12, are equivalent or strongly related to a specific stabilizer protocol (see, e.g., [21]). Interestingly, another generalization of such recurrence-type protocols has been suggested, so-called permutation-based schemes [15, 34]. Both approaches are related by their symplectic structure manifesting in the investigated properties of stabilizers and their encodings on the one side, and in the form of permutation matrices that correspond to local operations on the other side. The standard stabilizer and permutation protocols have been shown to be equivalent regarding the output fidelity in the case of Bell-diagonal input states for $d = 2$ [35]. For general dimension d , however, this equivalence is not expected. All permutation-based protocols map the set of Bell-diagonal states onto itself. However, with the presented results, one easily finds stabilizer codes that imply a mapping to non-Bell diagonal states. Conversely, it is unclear whether every permutation protocol can be realized by a stabilizer protocol with suitable encoding. This would imply that the class of stabilizer-distillation schemes is strictly larger than the permutation-based one. Further research in

this direction could contribute to the construction of optimal protocols for specific state families. Interestingly, the proposed FIMAX protocol does map the set of Bell-diagonal states onto itself, implying that the canonical encoding is part of the Clifford group and therefore can be efficiently constructed with quantum gates for $d = 2$ [23]. Whether this also holds for $d \geq 3$ remains an open question for future research. Extending the developed methods to the non-prime dimensional regime also provides an interesting challenge for the future. While generalization to prime-power dimensions may be possible by following the theory of nonbinary quantum stabilizer codes [17], other dimensions may be challenging due to the more complicated spectral properties of corresponding operators and their group structure.

The numerical results regarding the performance of FIMAX clearly demonstrate that the developed stabilizer approach offers great potential for effective distillation. While further investigations are needed for a general performance evaluation of FIMAX, the results clearly indicate high efficiency compared to the other protocols for certain state families. Interestingly, FIMAX also exhibits strong performance for pure states if a twirl to Bell-diagonal form is prepended. This is surprising, as such an operation generally reduces the fidelity due to its data processing inequality [36]. Investigating the impact of twirling on distillability and conducting a comparative analysis of performance relative to a protocol executed without prior twirling constitutes an intriguing subject for future research. Notable performance is especially evident in the distillation of low-fidelity Bell-diagonal states, as also recently affirmed [32]. These findings indicate the potential utility of the presented approach in addressing the broader challenges of general distillability and the phenomenon of bound entanglement [6]. Related research directions include further development of the theory, protocols and numerical investigations to the multi-copy, non-Bell-diagonal and non-prime dimensional regimes.

References

1. Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters* **70**. Publisher: American Physical Society, 1895–1899. DOI: [10.1103/PhysRevLett.70.1895](https://link.aps.org/doi/10.1103/PhysRevLett.70.1895). <https://link.aps.org/doi/10.1103/PhysRevLett.70.1895> (2024) (Mar. 1993).
2. Bennett, C. H. & Wiesner, S. J. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters* **69**. Publisher: American Physical Society, 2881–2884. DOI: [10.1103/PhysRevLett.69.2881](https://link.aps.org/doi/10.1103/PhysRevLett.69.2881). <https://link.aps.org/doi/10.1103/PhysRevLett.69.2881> (2024) (Nov. 1992).
3. Werner, R. F. All teleportation and dense coding schemes. en. *Journal of Physics A: Mathematical and General* **34**, 7081. ISSN: 0305-4470. DOI: [10.1088/0305-4470/34/35/332](https://dx.doi.org/10.1088/0305-4470/34/35/332). <https://dx.doi.org/10.1088/0305-4470/34/35/332> (2024) (Aug. 2001).
4. Briegel, H. J., Browne, D. E., Dür, W., Raussendorf, R. & Van den Nest, M. Measurement-based quantum computation. en. *Nature Physics* **5**. Publisher: Nature Publishing Group, 19–26. ISSN: 1745-2481. DOI: [10.1038/nphys1157](https://www.nature.com/articles/nphys1157). <https://www.nature.com/articles/nphys1157> (2024) (Jan. 2009).
5. Horodecki, M., Horodecki, P. & Horodecki, R. Mixed-State Entanglement and Distillation: Is there a “Bound” Entanglement in Nature? en. *Physical Review Letters* **80**, 5239–5242. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.80.5239](https://link.aps.org/doi/10.1103/PhysRevLett.80.5239). <https://link.aps.org/doi/10.1103/PhysRevLett.80.5239> (2024) (June 1998).
6. Hiesmayr, B. C., Popp, C. & Sutter, T. C. Bipartite bound entanglement. *International Journal of Quantum Information* **23**. Publisher: World Scientific Publishing Co., 2530003. ISSN: 0219-7499. DOI: [10.1142/S0219749925300037](https://www.worldscientific.com/doi/10.1142/S0219749925300037). <https://www.worldscientific.com/doi/10.1142/S0219749925300037> (2025) (Aug. 2025).
7. Bennett, C. H. *et al.* Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. en. *Physical Review Letters* **76**, 722–725. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.76.722](https://link.aps.org/doi/10.1103/PhysRevLett.76.722). <https://link.aps.org/doi/10.1103/PhysRevLett.76.722> (2024) (Jan. 1996).
8. Bennett, C. H., Bernstein, H. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. en. *Physical Review A* **53**, 2046–2052. ISSN: 1050-2947, 1094-1622. DOI: [10.1103/PhysRevA.53.2046](https://link.aps.org/doi/10.1103/PhysRevA.53.2046). <https://link.aps.org/doi/10.1103/PhysRevA.53.2046> (2024) (Apr. 1996).
9. Deutsch, D. *et al.* Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels. *Physical Review Letters* **77**. Publisher: American Physical Society, 2818–2821. DOI: [10.1103/PhysRevLett.77.2818](https://link.aps.org/doi/10.1103/PhysRevLett.77.2818). <https://link.aps.org/doi/10.1103/PhysRevLett.77.2818> (2024) (Sept. 1996).
10. Yan, P.-S., Zhou, L., Zhong, W. & Sheng, Y.-B. Measurement-based logical qubit entanglement purification. *Physical Review A* **105**. Publisher: American Physical Society, 062418. DOI: [10.1103/PhysRevA.105.062418](https://link.aps.org/doi/10.1103/PhysRevA.105.062418). <https://link.aps.org/doi/10.1103/PhysRevA.105.062418> (2024) (June 2022).

11. Zhou, L., Zhong, W. & Sheng, Y.-B. Purification of the residual entanglement. EN. *Optics Express* **28**. Publisher: Optica Publishing Group, 2291–2301. ISSN: 1094-4087. DOI: [10.1364/OE.383499](https://doi.org/10.1364/OE.383499). <https://opg.optica.org/oe/abstract.cfm?uri=oe-28-2-2291> (2024) (Jan. 2020).
12. Alber, G., Delgado, A., Gisin, N. & Jex, I. Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces. en. *Journal of Physics A: Mathematical and General* **34**, 8821. ISSN: 0305-4470. DOI: [10.1088/0305-4470/34/42/307](https://doi.org/10.1088/0305-4470/34/42/307). <https://doi.org/10.1088/0305-4470/34/42/307> (2025) (Oct. 2001).
13. Horodecki, M. & Horodecki, P. Reduction criterion of separability and limits for a class of distillation protocols. *Physical Review A* **59**. Publisher: American Physical Society, 4206–4216. DOI: [10.1103/PhysRevA.59.4206](https://link.aps.org/doi/10.1103/PhysRevA.59.4206). <https://link.aps.org/doi/10.1103/PhysRevA.59.4206> (2024) (June 1999).
14. Vollbrecht, K. G. H. & Wolf, M. M. Efficient distillation beyond qubits. en. *Physical Review A* **67**, 012303. ISSN: 1050-2947, 1094-1622. DOI: [10.1103/PhysRevA.67.012303](https://link.aps.org/doi/10.1103/PhysRevA.67.012303). <https://link.aps.org/doi/10.1103/PhysRevA.67.012303> (2024) (Jan. 2003).
15. Dehaene, J., Van Den Nest, M., De Moor, B. & Verstraete, F. Local permutations of products of Bell states and entanglement distillation. en. *Physical Review A* **67**, 022310. ISSN: 1050-2947, 1094-1622. DOI: [10.1103/PhysRevA.67.022310](https://link.aps.org/doi/10.1103/PhysRevA.67.022310). <https://link.aps.org/doi/10.1103/PhysRevA.67.022310> (2024) (Feb. 2003).
16. Gottesman, D. Stabilizer Codes and Quantum Error Correction. en. arXiv:quant-ph/9705052. DOI: <https://doi.org/10.48550/arXiv.quant-ph/9705052>. <http://arxiv.org/abs/quant-ph/9705052> (2024) (May 1997).
17. Ashikhmin, A. & Knill, E. Nonbinary quantum stabilizer codes. *IEEE Transactions on Information Theory* **47**, 3065–3072. ISSN: 1557-9654. DOI: [10.1109/18.959288](https://ieeexplore.ieee.org/document/959288). <https://ieeexplore.ieee.org/document/959288> (2025) (Nov. 2001).
18. Wilde, M. M. *Quantum Coding with Entanglement* en. arXiv:0806.4214 [quant-ph]. June 2008. DOI: <https://doi.org/10.48550/arXiv.0806.4214>. <http://arxiv.org/abs/0806.4214> (2024).
19. Dür, W. & Briegel, H.-J. Entanglement Purification for Quantum Computation. en. *Physical Review Letters* **90**, 067901. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.90.067901](https://link.aps.org/doi/10.1103/PhysRevLett.90.067901). <https://link.aps.org/doi/10.1103/PhysRevLett.90.067901> (2024) (Feb. 2003).
20. Dür, W. & Briegel, H. J. Entanglement purification and quantum error correction. en. *Reports on Progress in Physics* **70**, 1381. ISSN: 0034-4885. DOI: [10.1088/0034-4885/70/8/R03](https://dx.doi.org/10.1088/0034-4885/70/8/R03). <https://dx.doi.org/10.1088/0034-4885/70/8/R03> (2024) (July 2007).
21. Matsumoto, R. Conversion of a general quantum stabilizer code to an entanglement distillation protocol. en. *Journal of Physics A: Mathematical and General* **36**. arXiv:quant-ph/0209091, 8113–8127. ISSN: 0305-4470, 1361-6447. DOI: [10.1088/0305-4470/36/29/316](https://arxiv.org/abs/quant-ph/0209091). [http://arxiv.org/abs/quant-ph/0209091](https://arxiv.org/abs/quant-ph/0209091) (2024) (July 2003).
22. Matsumoto, R. *Breeding protocols are advantageous for finite-length entanglement distillation* arXiv:2401.02265 [quant-ph]. Feb. 2024. DOI: [10.48550/arXiv.2401.02265](https://arxiv.org/abs/2401.02265). [http://arxiv.org/abs/2401.02265](https://arxiv.org/abs/2401.02265) (2025).

23. Watanabe, S., Matsumoto, R. & Uyematsu, T. Improvement of stabilizer-based entanglement distillation protocols by encoding operators. en. *Journal of Physics A: Mathematical and General* **39**, 4273. ISSN: 0305-4470. DOI: [10.1088/0305-4470/39/16/013](https://dx.doi.org/10.1088/0305-4470/39/16/013). <https://dx.doi.org/10.1088/0305-4470/39/16/013> (2024) (Mar. 2006).
24. Miguel-Ramiro, J. & Dür, W. Efficient entanglement purification protocols for d -level systems. en. *Physical Review A* **98**, 042309. ISSN: 2469-9926, 2469-9934. DOI: [10.1103/PhysRevA.98.042309](https://link.aps.org/doi/10.1103/PhysRevA.98.042309). <https://link.aps.org/doi/10.1103/PhysRevA.98.042309> (2024) (Oct. 2018).
25. Knill, E. *Non-binary unitary error bases and quantum codes* en. Tech. rep. LA-UR-96-2717, 373768 (June 1996), LA-UR-96-2717, 373768. DOI: [10.2172/373768](http://www.osti.gov/servlets/purl/373768-BfsVyz/webviewable/). <http://www.osti.gov/servlets/purl/373768-BfsVyz/webviewable/> (2024).
26. Rains, E. Nonbinary quantum codes. *IEEE Transactions on Information Theory* **45**. Conference Name: IEEE Transactions on Information Theory, 1827–1832. ISSN: 1557-9654. DOI: [10.1109/18.782103](https://ieeexplore.ieee.org/document/782103). <https://ieeexplore.ieee.org/document/782103> (2024) (Sept. 1999).
27. Baumgartner, B., Hiesmayr, B. & Narnhofer, H. A special simplex in the state space for entangled qudits. en. *Journal of Physics A: Mathematical and Theoretical* **40**. arXiv:quant-ph/0610100, 7919–7938. ISSN: 1751-8113, 1751-8121. DOI: [10.1088/1751-8113/40/28/S03](http://arxiv.org/abs/quant-ph/0610100). <http://arxiv.org/abs/quant-ph/0610100> (2024) (July 2007).
28. Popp, C. & Hiesmayr, B. C. Comparing bound entanglement of bell diagonal pairs of qutrits and ququarts. en. *Scientific Reports* **13**. Number: 1 Publisher: Nature Publishing Group, 2037. ISSN: 2045-2322. DOI: [10.1038/s41598-023-29211-w](https://www.nature.com/articles/s41598-023-29211-w). <https://www.nature.com/articles/s41598-023-29211-w> (2024) (Feb. 2023).
29. Popp, C. & Hiesmayr, B. C. Special features of the Weyl–Heisenberg Bell basis imply unusual entanglement structure of Bell-diagonal states. en. *New Journal of Physics* **26**. Publisher: IOP Publishing, 013039. ISSN: 1367-2630. DOI: [10.1088/1367-2630/ad1d0e](https://dx.doi.org/10.1088/1367-2630/ad1d0e). <https://dx.doi.org/10.1088/1367-2630/ad1d0e> (2024) (Jan. 2024).
30. Karush, W. *Minima of functions of several variables with inequalities as side conditions* OCLC: 43268508. PhD thesis (1939). <https://catalog.lib.uchicago.edu/vufind/Record/4111654> (2024).
31. Kuhn, H. W. & Tucker, A. W. en. in *Traces and Emergence of Nonlinear Programming* (eds Giorgi, G. & Kjeldsen, T. H.) 247–258 (Springer, Basel, 2014). ISBN: 978-3-0348-0439-4. DOI: [10.1007/978-3-0348-0439-4_11](https://doi.org/10.1007/978-3-0348-0439-4_11). https://doi.org/10.1007/978-3-0348-0439-4_11 (2025).
32. Popp, C., Sutter, T. C. & Hiesmayr, B. C. Low-fidelity entanglement distillation with FIMAX. *International Journal of Quantum Information* **23**. Publisher: World Scientific Publishing Co., 2550017. ISSN: 0219-7499. DOI: [10.1142/S0219749925500170](https://www.worldscientific.com/doi/10.1142/S0219749925500170). <https://www.worldscientific.com/doi/10.1142/S0219749925500170> (2025) (Sept. 2025).
33. Popp, C. BellDiagonalQudits: A package for entanglement analyses of mixed maximally entangled qudits. en. *Journal of Open Source Software* **8**, 4924. ISSN: 2475-9066. DOI: [10.21105/joss.04924](https://joss.theoj.org/papers/10.21105/joss.04924). <https://joss.theoj.org/papers/10.21105/joss.04924> (2024) (Jan. 2023).
34. Bombin, H. & Martin-Delgado, M. A. Entanglement distillation protocols and number theory. en. *Physical Review A* **72**, 032313. ISSN: 1050-2947, 1094-1622. DOI: [10.1103/PhysRevA.72.032313](https://link.aps.org/doi/10.1103/PhysRevA.72.032313). <https://link.aps.org/doi/10.1103/PhysRevA.72.032313> (2024) (Sept. 2005).

35. Hostens, E., Dehaene, J. & De Moor, B. *The equivalence of two approaches to the design of entanglement distillation protocols* arXiv:quant-ph/0406017. June 2004. DOI: [10.48550/arXiv.quant-ph/0406017](https://doi.org/10.48550/arXiv.quant-ph/0406017). <http://arxiv.org/abs/quant-ph/0406017> (2024).
36. Khatri, S. & Wilde, M. M. *Principles of Quantum Communication Theory: A Modern Approach* arXiv:2011.04672 [cond-mat, physics:hep-th, physics:math-ph, physics:quant-ph]. Feb. 2024. DOI: [10.48550/arXiv.2011.04672](https://doi.org/10.48550/arXiv.2011.04672). <http://arxiv.org/abs/2011.04672> (2024).

Acknowledgments

C.P. and B.C.H. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) project P36102-N (Grant DOI: 10.55776/P36102). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Author Contributions Statement

C.P. developed the methods, carried out the analytical and numerical analyses, implemented the software and edited the manuscript. B.C.H. and T.C.S. revised the analyses and proposed improvements.

Competing Interests

All authors declare no financial or non-financial competing interests.

Data availability statement

All analyzed datasets were generated during the current study and are available from the corresponding author on reasonable request.

Code availability statement

The software used to generate the reported results is published as a repository and registered open source package “BellDiagonalQudits.jl” [33] available at <https://github.com/kungfugo/BellDiagonalQudits.jl>.

Additional information

Correspondence and requests for materials should be addressed to C.P..

A Example: Stabilizer Objects and Action Operators

This appendix aims to provide a better understanding of the concepts introduced in Sections 2 and 3 by giving a specific example.

Let $d = 3$ and $N = 2$, $\omega = e^{2\pi i/3}$. In this case, the group of Weyl errors reads:

$$\mathcal{E}_2 = \{W(e) \mid e \in \mathcal{Z}_3^2 \times \mathcal{Z}_3^2\} = \{W_{k_1, l_1} \otimes W_{k_2, l_2} \mid k_1, l_1, k_2, l_2 \in \mathcal{Z}_3\},$$

where we write the error elements $e \in \mathcal{Z}_3^2 \times \mathcal{Z}_3^2$ in the form $e = \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right]$.

Consider the following stabilizer $S \subset \mathcal{E}_2$ generated by the generator $W(g) = W_{1,0} \otimes W_{1,0}$ and the corresponding subgroup of error elements $G_S \subset \mathcal{Z}_3^2 \times \mathcal{Z}_3^2$ with generating element $g = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]$:

$$\begin{aligned} S &= \{W(0), W(g), W(2g)\} = \{\mathbb{1}_3 \otimes \mathbb{1}_3, W_{1,0} \otimes W_{1,0}, W_{2,0} \otimes W_{2,0}\}, \\ G_S &= \{0, g, 2g\} = \left\{ \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right], \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right], \left[\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right] \right\}, \end{aligned}$$

Since the stabilizer is generated by one generator, we have $p = 1$.

The generator implies the decomposition of errors into cosets $C(e) = e + G_S$ and according to the symplectic product $\langle g, e \rangle$, reading $\mathcal{E}(s) = \left\{ e = \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right] \mid \langle g, e \rangle = -l_1 - l_2 = s \right\}$.

The eigenvalues of $W(g)$ are $\{\omega^x \mid x \in \mathcal{Z}_3\}$, each threefold degenerated. The corresponding codespaces of $\mathcal{H}_A^{\otimes 2}$, $\mathcal{Q}(x) = \{|\phi\rangle \in \mathcal{H}_A^{\otimes 2} \mid W(g)|\phi\rangle = \omega^x|\phi\rangle\}$, have dimension $d^{N-p} = d$. We can define an encoding for this stabilizer by the mapping

$$U := |x\rangle \otimes |k\rangle \mapsto |u_{x,k}\rangle := |k\rangle \otimes |x - k\rangle,$$

defining an orthonormal basis of eigenstates of $W(g)$, i.e., codewords with $|u_{x,k}\rangle \in \mathcal{Q}(x) \forall x, k \in \mathcal{Z}_3$. Since for this choice of stabilizer S , the stabilizer with complex conjugated elements S^* is identical to S , also the codespaces of $\mathcal{H}_B^{\otimes 2}$ and corresponding encoding/codewords can be defined in this form.

We proceed by determining the effective error action operators $T_x^{U,e}$. Let $e = \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right]$. Consider

$$\begin{aligned} U^\dagger W(e) U (|b\rangle \otimes |j\rangle) &= U^\dagger W_{k_1, l_1} \otimes W_{k_2, l_2} (|j\rangle \otimes |b - j\rangle) = U^\dagger \omega^{k_1(j-l_1)} \omega^{k_2(b-j-l_2)} |j - l_1\rangle \otimes |b - j - l_2\rangle \\ &= \omega^{-k_1 l_1 - k_2 l_2} \omega^{k_2 b} \omega^{j(k_1 - k_2)} |b + (-l_1 - l_2)\rangle \otimes |j - l_1\rangle \\ &= \omega^{k_2(b-l_1-l_2)} |b + s\rangle \otimes W_{k_1-k_2, l_1} |j\rangle, \end{aligned}$$

where we used that $s = \langle g, e \rangle = -l_1 - l_2$. This shows $U^\dagger W(e) U = \sum_{x \in \mathcal{Z}_3} |x + s\rangle \langle x| \otimes T_{x+s}^{U,e}$ with

$$T_{x+s}^{U,e} = \omega^{k_2(x-l_1-l_2)} W_{k_1-k_2, l_1} \propto W_{k_1-k_2, l_1}.$$

Note that in this case, the action operators are Weyl operators. This is due to the fact, that the chosen encoding U is the canonical encoding.

B Example: FIMAX routine

Using the example of Appendix A, with $d = 3$, $N = 2$ and $g = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]$, we demonstrate the application of FIMAX for one iteration.

Let ρ_{in} be a Bell-diagonal input state $\rho_{in} = \sum_{k,l \in \mathcal{Z}_3} p_{k,l} |\Omega_{k,l}\rangle\langle\Omega_{k,l}|$, which can be represented in the Bell basis $\{|\Omega_{0,0}\rangle, |\Omega_{1,0}\rangle, \dots, |\Omega_{1,2}\rangle, |\Omega_{2,2}\rangle\}$ by its mixing probabilities $p_{k,l}$. In the case of two copies of ρ_{in} , the induced probability distribution for two copy errors $e = \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right] \in \mathcal{E}_2$ is given by $\mathbb{P}(e) = p_{k_1,l_1} p_{k_2,l_2}$. Consider the Bell-diagonal input state given by the following mixing probabilities (fidelities):

$$(p_{k,l})_{k,l \in \mathcal{Z}_3} = (0.06, 0.06, 0.06, 0.06, 0.06, \mathbf{0.56}, 0.06, 0.06, 0.06)$$

and the coset

$$C_{max} = C\left(\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]\right) = \left\{ \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right], \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right], \left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] \right\} \in \mathcal{C}(s_{max}) = \mathcal{C}(1).$$

Calculating the probabilities $\mathbb{P}(\mathcal{E}(s_{max})) = \sum_{e \in \mathcal{E}(1)} \mathbb{P}(e) = 0.5$ and similarly $\mathbb{P}(C_{max}) = \sum_{e \in C_{max}} \mathbb{P}(e) = 0.314$, one obtains $\frac{\mathbb{P}(C_{max})}{\mathbb{P}(\mathcal{E}(s_{max}))} = 0.63$, which is the highest value among all stabilizers and cosets. For this stabilizer, there are two other cosets taking the value 0.13 and the remaining six cosets take the value 0.02.

According to the example in Appendix A, for this stabilizer, the error action operators are $T_x^{U,e} \propto W_{k_1-k_2, l_1}$ for $e = \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right]$. According to Lemma 5, the same holds for all errors of the same coset $C(e)$, so we write shorthand $T_{C(e)} \propto W_{k_1-k_2, l_1}$. In the case of $C_{max} = C\left(\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]\right)$, this implies $T_{C_{max}} \propto W_{0,1}$.

Identifying the stabilizer generated by $W(g)$ and the coset C_{max} according to steps 1. and 2., Alice and Bob perform the stabilizer measurements in step 3., for which we assume outcomes a and b with $a - b = s_{max} = 1$ in step 4. After application of the inverse encoding by both parties, Alice finally applies $T_{C_{max}}^\dagger = W_{0,2}$ to her second qudit. Following this routine, the output state is again of Bell-diagonal form, given by the mixing probabilities

$$(\hat{p}_{k,l})_{k,l \in \mathcal{Z}_3} = (\mathbf{0.63}, 0.13, 0.13, 0.02, 0.02, 0.02, 0.02, 0.02, 0.02).$$

Note that the fidelity with the maximally entangled state $|\Omega_{0,0}\rangle$ is now precisely $\frac{\mathbb{P}(C_{max})}{\mathbb{P}(\mathcal{E}(s_{max}))} = 0.63$.

B.7. Low-fidelity entanglement distillation with FIMAX

Low-fidelity entanglement distillation with FIMAX

Christopher Popp ^{*}, Tobias C. Sutter [†] and Beatrix C. Hiesmayr [‡]

*University of Vienna, Faculty of Physics, Währingerstrasse 17
Vienna 1090, Austria*

**christopher.popp@univie.ac.at*

†tobias.christoph.sutter@univie.ac.at

‡beatrix.hiesmayr@univie.ac.at

Received 14 April 2025

Revised 24 June 2025

Accepted 9 July 2025

Published 4 September 2025

Uncontrolled interactions with the environment introduce errors that remain a significant challenge to the reliability of quantum technologies using entanglement. An essential method to overcome or mitigate these errors is entanglement distillation, the transformation of multiple copies of weakly entangled states into a smaller number of approximately maximally entangled states. We present a comparative analysis of the distillation effectivity for the stabilizer-based two-copy entanglement distillation protocol, FIMAX, against other recurrent two-copy protocols, including ADGJ, DEJMPS, P1–P2 and the generalized BBPSSW protocol. We focus on low-fidelity bipartite quantum states in dimensions $d = 2$ and $d = 3$, which are particularly challenging to distill. Our findings demonstrate that FIMAX exhibits superior performance regarding the distillable share for these states. While other protocols struggle with highly noisy states, FIMAX successfully distills entanglement even when the initial state quality is severely compromised. These results highlight the protocol's capability to address the effects of environmental noise, advancing the robustness and scalability of quantum technologies leveraging entanglement distillation.

Keywords: Entanglement distillation; stabilizer codes; qudits.

1. Introduction

Overcoming the pervasive challenge of environment-induced errors is crucial for quantum technologies to revolutionize applications like secure communication¹ and computation.²

*Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Quantum information processing tasks often require quantum-specific resources, such as entanglement and coherence, to function reliably.^{3,4} Affecting the entanglement structure,⁵ noise in prepared quantum states, operations, or measurements can degrade these resources and, consequently, the performance of reliant applications. Even in the fundamental case of a bipartite qudit system, a d^2 -dimensional quantum system shared between two parties, the correction or mitigation of errors is essential for realizing tasks on Noisy Intermediate-Scale Quantum (NISQ) devices.^{6–8} To tackle this challenge, two promising approaches have emerged: Quantum error correction⁹ and entanglement distillation.¹⁰ Error correction involves encoding logical information into states resilient to specific forms and levels of noise, enabling the detection and subsequent correction of errors. Entanglement distillation aims to extract pure, maximally entangled states from noisy shared states through local operations and classical operations, thereby facilitating reliable implementation of various quantum information processing tasks. Notably, these approaches are interconnected:¹¹ Any quantum error-correcting code can be utilized in entanglement distillation schemes, and conversely, maximally entangled states produced by any distillation protocol enable reliable quantum communication, e.g. via teleportation.¹² Recently, a highly efficient entanglement distillation protocol named FIMAX¹³ has been proposed, leveraging this relationship. Based on stabilizer codes, a class of quantum error correction codes,^{14,15} FIMAX exploits algebraic relations between these codes and maximally entangled states to efficiently use noisy bipartite quantum states for distillation. This protocol demonstrates significant promise in enhancing the fidelity, i.e. the overlap of used states with the maximally entangled state, even if the initial states are substantially affected by noise. The numerical analyses of Ref. 13 indicate a high *distillation efficiency* of FIMAX, i.e. the number of distilled maximally entangled states relative to the expected number of required input states. However, the analysis of *distillation effectivity*, i.e. the question of which states are distillable and which are not, remains an open question.

In this work, we address this question by performing a comprehensive analysis of the distillation capabilities of FIMAX regarding distillation effectivity and by comparing it to other recurrent entanglement distillation protocols that operate on two copies of the input state. We focus on entangled bipartite qudit states subjected to very high levels of noise, yet still exhibiting entanglement. Unlike hashing or breeding protocols,^{10,16} which typically require states with higher fidelities, recurrent distillation protocols like FIMAX can distill very noisy states, although their effectiveness is influenced by the specific characteristics of the noise. Note that there exist entangled, yet generally undistillable states, so-called bound entangled states¹⁷ (for a recent review, see Ref. 18). The proportion of entangled low-fidelity states that can be successfully distilled depends on the protocol applied, and is the main quantity of interest in this work.

This paper is organized as follows. In Sec. 2, we introduce entanglement distillation using FIMAX, emphasizing the key concepts necessary to understand its

principles and subsequent analyses. Section 3 presents our main results, comparing FIMAX to other distillation protocols regarding the share of distillable states across various families of low-fidelity bipartite states. We also report several interesting observations about FIMAX's distillation properties. Section 4 concludes our numerical analysis and the implications of our findings for future quantum technologies. In Sec. 5, we detail the technical aspects and methods concerning the analyzed error-affected states, the stabilizer theory underpinning FIMAX and the protocol itself. Additionally, we provide a simple example to enhance comprehension of these techniques.

2. Preliminaries

In this section, we introduce the basic notation for the application of the entanglement distillation protocol FIMAX¹³ and leave the details to Sec. 5. The goal of entanglement distillation protocols is to create maximally entangled states from potentially many copies of the generally error-affected mixed input state ρ_{in} that is shared between two parties named Alice and Bob. For this task, only local operations and classical communication (LOCC) are allowed to be used for Alice and Bob.

We choose the specific maximally entangled d^2 -dimensional bipartite state

$$|\Omega_{0,0}\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes^{\text{AB}} |i\rangle, \quad (1)$$

to be the target state. This state is one of the so-called Bell states, denoted by $|\Omega_{k,l}\rangle$ and defined in Sec. 5.1.2. Note that multiple copies of this target state with $d > 2$ can always be transformed to the often considered maximally entangled qubit state $|\Phi^+\rangle := \frac{1}{\sqrt{2}} \sum_{i=0}^1 |i\rangle \otimes^{\text{AB}} |i\rangle (\equiv |\Omega_{0,0}\rangle \text{ for } d = 2)$ via LOCC. Hence, any protocol that produces the states (1) for $d > 2$ can always be extended to distill $|\Phi^+\rangle$ in a lower-dimensional subspace.

We shortly describe the principles of the two-copy stabilizer-based recurrent entanglement distillation protocol, FIMAX. The protocol consists of the following main steps: First, Alice and Bob transform two copies of the input state via LOCC into Bell-diagonal form, in which states can be interpreted as mixtures of Bell states that are affected by so-called Weyl errors. Second, Alice and Bob perform local measurements using both copies of their states. Thereby, they choose a specific measurement basis, which is determined by the properties of the input state and related to a certain stabilizer, a subgroup of the operators representing the Weyl errors. Finally, they perform local operations based on the stabilizer and their measurement outcomes. If successful, the protocol results in an output state ρ_{out} , having a higher fidelity with the target state, i.e. $\mathcal{F}_{\text{out}} = \langle \Omega_{0,0} | \rho_{\text{out}} | \Omega_{0,0} \rangle \geq \langle \Omega_{0,0} | \rho_{\text{in}} | \Omega_{0,0} \rangle = \mathcal{F}_{\text{in}}$. FIMAX has been implemented as part of the open source Julia package “BellDiagonalQudits.jl”.¹⁹

We have summarized the details regarding the relevant theory and application of the protocol in Sec. 5. The Weyl–Heisenberg operators, their use as error operators

and for the construction of Bell states are introduced in Sec. 5.1 together with related state families used in this analysis. The stabilizer theory that FIMAX is built upon is revised in Sec. 5.2. Finally, Sec. 5.3 contains the precise FIMAX protocol and provides an example for its application. For more information, the reader is referred to Ref. 13.

3. Results

FIMAX was shown to perform entanglement distillation with high efficiency compared to other recurrence-based protocols and can distill certain strongly mixed states, for which other protocols fail.¹³ In this section, we analyze the performance of the described protocol with focus on the share of distillable states for more general classes of states. It is known that any d^2 -dimensional bipartite state ρ_{in} with fidelity $\mathcal{F}_{\text{in}} > \frac{1}{d}$ can be distilled by recurrent two-copy distillation, e.g. by the protocol given in Ref. 20. On the other hand, any entangled state with positive partial transposition (PPT) is bound entangled, i.e. cannot be distilled by any protocol. In this contribution, we thus restrict the analyses to states with negative partial transposition (NPT) and low fidelity. We apply FIMAX to bipartite systems of dimension $d = 2$ and $d = 3$, and compare the share of FIMAX-distillable states to the shares of other two-copy recurrence-based protocols.

For this purpose, we sample random states in several state families (cf. Sec. 5.1 for details). The first analyzed state family is the set of pure bipartite states, which we sample uniformly according to the Haar measure, serving a natural measure on the space of pure states.²¹ As a second family, we consider uniformly distributed mixtures of the standard Bell basis states $\{|\Omega_{k,l}\rangle\}$, constructed via the Weyl operators according to (6) in Sec. 5.1.2. These Bell-diagonal states (BDS) of the form (9) have a special entanglement structure, which is related to the properties of the Weyl operators. Finally, we analyze uniformly distributed mixtures of different Bell states, called “generalized Bell states”, which are defined in Sec. 5.1.3. Similar to the standard Bell basis, these generalized Bell states are also maximally entangled states that form a basis and can be constructed via the so-called “generalized Weyl operators” as in Eq. (12). However, due to additional phases, the strong algebraic relations between the Bell states are generally lost.⁵

To focus on low-fidelity states, we impose one of two restrictions on the set of sampled NPT states. In Sec. 3.1, we require the fidelity $\mathcal{F}_{\text{in}} = \langle \Omega_{0,0} | \rho_{\text{in}} | \Omega_{0,0} \rangle$ with the specific target state $|\Omega_{0,0}\rangle$ to be smaller than $\frac{1}{d}$. As a stricter condition for low-fidelity, in Sec. 3.2, we require $\langle \Omega_{k,l} | \rho_{\text{in}} | \Omega_{k,l} \rangle \leq \frac{1}{d}$ for all k, l . The latter means, that the fidelity of the input state with any of the maximally entangled Bell states (6) is less than the critical value for certain two-copy distillability.

For the families of pure states and standard BDS, we sample uniformly distributed states and filter this set to 10,000 NPT low-fidelity states. For the generalized Bell-diagonal states (gBDS), we construct 1000 generalized Bell-bases and for each sample 100 diagonal NPT states with low-fidelity. For these states,

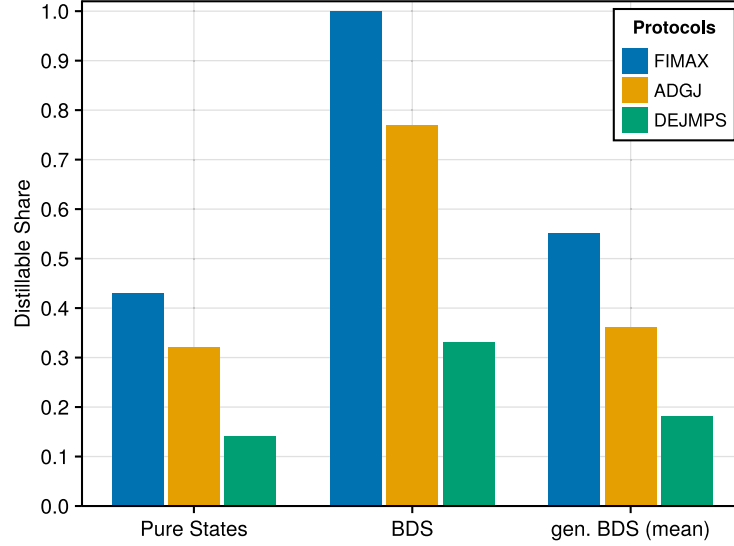


Fig. 1. Share of distillable states for $d = 2$ and $\mathcal{F}_{\text{in}} \leq \frac{1}{2}$.

the shares of distillable states are then evaluated. Here, we present maximum, minimum, mean and median share of the 1000 systems corresponding to the different generalized Bell bases. Note that due to the limited sample size of 100 states, the sample statistics have limited accuracy, especially for the max and min values.

The share of distillable states of this set is then determined for the following recurrent two-copy protocols: FIMAX,¹³ ADGJ,²² DEJMPS²³ ($d = 2$ only), P1–P2²⁴ and BBPSSW in its generalization to $d \geq 2$.^{20,25}

3.1. Distillability of low-fidelity states

First, consider NPT states for $d = 2$ that fulfill $\mathcal{F}_{\text{in}} \leq \frac{1}{2}$. The results are summarized in Fig. 1. For the family of pure bipartite states, FIMAX achieves the highest share of distillable states with 43% of analyzed low-fidelity states, followed by ADGJ with 32% and DEJMPS with 14%. P1–P2 and BBPSSW cannot distill any such states. In the set of standard Bell-diagonal states, FIMAX can distill all states, ADGJ 77%, DEJMPS 33%, P1–P2 0% and also BBPSSW 0%. Considering generalized Bell-diagonal states, the distillable shares are again nonzero only for FIMAX (max: 100%, min: 0%, mean: 55%, median: 55%), ADGJ (max: 85%, min: 0%, mean: 36%, median: 35%) and DEJMPS (max: 43%, min: 0%, mean: 18%, median: 20%).

The equivalent analysis for $d = 3$ is shown in Fig. 2. FIMAX can distill 54% of uniformly sampled pure states sufficing the restrictions of NPT and $\mathcal{F}_{\text{in}} \leq \frac{1}{3}$. ADGJ achieves a rate of 2%, while P1–P2 and BBPSSW cannot distill any of the sampled states. Considering the standard Bell-diagonal states, the distillable shares are as follows: FIMAX: 92%, ADGJ: 3%, P1–P2: 0%, BBPSSW: 0%. For generalized Bell-diagonal states with $d = 3$ and $\mathcal{F}_{\text{in}} \leq \frac{1}{3}$, the results for the maximal, minimal,

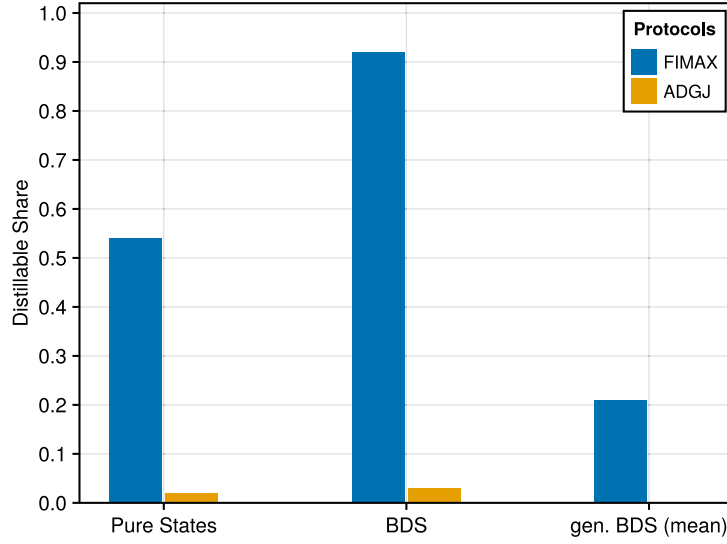


Fig. 2. Share of distillable states for $d = 3$ and $\mathcal{F}_{\text{in}} \leq \frac{1}{3}$.

mean and median share among created generalized Bell-diagonal systems are the following: FIMAX: (max: 54%, min: < 1%, mean: 21%, median: 21%), ADGJ: (max: 3%, min: 0%, mean: < 1%, median: < 1%). The other protocols (P1–P2, BBPSSW) cannot distill any states in all systems.

3.2. Distillability of strictly low-fidelity states

Imposing the stricter fidelity restriction $\langle \Omega_{k,l} | \rho_{\text{in}} | \Omega_{k,l} \rangle \leq \frac{1}{d}$ for all Bell states, none of the protocols can distill any state for $d = 2$. For FIMAX, BBPSSW and P12 this is expected, as any state whose fidelities with any Bell state is smaller than $\frac{1}{2}$ is PPT after being projected to Bell-diagonal form, and therefore undistillable.

For $d = 3$, the results are shown in Fig. 3. FIMAX is the only protocol to distill a significant share of low-fidelity pure states, with a share of 33%. ADGJ can distill < 1%, while for P1–P2 and BBPSSW not a single distillable state is observed. FIMAX can distill 83% of low-fidelity standard Bell-diagonal states satisfying the strict fidelity restriction. In contrast to the weaker fidelity restriction, none of the other protocols can distill a significant share of such states, as the share of ADGJ is < 1% and the shares of other protocols are equal to 0%. The same holds for generalized Bell-diagonal states, for which FIMAX distills a relevant number of states (max: 33%, min: 1%, mean: 11%, median: 9%).

3.3. Additional remarks

With the given sample sizes, the estimated deviation from the reported numbers for the distillable share is $\leq 1\%$ for equivalent analyses. In addition, we make some qualitative observations regarding the analysis above. First, all states that can be

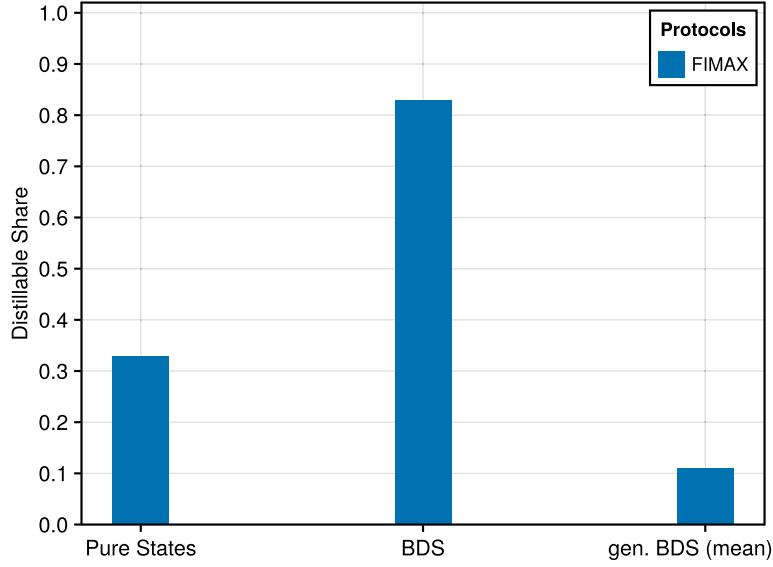


Fig. 3. Share of distillable states for $d = 3$ and $\langle \Omega_{k,l} | \rho_{\text{in}} | \Omega_{k,l} \rangle \leq \frac{1}{3}$ for all k, l .

distilled with a different protocol can also be distilled with FIMAX. Second, all NPT states that cannot be distilled with FIMAX, result in a PPT output state within the numerical precision after one iteration. Third, all analyzed states that can be distilled with FIMAX produce an output state that has fidelity $\geq \frac{1}{d}$ within the numerical precision after one iteration. Finally, all observed states that are not distillable with FIMAX converge to a Bell-diagonal state which is a fixpoint for recurrent application of the protocol.

4. Conclusion

In this work, we conducted a comprehensive comparison of the distillation effectivity of the recently introduced stabilizer-based two-copy bipartite entanglement distillation protocol FIMAX¹³ against other recurrent two-copy protocols, including ADGJ,²² DEJMPS,²³ P1–P2,²⁴ and the generalized BBPSSW protocol.²⁰ We evaluated their performance across three state families: pure states, standard Bell-diagonal states and generalized Bell-diagonal states.

Our analysis focused on low-fidelity states with negative partial transposition in bipartite systems of dimensions $d = 2$ and $d = 3$. We found that FIMAX outperforms all other protocols in terms of the share of distillable states. For states with initial fidelity less than $\frac{1}{d}$, only FIMAX, ADGJ and DEJMPS could distill entanglement. Notably, for $d = 3$, FIMAX's performance surpassed the only other protocol capable of distilling any states by more than an order of magnitude. The substantial share of random pure bipartite states with low fidelity distillable by FIMAX underscores its broad applicability, while its effectiveness with Bell-diagonal states highlights how it leverages the unique algebraic structure of these states. FIMAX's

advantage becomes even more pronounced under stricter fidelity constraints. When all elements of the Bell basis have low fidelity with the input state, it stands as the sole considered protocol capable of distilling entanglement, distilling a significant share of states in all analyzed state families.

The fact that every state distillable by other protocols was also distillable by FIMAX underlines its exceptional performance in handling low-fidelity states and emphasizes the protocol's robustness. Importantly, we observed that any state distillable by FIMAX achieves a fidelity greater than $\frac{1}{d}$ after the first iteration, enabling other protocols to further distill these states. Conversely, states undistillable by this protocol become positive under partial transposition (PPT) after the first application, hence becoming inherently undistillable. Leveraging these properties can lead to a resource-efficient application of the protocol, as the distinction of two-copy distillability can principally be made after just one iteration. Our observations provide valuable insights for future research into stabilizer-based approaches for entanglement distillation and the structure of entanglement.

In conclusion, combined with its confirmed high efficiency,¹³ our findings demonstrate that FIMAX is also very effective for two-copy distillation, particularly in high-noise regimes where other protocols fail. This positions FIMAX as a powerful tool in quantum information processing, potentially accelerating the development of reliable quantum communication in practically relevant, thus noisy, environments.

5. Methods

Finally, we provide details about the analyzed states and systems, the stabilizer-based methods and the application of the protocol, which we demonstrate by an explicit example.

5.1. Considered systems and states

Consider the d^2 -dimensional Hilbert space $\mathcal{H} \equiv \mathcal{H}_A \otimes^{AB} \mathcal{H}_B \cong \mathbb{C}^d \otimes \mathbb{C}^d$ of bipartite quantum states and the corresponding Hilbert space $\mathcal{H}^{\otimes 2} \equiv \bigotimes_{n=1}^2 \mathcal{H} \cong \mathcal{H}_A^{\otimes 2} \otimes^{AB} \mathcal{H}_B^{\otimes 2}$

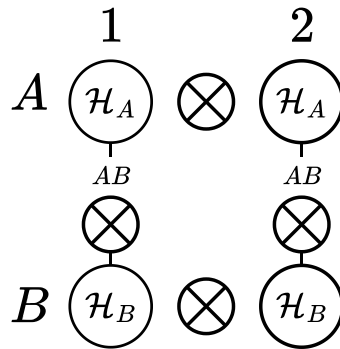


Fig. 4. Hilbert space $\mathcal{H}^{\otimes 2}$ of two-copies of a bipartite system $\mathcal{H} = \mathcal{H}_A \otimes^{AB} \mathcal{H}_B$.

of two-copy bipartite quantum states. Figure 4 visualizes the structure of $\mathcal{H}^{\otimes 2}$. In this contribution, we restrict d to be prime. $\mathcal{D}(\mathcal{H})$ denotes the set of density operators on $\mathcal{H}$. Further, $\omega \equiv \exp(\frac{2\pi i}{d})$ and $\mathcal{Z}_d \equiv \mathbb{Z}/d\mathbb{Z}$ denotes the quotient ring of integers with addition and multiplication modulo d .

5.1.1. Weyl errors

The *Weyl(-Heisenberg) operators* are a unitary generalization of the two-dimensional Pauli operators to higher dimensions d . They are defined as

$$W_{k,l} := \sum_j \omega^{jk} |j\rangle \langle j+l|, \quad j, k, l \in \mathcal{Z}_d, \quad (2)$$

and obey the so-called *Weyl relations*

$$\begin{aligned} W_{k_1, l_1} W_{k_2, l_2} &= \omega^{l_1 k_2} W_{k_1+k_2, l_1+l_2}, \\ W_{k,l}^\dagger &= \omega^{kl} W_{-k, -l} = W_{k,l}^{-1}. \end{aligned} \quad (3)$$

The Weyl operators form a so-called “nice error basis”,²⁶ representing both phase and shift errors, corresponding to the indices k, l , respectively. As both copies of the two-copy input state may be affected by errors, we define

$$e := \left[\begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \begin{pmatrix} k_2 \\ l_2 \end{pmatrix} \right] \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2, \quad (4)$$

$$W(e) := W_{k_1, l_1} \otimes W_{k_2, l_2},$$

as a *two-copy error basis*. The elements $e \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2$ are called *error elements*. Due to the Weyl relations (3) and ignoring the for this work irrelevant phases, these two-copy error operators form a group, which naturally extends to the error elements via elementwise addition modulo d by $W(e)W(f) \propto W(e+f)$. This group is denoted by

$$\mathcal{E}_2 := \{W(e) \mid e \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2\}. \quad (5)$$

5.1.2. Bell states

Besides their role as error operators, the Weyl operators can also be used to define a basis of the state space for pure bipartite qudits, whose elements are maximally entangled and called *Bell states*. They are defined as

$$|\Omega_{0,0}\rangle := \frac{1}{\sqrt{d}} \sum_i |i\rangle^{AB} \otimes |i\rangle, \quad i \in \mathcal{Z}_d, \quad (6)$$

$$|\Omega_{k,l}\rangle := (W_{k,l}^{AB} \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle, \quad k, l \in \mathcal{Z}_d.$$

The *two-copy Bell basis* is then defined by the two-copy states

$$\begin{aligned} |\Omega(e)\rangle &:= |\Omega_{k_1, l_1}\rangle \overset{AB}{\otimes} |\Omega_{k_2, l_2}\rangle \\ &\equiv (W(e) \overset{AB}{\otimes} \mathbb{1}_{d^2}) |\Omega_{0,0}\rangle^{\otimes 2}, \end{aligned} \quad (7)$$

where we use the notation introduced in Eq. (4). In this basis, any bipartite two-copy state $\rho^{\otimes 2}$ can be then written as

$$\rho^{\otimes 2} = \sum_{e,f} \rho(e, f) |\Omega(e)\rangle \langle \Omega(f)|, \quad (8)$$

where $e, f \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2$ and we use the notation $\rho(\cdot, \cdot)$ for the elements of the density matrix representing the state in the two-copy Bell basis. A state $\rho \in \mathcal{D}(\mathcal{H})$ is called Bell-diagonal if it is of the following form:

$$\rho = \sum_{k,l} p_{k,l} |\Omega_{k,l}\rangle \langle \Omega_{k,l}|, \quad (9)$$

where $p_{k,l} \geq 0$ and $\sum_{k,l} p_{k,l} = 1$. In this case, it can be interpreted as a mixture of Bell states (6), allowing to generate those states in experiment.²⁷ Due to its algebraic and geometric properties implied by the Weyl-relations and its special entanglement structure, this family of states is also known as “Magic Simplex”.^{28–30}

The two-copy state $\rho^{\otimes 2} \in \mathcal{D}(\mathcal{H}^{\otimes 2})$ of a Bell-diagonal state takes the following diagonal form:

$$\rho^{\otimes 2} = \sum_e \rho(e, e) |\Omega(e)\rangle \langle \Omega(e)|. \quad (10)$$

Any state $\rho \in \mathcal{D}(\mathcal{H})$ can be projected to Bell-diagonal form by application of the “Weyl twirl channel”.⁵

$$\rho \mapsto \frac{1}{d^2} \sum_{i,j} W_{i,j} \overset{AB}{\otimes} W_{-i,j} \rho (W_{i,j} \overset{AB}{\otimes} W_{-i,j})^\dagger, \quad (11)$$

with $i, j \in \mathcal{Z}_d$. Crucially for entanglement distillation, this channel can be realized by LOCC operations.

5.1.3. Generalized bell states

The family of generalized Bell basis states and corresponding “gBDS” is defined as follows (cf.⁵). Let $(\alpha_{s,t})_{s,t \in \mathbb{Z}_d}$ with $|\alpha_{s,t}| = 1 \ \forall s, t \in \mathbb{Z}_d$ be a matrix of complex phases. For each α , we can define an orthonormal basis of maximally entangled

states and corresponding diagonal states:

$$\begin{aligned}
 V_{k,l}^\alpha &:= \sum_j w^{jk} \alpha_{j+l,l} |j\rangle \langle j+l|, \\
 |\Phi_{k,l}^\alpha\rangle &:= (V_{k,l}^\alpha \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle, \\
 \rho &= \sum_{k,l} p_{k,l} |\Phi_{k,l}^\alpha\rangle \langle \Phi_{k,l}^\alpha|,
 \end{aligned} \tag{12}$$

where $j, k, l \in \mathcal{Z}_d$, $p_{k,l} \geq 0$ and $\sum_{k,l} p_{k,l} = 1$. The main difference to the standard Bell states is that the generalized Weyl operators $V_{k,l}^\alpha$ do not obey the Weyl relations (3), resulting in a loss of algebraic structure and a change in the entanglement structure of diagonal states.

5.2. Stabilizer theory

In the following, we shortly introduce the relevant stabilizer-related objects, required to apply the FIMAX distillation protocol. For detailed information about stabilizer theory and its application to entanglement distillation, see, e.g. Refs. 13, 15, 26, and 31.

Given the error group of Weyl error operators $\mathcal{E}_2$ as in Eq. (5), a *stabilizer* S is an abelian subgroup of $\mathcal{E}_2$. The authors of Ref. 13 showed that for prime dimensions, any stabilizer relevant for entanglement distillation is generated by one error operator $W(g)$ with a corresponding error element $g \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2$. Such stabilizers contain exactly d elements and are of the form $S = \{\mathbb{1}, W(g), W(2g), \dots, W((d-1)g)\}$.

A stabilizer S and the corresponding generating element g allow partitioning the set of two-copy error elements in two ways that are relevant for the FIMAX protocol. First, it decomposes into *cosets* of the stabilizer S , which are defined as the sets

$$C(e) := \{e, e+g, e+2g, \dots, e+(d-1)g\}, \tag{13}$$

where $e \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2$. Second, a stabilizer implies a decomposition of error elements by their *symplectic product* with the generator $g = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$. For an error element $e = \begin{bmatrix} k_1 \\ l_1 \end{bmatrix}, \begin{bmatrix} k_2 \\ l_2 \end{bmatrix}$, it is defined as $\langle g, e \rangle := \sum_{n=1}^2 b_n k_n - a_n l_n$. We have a decomposition of error elements into d subsets of the following form:

$$\varepsilon(s) := \{e \in \mathcal{Z}_d^2 \times \mathcal{Z}_d^2 \mid \langle g, e \rangle = s\}, \quad s \in \mathcal{Z}_d. \tag{14}$$

Note that both decompositions depend on the generator g and that if $e \in \varepsilon(s)$, then all error elements of the error coset $C(e)$ are in $\varepsilon(s)$, as well.

The eigenspaces of the generator $W(g)$ decompose the Hilbert space on which the error operators act in so-called *codespaces* $\mathcal{Q}$. Using the two-copy Bell basis to represent the input states (7), it is clear that $W(e)$ acts on the two parts of the A -side of the bipartite system. Denoting the eigenspaces of $W(g)$ with eigenvalue corresponding to x by $\mathcal{Q}(x)$, we have $\mathcal{H}_A^{\otimes 2} \cong \bigoplus_x \mathcal{Q}(x)$. For nontrivial generators,

the codespaces are of dimension d and can be spanned by an orthonormal basis of eigenvectors of $W(g)$, the so-called *codewords* $\{|u_{x,k}\rangle\}_{x,k \in \mathbb{Z}_d}$. A unitary map $U : |x\rangle \otimes |k\rangle \mapsto |u_{x,k}\rangle \in \mathcal{Q}(x) \subset \mathcal{H}_A^{\otimes 2}$ is called *encoding* for S . The authors of Ref. 13 proposed a *canonic encoding* U_c , based on the product eigenvectors of the given generator $W(g) = W_{a_1,b_1} \otimes W_{a_2,b_2}$. U_c has special properties that are leveraged in FIMAX. A *stabilizer measurement* performed on $\mathcal{H}_A^{\otimes 2}$ corresponds to a projection on a certain codespace $\mathcal{Q}(x)$ and can be expressed as the projection-valued measure $\{\mathcal{P}(x)\}_{x \in \mathbb{Z}_d}$, $\mathcal{P}(x) = \sum_k |u_{x,k}\rangle\langle u_{x,k}|$. Given ρ , the unnormalized post-measurement state for measurement outcome x is $\mathcal{P}(x)\rho\mathcal{P}(x) \in \mathcal{Q}(x)$.

Note that the objects defined above relate to the A side of the bipartition, as the elements of S are operators acting on $\mathcal{H}_A^{\otimes 2}$. To use stabilizer codes for distillation, we also need the corresponding objects for the other side B . Denote by S^* the stabilizer with complex conjugated elements of S , generated by $W(g)^*$ and implying a decomposition $\mathcal{H}_B^{\otimes 2} \cong \bigoplus_y \mathcal{Q}^*(y)$. Then, $U^* : |y\rangle \otimes |l\rangle \mapsto |u_{y,l}^*\rangle \in \mathcal{Q}^*(y)$ defines an encoding for S^* and $\mathcal{P}^*(y) = \sum_l |u_{y,l}^*\rangle\langle u_{y,l}^*|$ defines the projections for the stabilizer measurements corresponding to S^* .

5.3. FIMAX

Having introduced the relevant objects in the previous section, we now specify the two-copy recurrence protocol FIMAX and provide an example for its application.

5.3.1. Protocol

Let $\rho_{\text{in}} \in \mathcal{D}(\mathcal{H})$ be a bipartite state in prime dimension d and $\rho_{\text{in}}^{\otimes 2} \in \mathcal{D}(\mathcal{H}^{\otimes 2})$ be the two-copy state as in Eq. (8). Define the probability distribution $\mathbb{p}(e) := \rho_{\text{in}}^{\otimes 2}(e, e)$ and denote corresponding probabilities by $\mathbb{P}$.

- (1) If ρ_{in} is not of Bell-diagonal form (9), then we apply the Weyl twirl channel (11) to both copies.
- (2) For each stabilizer S with generating element $g \in \mathbb{Z}_d^2 \times \mathbb{Z}_d^2$:
 - (i) Partition the error elements e according to their symplectic product (14) and calculate $\mathbb{P}(\varepsilon(s))$.
 - (ii) For each s with $\mathbb{P}(\varepsilon(s)) > 0$ and for each error coset $C = C(e)$, $e \in \varepsilon(s)$ (13), determine $(C_{\text{max}}, s_{\text{max}}) := \arg \max \frac{\mathbb{P}(C)}{\mathbb{P}(\varepsilon(s))}$.
- (3) Alice chooses the stabilizer S_{max} , maximizing $\frac{\mathbb{P}(C_{\text{max}})}{\mathbb{P}(\varepsilon(s_{\text{max}}))}$ among stabilizers and communicates her choice classically to Bob.
- (4) Alice and Bob perform stabilizer measurements for S_{max} and S_{max}^* with measurement outcomes a, b , respectively.
- (5) Bob sends b to Alice. Alice declares failure of the protocol if $s_{\text{max}} \neq a - b$.
- (6) Alice and Bob apply the inverse of the canonic encoding U_c and U_c^* for S_{max} and S_{max}^* to their qudits, respectively.
- (7) Alice and Bob discard their first qudit containing the measurement outcome.

- (8) Alice applies $W_{k_{\max}, l_{\max}}^\dagger$ to the remaining qudit, where $(k_{\max}, l_{\max})$ is determined from $(C_{\max}, s_{\max})$ as described in Ref. 13.

Note that FIMAX maps Bell-diagonal states to Bell-diagonal states, so the first step of the protocol is only required once for recurrent application of the protocol.

5.3.2. Example

As an explicit example to better understand the relevant stabilizer-related objects leveraged by FIMAX, consider the case $d = 3$ and the Bell-diagonal input state, taking the form in the basis (6), ordered first by shift then by phase index as $|\Omega_{0,0}\rangle, |\Omega_{1,0}\rangle, \dots, |\Omega_{1,2}\rangle, |\Omega_{2,2}\rangle$:

$$\rho_{\text{in}} = \begin{pmatrix} 0.06 & & & & & & \\ & 0.06 & & & & & \\ & & 0.06 & & & & \\ & & & 0.06 & & & \\ & & & & 0.06 & & \\ & & & & & \mathbf{0.56} & \\ & & & & & & 0.06 \\ & & & & & & & 0.06 \\ & & & & & & & & 0.06 \end{pmatrix}.$$

This state has small fidelity $\mathcal{F} = 0.06$, but a large overlap of 0.56 with the Bell state $|\Omega_{2,1}\rangle$. Comparing the coset probabilities for all stabilizers according to step 2 of FIMAX, the stabilizer with the generating element $g = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathcal{Z}_3^2 \times \mathcal{Z}_3^2$ implies the best partition of error elements. More precisely, g corresponds to the operator $W(g) = W_{1,0} \otimes W_{1,0}$ which generates the stabilizer

$$\begin{aligned} S_{\max} &= \{W(0), W(g), W(2g)\} \\ &= \{\mathbb{1}_3 \otimes \mathbb{1}_3, W_{1,0} \otimes W_{1,0}, W_{2,0} \otimes W_{2,0}\}. \end{aligned} \quad (15)$$

The coset

$$C_{\max} = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\},$$

is a subset of $\varepsilon(1)$, i.e. all error elements having symplectic product $\langle g, e \rangle = 1 = s_{\max}$. Based on the probabilities $\mathbb{P}$ induced by $\rho_{\text{in}}^{\otimes 2}$, $\mathbb{P}(\varepsilon(1)) = 0.5$ and $\mathbb{P}(C) = 0.315$, which correspond to the largest value of $\frac{\mathbb{P}(C)}{\mathbb{P}(\varepsilon(s))}$ for all stabilizers S , cosets C and values of s .

The eigenvalues of the generator $W(g)$ are ω^x , $x \in \mathbb{Z}_3$, each three-fold degenerate. It can be checked that the unitary

$$U_c := |x\rangle \otimes |k\rangle \mapsto |u_{x,k}\rangle := |k\rangle \otimes |x - k\rangle, \quad (16)$$

maps any computational basis state $|x\rangle \otimes |k\rangle$ to a codeword $|u_{x,k}\rangle$ of the codespace $\mathcal{Q}(x)$ corresponding to the eigenvalue ω^x and is therefore an encoding for S . In fact, this encoding is the canonic encoding U_c for S . Note that in this example, we have $S = S^*$ and therefore $U_c = U_c^*$.

Performing the stabilizer measurements having outcomes a, b with $a - b = s_{\max} = 1$, according to steps 4 and 5, followed by the application of $U_c^\dagger \otimes U_c^\dagger$ in step 6 yields the following state of the remaining qudits:

$$\rho'_{\text{out}} = \begin{pmatrix} 0.02 & & & & & & & \\ & 0.02 & & & & & & \\ & & 0.02 & & & & & \\ & & & \mathbf{0.63} & & & & \\ & & & & \mathbf{0.13} & & & \\ & & & & & \mathbf{0.13} & & \\ & & & & & & 0.02 & \\ & & & & & & & 0.02 \\ & & & & & & & & 0.02 \end{pmatrix}.$$

Applying the correction operation $W_{0,2}^{AB} \otimes \mathbb{1}_3$ results in the final output state

$$\rho_{\text{out}} = \begin{pmatrix} \mathbf{0.63} & & & & & & & \\ & \mathbf{0.13} & & & & & & \\ & & \mathbf{0.13} & & & & & \\ & & & 0.02 & & & & \\ & & & & 0.02 & & & \\ & & & & & 0.02 & & \\ & & & & & & 0.02 & \\ & & & & & & & 0.02 \end{pmatrix}.$$

The state ρ_{out} has a higher fidelity and maximal overlap with any Bell state ($\mathcal{F} = 0.63$) compared to the input state ρ_{in} .

Acknowledgments

C.P. and B.C.H. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) Project P36102-N (Grant DOI:

10.55776/P36102). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this paper.

ORCID

Christopher Popp  <https://orcid.org/0000-0002-2352-3731>

Tobias C. Sutter  <https://orcid.org/0009-0008-3179-7979>

Beatrix C. Hiesmayr  <https://orcid.org/0000-0001-9062-6039>

References

1. S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi and P. Wallden, *Adv. Opt. Photonics* **12** (2020) 1012.
2. Y. Wang, Z. Hu, B. C. Sanders and S. Kais, *Front. Phys.* **8** (2020) 589504.
3. E. Chitambar and G. Gour, *Rev. Mod. Phys.* **91** (2019) 025001.
4. S. Bäuml, S. Das, X. Wang and M. M. Wilde, arXiv:1907.04181.
5. C. Popp and B. C. Hiesmayr, *New J. Phys.* **26** (2024) 013039.
6. J. Preskill, *Quantum* **2** (2018) 79.
7. R. Acharya, D. A. Abanin, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, N. Astrakhantsev, J. Atalaya, R. Babbush, D. Bacon, B. Ballard, J. C. Bardin, J. Bausch, A. Bengtsson, A. Bilmes, S. Blackwell, S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, D. A. Browne, B. Buchea, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, A. Cabrera, J. Campero, H.-S. Chang, Y. Chen, Z. Chen, B. Chiaro, D. Chik, C. Chou, J. Claes, A. Y. Cleland, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, S. Das, A. Davies, L. De Lorenzo, D. M. Debroy, S. Demura, M. Devoret, A. Di Paolo, P. Donohoe, I. Drozdov, A. Dunsworth, C. Earle, T. Edlich, A. Eickbusch, A. M. Elbag, M. Elzouka, C. Erickson, L. Faoro, E. Farhi, V. S. Ferreira, L. F. Burgos, E. Forati, A. G. Fowler, B. Foxen, S. Ganjam, G. Garcia, R. Gasca, É. Genois, W. Jiang, C. Gidney, D. Gilboa, R. Gosula, A. G. Dau, D. Graumann, A. Greene, J. A. Gross, S. Habegger, J. Hall, M. C. Hamilton, M. Hansen, M. P. Harrigan, S. D. Harrington, F. J. H. Heras, S. Heslin, P. Heu, O. Higgott, G. Hill, J. Hilton, G. Holland, S. Hong, H.-Y. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, S. Jordan, C. Joshi, P. Juhas, D. Kafri, H. Kang, A. H. Karamlou, K. Kechedzhi, J. Kelly, T. Khairé, T. Khattar, M. Khezri, S. Kim, P. V. Klimov, A. R. Klotz, B. Kobrin, P. Kohli, A. N. Korotkov, F. Kostritsa, R. Kothari, B. Kozlovskii, J. M. Kreikebaum, V. D. Kurilovich, N. Lacroix, D. Landhuis, T. Lange-Dei, B. W. Langley, P. Laptev, K.-M. Lau, L. Le Guevel, J. Ledford, J. Lee, K. Lee, Y. D. Lensky, S. Leon, B. J. Lester, W. Y. Li, Y. Li, A. T. Lill, W. Liu, W. P. Livingston, A. Locharla, E. Lucero, D. Lundahl, A. Lunt, S. Madhuk, F. D. Malone, A. Maloney, S. Mandrà, J. Manyika, L. S. Martin, O. Martin, S. Martin, C. Maxfield, J. R. McClean, M. McEwen, S. Meeks, A. Megrant, X. Mi, K. C. Miao, A. Mieszala, R. Molavi, S. Molina, S. Montazeri, A. Morvan, R. Movasagh, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, H. Neven, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, C.-H. Ni, M. Y. Niu, T. E. O'Brien,

- W. D. Oliver, A. Opremcak, K. Ottosson, A. Petukhov, A. Pizzuto, J. Platt, R. Potter, O. Pritchard, L. P. Pryadko, C. Quintana, G. Ramachandran, M. J. Reagor, J. Redding, D. M. Rhodes, G. Roberts, E. Rosenberg, E. Rosenfeld, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, A. W. Senior, M. J. Shearn, A. Shorter, N. Shutty, V. Shvarts, S. Singh, V. Sivak, J. Skrzynny, S. Small, V. Smelyanskiy, W. C. Smith, R. D. Somma, S. Springer, G. Sterling, D. Strain, J. Suchard, A. Szasz, A. Sztein, D. Thor, A. Torres, M. M. Torunbalci, A. Vaishnav, J. Vargas, S. Vdovichev, G. Vidal, B. Villalonga, C. V. Heidweiller, S. Waltman, S. X. Wang, B. Ware, K. Weber, T. Weidel, T. White, K. Wong, B. W. K. Woo, C. Xing, Z. J. Yao, P. Yeh, B. Ying, J. Yoo, N. Yosri, G. Young, A. Zalcman, Y. Zhang, N. Zhu, N. Zobrist and Google Quantum AI and Collaborators, *Nature* **638** (2024) 1.
8. S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall and T. J. Yoder, *Nature* **627** (2024) 778.
9. A. M. Steane, *Phys. Rev. Lett.* **77** (1996) 793.
10. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, *Phys. Rev. A* **54** (1996) 3824.
11. W. Dür and H.-J. Briegel, *Phys. Rev. Lett.* **90** (2003) 067901.
12. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, *Phys. Rev. Lett.* **70** (1993) 1895.
13. C. Popp, T. C. Sutter and B. C. Hiesmayr, arXiv:2408.02383.
14. D. Gottesman, arXiv:9705052.
15. R. Matsumoto, *J. Phys. A: Math. Gen.* **36** (2003) 8113.
16. K. G. H. Vollbrecht and M. M. Wolf, *Phys. Rev. A* **67** (2003) 012303.
17. M. Horodecki, P. Horodecki and R. Horodecki, *Phys. Rev. Lett.* **80** (1998) 5239.
18. B. C. Hiesmayr, C. Popp and T. C. Sutter, arXiv:2406.13491.
19. C. Popp, *J. Open Source Softw.* **8** (2023) 4924.
20. M. Horodecki and P. Horodecki, *Phys. Rev. A* **59** (1999) 4206.
21. K. Życzkowski, P. Horodecki, A. Sanpera and M. Lewenstein, *Phys. Rev. A* **58** (1998) 883.
22. G. Alber, A. Delgado, N. Gisin and I. Jex, *J. Phys. A: Math. Gen.* **34** (2001) 8821.
23. D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu and A. Sanpera, *Phys. Rev. Lett.* **77** (1996) 2818.
24. J. Miguel-Ramiro and W. Dür, *Phys. Rev. A* **98** (2018) 042309.
25. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin and W. K. Wootters, *Phys. Rev. Lett.* **76** (1996) 722.
26. E. Knill, arXiv:9608048.
27. B. C. Hiesmayr and W. Löffler, *New J. Phys.* **15** (2013) 083036.
28. B. Baumgartner, B. Hiesmayr and H. Narnhofer, *J. Phys. A: Math. Theor.* **40** (2007) 7919.
29. C. Popp and B. C. Hiesmayr, *Sci. Rep.* **12** (2022) 12472.
30. C. Popp and B. C. Hiesmayr, *Sci. Rep.* **13** (2023) 2037.
31. E. Rains, *IEEE Trans. Inf. Theory* **45** (1999) 1827.

B.8. Local Invariance of Divergence-based Quantum Information Measures

Article

Local Invariance of Divergence-Based Quantum Information Measures

Christopher Popp , Tobias C. Sutter  and Beatrix C. Hiesmayr 

Faculty of Physics, University of Vienna, Währingerstraße 17, 1090 Vienna, Austria; tobias.christoph.sutter@univie.ac.at (T.C.S.); beatrix.hiesmayr@univie.ac.at (B.C.H.)

* Correspondence: christopher.popp@univie.ac.at

Abstract

Quantum information quantities, such as mutual information and entropies, are essential for characterizing quantum systems and protocols in quantum information science. In this contribution, we identify types of information measures based on generalized divergences and prove their invariance under local isometric or unitary transformations. Leveraging the reversal channel for local isometries together with the data-processing inequality, we establish invariance for information quantities used in both asymptotic and one-shot regimes without relying on the specific functional form of the underlying divergence. These invariances can be applied to improve the computation of such information quantities or optimize protocols and their output states, whose performance is determined by some invariant measure. Our results improve the capability to characterize and compute many operationally relevant information measures with application across the field of quantum information processing.

Keywords: quantum information; entropy; mutual information; divergence; local invariance; reversal channel



Academic Editor: Rosario Lo Franco

Received: 10 September 2025

Revised: 6 October 2025

Accepted: 7 October 2025

Published: 10 October 2025

Citation: Popp, C.; Sutter, T.C.; Hiesmayr, B.C. Local Invariance of Divergence-Based Quantum Information Measures. *Entropy* **2025**, *27*, 1051. <https://doi.org/10.3390/e27101051>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A central goal of quantum information theory is the precise quantification of correlations, uncertainties, and distinguishability within quantum systems. Various information-theoretic quantities have been defined to capture the fundamental limits of quantum information processing tasks like communication, computation, or entanglement manipulation.

Traditionally, the fundamental quantity is the von Neumann entropy [1], from which many quantities relevant in classical information science, like the mutual information or the conditional entropy, can be generalized to the quantum regime (cf. [2]). These measures are relevant due to their operational meaning in quantifying optimal rates of achieving various information processing tasks in the asymptotic regime of infinitely many independent uses of a resource (e.g., quantum state). A notable example is the quantum relative entropy [3], which can be related to quantum hypothesis testing [4]. Several other information quantities can be expressed in terms of the relative entropy, such as the private information for the rate of secret-key distillation and the coherent information for the rate of entanglement distillation [5]. In the so-called one-shot setting (cf. Ref. [6]), only a finite amount of resources are considered. In this approach, the goal is to find optimal rates to transform the limited resources, such as quantum states, into the target states while allowing fixed error bounds. Moreover, the one-shot setting is crucial for security analysis in the context of quantum cryptography without any assumptions on the actions of a malicious third

party [7]. As in the asymptotic setting, several information quantities with operational meaning have been identified such as the ε -hypothesis testing mutual information [8] and the smooth max-mutual information [9] that can be related to secret-key distillation in the one-shot setting.

As quantum technologies advance, the landscape of information quantities and related processing tasks becomes increasingly diverse (see Ref. [10] for a comprehensive overview of information measures and related processing tasks) and a unified framework capable of accommodating both asymptotic and one-shot scenarios is essential. Generalized divergences offer such a unifying language. By definition, each divergence satisfies monotonicity under completely positive trace-preserving (CPTP) maps, i.e., a data-processing inequality under the action of quantum channels, that guarantee operational meaning and allow us to derive an entire spectrum of information measures. Examples include the Petz–Rényi relative entropy [11–13], the sandwiched Rényi relative entropy [14,15], the geometric Rényi relative entropy [16,17] and (smooth) max/min-relative entropies [7]. Note that the relative entropy also satisfies a data-processing inequality and thus many information quantities in the asymptotic setting can be expressed by this specific divergence. Replacing the relative entropy by other divergences in similar expressions allows one to define generalized quantities that are applicable in other regimes, e.g., the one-shot setting.

Despite their broad applicability, evaluating these divergence-based measures poses significant computational problems in the case of high-dimensional quantum states or if their definition involves challenging or infeasible optimizations. In this work, we define several types of divergence-based information measures and prove their invariance under local isometric or unitary transformations.

In Section 2, we introduce the setting and notation and define the types of generalized divergence-based information quantities. We continue to present and prove the main result of this contribution, namely the invariance of information quantities of the defined types under local isometric or unitary transformations in Section 3. Finally, we conclude our results in Section 4 and provide an outlook to potential applications and future research directions.

2. Methods: Generalized Divergences and Information Quantities

In this section, we briefly introduce the notation and necessary objects to define several types of generalized information quantities and prove their invariance under local isometric or unitary transformations.

2.1. Notation and Setting

We consider two parties A and B with corresponding Hilbert spaces $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. $L_+(\mathcal{H})$ and $\mathcal{D}(\mathcal{H})$ denote the spaces of positive-semidefinite operators and density operators (i.e., quantum states) acting on $\mathcal{H}$, respectively. Isometries are denoted by a capital V . Corresponding channels, i.e., completely positive and trace-preserving maps, are written as $\mathcal{V}(\cdot) \equiv V(\cdot)V^\dagger$. Unitary operators are referred to with a capital U . The identity map is written as id and $\circ$ denotes the composition of maps. States or operators are labeled with the systems they act on. For any multipartite state or operator, e.g., the state ρ_{AB} , the marginal state or operator is given by $\rho_A := \text{Tr}_B[\rho_{AB}]$.

2.2. Generalized Divergence-Based Types of Information Quantities

The information quantities we analyze in this contribution are defined via so-called generalized divergences $\mathbf{D}$ [18,19].

Definition 1 (Generalized divergence, data-processing inequality). For quantum states $\rho \in \mathcal{D}(\mathcal{H})$ and positive-semidefinite operators $\zeta \in L_+(\mathcal{H})$, a function $\mathbf{D} : \mathcal{D}(\mathcal{H}) \times L_+(\mathcal{H}) \rightarrow \mathbb{R} \cup \infty$ that satisfies the data-processing inequality under any channel $\mathcal{N}$, $\mathbf{D}(\rho||\zeta) \geq \mathbf{D}(\mathcal{N}(\rho)||\mathcal{N}(\zeta))$, is called (generalized) divergence.

For quantum states, the data-processing inequality can be interpreted as the property that no physical transformation can make two states more distinguishable. It is known that any divergence is invariant under the application of isometric transformations $\mathcal{V}$ [10]:

$$\mathbf{D}(\mathcal{V}(\rho)||\mathcal{V}(\zeta)) = \mathbf{D}(\rho||\zeta). \quad (1)$$

Generalized divergences are used to define generalized information quantities for bipartite quantum states ρ_{AB} . There exist several ways to define information quantities based on divergences with operational meaning. In this work, we consider quantities that relate a bipartite state and its local marginals. Using specific instances for the divergences, these quantities can be related to various information-processing tasks (cf. [10] for a comprehensive overview). The method to prove local invariance of the quantities we present in this work, however, does not depend on the specific form of the divergence, but only on how it is applied to the state. We therefore define several types of information quantities independent of the specific form of the divergence. Some types involve a so-called smoothing, i.e., an optimization over an environment of the quantum state. For this, the sine distance [20] is often used as a distance measure that is closely related to the fidelity [21] as defined below.

In the following, let $\varepsilon \in [0, 1]$ and let ρ and σ be states.

Definition 2 (Fidelity, sine distance). The fidelity of two quantum states ρ and σ is defined as follows:

$$\mathcal{F}(\rho, \sigma) := \left(\text{Tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}] \right)^2 \quad (2)$$

The sine distance is defined as:

$$P(\rho, \sigma) := \sqrt{1 - \mathcal{F}(\rho, \sigma)} \quad (3)$$

However, other distinguishability functions potentially with different properties can be used as well. In this work, we only require that the function obeys the data-processing inequality and invariance under isometric transformations. As both properties are satisfied by any divergence (cf. Definition 1 and (1)), including the sine distance [10], the smoothing environment $B^\varepsilon(\rho)$ around a state ρ is defined as follows:

Definition 3 (Smoothing environment). Let $\mathbf{D}$ be any divergence.

$$B^\varepsilon(\rho) := \{\hat{\rho} : \mathbf{D}(\rho||\hat{\rho}) \leq \varepsilon\} \quad (4)$$

For the sine distance, it specifically reads:

$$B_P^\varepsilon(\rho) := \{\hat{\rho} : P(\rho, \hat{\rho}) \leq \varepsilon\} = \left\{ \hat{\rho} : \mathcal{F}(\rho, \hat{\rho}) \geq 1 - \varepsilon^2 \right\}. \quad (5)$$

In this work, we define and analyze the following types of information quantities:

Definition 4 (Types of generalized mutual information).

$$\mathbf{I}_1(\rho_{AB}) := \mathbf{D}(\rho_{AB} || \rho_A \otimes \rho_B) \quad (6)$$

$$\mathbf{I}_2(\rho_{AB}) := \inf_{\sigma_B} \mathbf{D}(\rho_{AB} || \rho_A \otimes \sigma_B) \quad (7)$$

$$\mathbf{I}_3^\varepsilon(\rho_{AB}) := \inf_{\substack{\hat{\rho}_{AB} \in B^\varepsilon(\rho_{AB}) \\ \sigma_B}} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \sigma_B) \quad (8)$$

$$\mathbf{I}_4^\varepsilon(\rho_{AB}) := \inf_{\hat{\rho}_{AB} \in B^\varepsilon(\rho_{AB})} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \hat{\rho}_B) \quad (9)$$

Definition 5 (Types of generalized conditional entropies).

$$\mathbf{H}_1(\rho_{AB}) := -\mathbf{D}(\rho_{AB} || \mathbb{1}_A \otimes \rho_B) \quad (10)$$

$$\mathbf{H}_2(\rho_{AB}) := -\inf_{\sigma_B} \mathbf{D}(\rho_{AB} || \mathbb{1}_A \otimes \sigma_B) \quad (11)$$

$$\mathbf{H}_3^\varepsilon(\rho_{AB}) := -\inf_{\substack{\hat{\rho}_{AB} \in B^\varepsilon(\rho_{AB}) \\ \sigma_B}} \mathbf{D}(\hat{\rho}_{AB} || \mathbb{1}_A \otimes \sigma_B) \quad (12)$$

The defined types include information quantities in both the asymptotic and the one-shot setting. Examples of this are the (generalized) quantum mutual information ($\mathbf{I}_1, \mathbf{I}_2$), the coherent information ($\mathbf{H}_2$), the Petz-Rényi mutual information, the sandwiched Rényi mutual information, the geometric Rényi mutual information, the smooth min-mutual information (see Ref. [22] for these quantities), the ε -hypothesis testing mutual information ($\mathbf{I}_3^\varepsilon$), the smooth max-conditional and min-conditional entropy ($\mathbf{H}_3^\varepsilon$), and the smooth max-mutual information ($\mathbf{I}_4^\varepsilon$).

3. Results: Local Invariance of Information Quantities

We show that the types $\mathbf{I}_1, \mathbf{I}_2$ and $\mathbf{I}_3^\varepsilon$ are invariant under any local isometric transformation, and that the types $\mathbf{I}_4^\varepsilon, \mathbf{H}_1, \mathbf{H}_2$ and $\mathbf{H}_3^\varepsilon$ are invariant if the first subsystem is transformed unitarily, while the second subsystem can be transformed by any isometry. More precisely, we prove below:

Proposition 1. Any information quantity of type $\mathbf{I}_1, \mathbf{I}_2$ and $\mathbf{I}_3^\varepsilon$ as in Definition 4 is invariant under local isometric transformations of the form $\mathcal{V}_{AB}(\cdot) = V_A \otimes V_B (\cdot) V_A^\dagger \otimes V_B^\dagger$.

Proposition 2. Any information quantity of type $\mathbf{I}_4^\varepsilon, \mathbf{H}_1, \mathbf{H}_2$ and $\mathbf{H}_3^\varepsilon$ as in Definitions 4 and 5 is invariant under local unitary transformations in the first system and isometric transformations in the second system of the form $\mathcal{V}_{AB}(\cdot) = U_A \otimes V_B (\cdot) U_A^\dagger \otimes V_B^\dagger$.

The proofs for Propositions 1 and 2 use the technical Lemmas 1, 2 and 4 related to the isometric transformations and the corresponding so-called reversal channel [2], defined in the following. Due to $VV^\dagger \neq \mathbb{1}$ in general, the map $\mathcal{V}^\dagger$ is not necessarily trace-preserving, although it is completely positive. The reversal channel $\mathcal{R}_\mathcal{V}$ corresponding to an isometric channel $\mathcal{V}$ is a completely positive and trace-preserving map satisfying $\mathcal{R}_\mathcal{V} \circ \mathcal{V} = \text{id}$, hence reversing the action of an isometric channel.

Definition 6 (Reversal channel). Let $\mathcal{V} : L_+(\mathcal{H}) \rightarrow L_+(\tilde{\mathcal{H}})$ be an isometric channel and $\omega \in \mathcal{D}(\mathcal{H})$. The reversal channel corresponding to $\mathcal{V}$ and ω is defined as

$$\mathcal{R}_\mathcal{V}(\sigma) := \mathcal{V}^\dagger(\sigma) + \text{Tr}[(\mathbb{1} - VV^\dagger)\sigma]\omega, \quad (13)$$

for any $\sigma \in L_+(\tilde{\mathcal{H}})$.

Note, that this channel is not unique and $\mathcal{V}(\mathcal{R}_\mathcal{V}(\rho)) \neq \rho$, in general. Also note that while $\sigma \in L_+(\tilde{\mathcal{H}})$, $\omega \in \mathcal{D}(\mathcal{H})$ needs to be a quantum state such that the map is trace-preserving. Given an isometric channel $\mathcal{V}$, we define the following sets that relate a smoothing environment to its image and pre-image under the isometric channel and its reversal channel:

Definition 7. For a state ρ with ε -environment $B^\varepsilon(\rho)$ as in Definition 3, an isometric channel $\mathcal{V}$, and the reversal channel $\mathcal{R}_\mathcal{V}$, we define:

$$B_\mathcal{V}^\varepsilon(\rho) := \{\mathcal{V}(\hat{\rho}) \mid \hat{\rho} \in B^\varepsilon(\rho)\}, \quad (14)$$

$$B_{\mathcal{R}_\mathcal{V}}^\varepsilon(\rho) := \{\tilde{\rho} \mid \mathcal{R}_\mathcal{V}(\tilde{\rho}) \in B^\varepsilon(\rho)\}. \quad (15)$$

We now show that these sets can be ordered (see Figure 1), implying inequalities for smoothed information quantities as used in the proofs of Propositions 1 and 2.

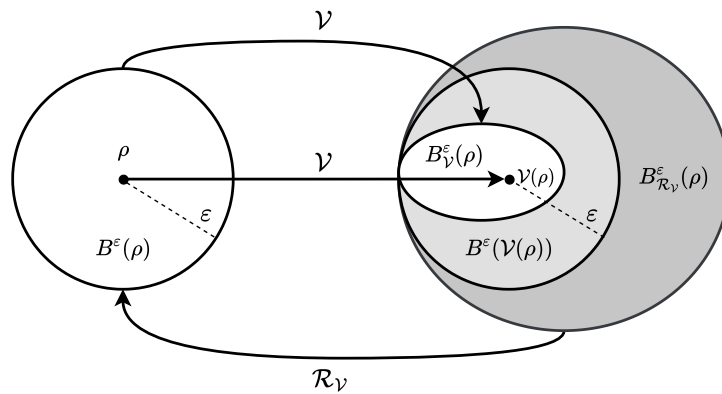


Figure 1. Schematic visualization of the smoothing environment B^ε and its image $B_\mathcal{V}^\varepsilon$ and pre-image $B_{\mathcal{R}_\mathcal{V}}^\varepsilon$ under the isometric channel $\mathcal{V}$ and its reversal channel $\mathcal{R}_\mathcal{V}$.

Lemma 1. Regarding the sets from Definitions 3 and 7, the following set relations hold:

$$B_\mathcal{V}^\varepsilon(\rho) \subset B^\varepsilon(\mathcal{V}(\rho)) \subset B_{\mathcal{R}_\mathcal{V}}^\varepsilon(\rho). \quad (16)$$

Proof. For the first relation let $\tilde{\rho} \in B_\mathcal{V}^\varepsilon(\rho)$, $\tilde{\rho} = \mathcal{V}(\hat{\rho})$, $\hat{\rho} \in B^\varepsilon(\rho)$. Using the invariance under isometric transformations (1), we have $\mathbf{D}(\mathcal{V}(\rho) \parallel \tilde{\rho}) = \mathbf{D}(\mathcal{V}(\rho) \parallel \mathcal{V}(\hat{\rho})) = \mathbf{D}(\rho \parallel \hat{\rho}) \leq \varepsilon \implies \tilde{\rho} \in B^\varepsilon(\mathcal{V}(\rho))$. Now let $\tilde{\rho} \in B^\varepsilon(\mathcal{V}(\rho)) \implies \mathbf{D}(\mathcal{V}(\rho) \parallel \tilde{\rho}) \leq \varepsilon$. Using the data-processing inequality for any quantum channel, we have $\mathbf{D}(\mathcal{R}_\mathcal{V}(\mathcal{V}(\rho)) \parallel \mathcal{R}_\mathcal{V}(\tilde{\rho})) \leq \varepsilon \implies \mathbf{D}(\rho \parallel \mathcal{R}_\mathcal{V}(\tilde{\rho})) \leq \varepsilon \implies \tilde{\rho} \in B_{\mathcal{R}_\mathcal{V}}^\varepsilon(\rho)$. $\square$

A second technical property of the reversal channels is used for proving invariance under local transformations. There exist reversal channels for local transformations that preserve the reversal property $\mathcal{R}_\mathcal{V} \circ \mathcal{V} = \text{id}$ locally, in the sense $\mathcal{R}_\mathcal{V}(\mathcal{V}_A(\cdot) \otimes (\cdot)) \sim \text{id}_A \otimes (\cdot)$.

For mutual information quantities of types I using quantum states as arguments of the divergence, the following fact is used for proving Proposition 1.

Lemma 2. Let $V = V_A \otimes V_B$ be a local isometry. For any state ρ_A and any local quantum states σ_A , σ_B , there exists a reversal channel $\mathcal{R}_\mathcal{V}$ satisfying

$$\mathcal{R}_\mathcal{V}(\mathcal{V}_A(\rho_A) \otimes \sigma_B) = \rho_A \otimes \text{Tr}_A[\mathcal{R}_\mathcal{V}(\mathcal{V}_A(\sigma_A) \otimes \sigma_B)]. \quad (17)$$

Proof. Define the reversal channel with $\omega = \rho_A \otimes \omega_B$ for any local quantum state ω_B . Using the definition of $\mathcal{R}_V$ and $V_A^\dagger V_A = \mathbb{1}$, one can directly calculate:

$$\mathcal{R}_V(\mathcal{V}_A(\rho_A) \otimes \sigma_B) = \rho_A \otimes (V_B^\dagger \sigma_B V_B + (1 - \text{Tr}_B[V_B V_B^\dagger \sigma_B]) \omega_B) = \rho_A \otimes \text{Tr}_A[\mathcal{R}_V(\mathcal{V}_A(\sigma_A) \otimes \sigma_B)]. \quad (18)$$

□

For conditional entropies of types **H** using both quantum states and positive-semidefinite operators as arguments of the divergence, the following Lemma is used for proving invariance in the proof of Proposition 2.

Lemma 3. Let $V = U_A \otimes V_B$ with unitary U_A and isometric V_B . For any local quantum state σ_B , there exists a reversal channel $\mathcal{R}_V$ satisfying

$$\mathcal{R}_V(\mathbb{1}_A \otimes \sigma_B) = \mathbb{1}_A \otimes \text{Tr}_A[\mathcal{R}_V(\pi_A \otimes \sigma_B)], \quad (19)$$

where $\pi_A = \frac{1}{d_A} \mathbb{1}_A$ is the maximally mixed state of A with dimension d_A .

Proof. Define the reversal channel with $\omega = \pi_A \otimes \omega_B$ for any local quantum state ω_B . Considering $U_A^\dagger U_A = U_A U_A^\dagger = \mathbb{1}_A$, using the definition of $\mathcal{R}_V$ and calculating as in the proof of Lemma 2, one finds:

$$\mathcal{R}_V(\mathbb{1}_A \otimes \sigma_B) = \mathbb{1}_A \otimes (V_B^\dagger \sigma_B V_B + (1 - \text{Tr}_B[V_B V_B^\dagger \sigma_B]) \omega_B) = \mathbb{1}_A \otimes \text{Tr}_A[\mathcal{R}_V(\pi_A \otimes \sigma_B)]. \quad (20)$$

□

In the special case of smoothed states present in both arguments of the divergence as in $\mathbf{I}_4^\epsilon$, the following property is used for proving Proposition 2.

Lemma 4. Let $V = U_A \otimes V_B$ with unitary U_A and isometric V_B . For any state ρ_A and any quantum state σ_{AB} , there exists a reversal channel $\mathcal{R}_V$ satisfying

$$\mathcal{R}_V(\mathcal{V}_A(\rho_A) \otimes \sigma_B) = \rho_A \otimes \text{Tr}_A[\mathcal{R}_V(\sigma_{AB})]. \quad (21)$$

Proof. Set $\omega = \rho_A \otimes \omega_B$ for any quantum state ω_B . Using the definition of $\mathcal{R}_V$ and $U_A^\dagger U_A = \mathbb{1}$, we then have for any local isometry $V = U_A \otimes V_B$:

$$\mathcal{R}_V(\mathcal{V}_A(\rho_A) \otimes \sigma_B) = \rho_A \otimes V_B^\dagger \sigma_B V_B + \text{Tr}[\mathbb{1}_{AB} - \mathcal{V}_A(\rho_A) \otimes V_B V_B^\dagger \sigma_B] \rho_A \otimes \omega_B \quad (22)$$

$$= \rho_A \otimes (V_B^\dagger \sigma_B V_B + (1 - \text{Tr}_B[V_B V_B^\dagger \sigma_B]) \omega_B). \quad (23)$$

On the other hand, using $U_A U_A^\dagger = \mathbb{1}_A$, one finds

$$\rho_A \otimes \text{Tr}_A[\mathcal{R}_V(\sigma_{AB})] = \rho_A \otimes \text{Tr}_A[U_A^\dagger \otimes V_B^\dagger \sigma_{AB} U_A \otimes V_B + \text{Tr}[(\mathbb{1}_{AB} - U_A U_A^\dagger \otimes V_B V_B^\dagger) \sigma_{AB}] \rho_A \otimes \omega_B] \quad (24)$$

$$= \rho_A \otimes \text{Tr}_A[\mathbb{1}_A \otimes V_B^\dagger \sigma_{AB} \mathbb{1}_A \otimes V_B + (1 - \text{Tr}[\mathbb{1}_A \otimes V_B V_B^\dagger \sigma_{AB}]) \rho_A \otimes \omega_B] \quad (25)$$

$$= \rho_A \otimes (V_B^\dagger \sigma_B V_B + (1 - \text{Tr}_B[V_B V_B^\dagger \sigma_B]) \omega_B), \quad (26)$$

showing the claimed equality. □

Using these results, we now prove Propositions 1 and 2.

Proof of Proposition 1. Let $\mathcal{V} = \mathcal{V}_A \otimes \mathcal{V}_B$, where $\mathcal{V}_{(A/B)} : \mathcal{D}(\mathcal{H}_{(A/B)}) \rightarrow \mathcal{D}(\tilde{\mathcal{H}}_{(A/B)})$. Let $\text{im}(\mathcal{V}_{(A/B)}) \subset \mathcal{D}(\tilde{\mathcal{H}}_{(A/B)})$ denote the image of $\mathcal{V}_{(A/B)}$.

First consider type I_1 . Using $\text{Tr}_{(A/B)}[\mathcal{V}_A \otimes \mathcal{V}_B(\rho_{AB})] = \mathcal{V}_{(B/A)}(\rho_{B/A})$ and the invariance of any divergence under isometric evolution (1), one has:

$$I_1(\mathcal{V}(\rho_{AB})) = \mathbf{D}(\mathcal{V}(\rho_{AB}) || \text{Tr}_B[\mathcal{V}(\rho_{AB})] \otimes \text{Tr}_A[\mathcal{V}(\rho_{AB})]) \quad (27)$$

$$= \mathbf{D}(\mathcal{V}(\rho_{AB}) || \mathcal{V}(\rho_A \otimes \rho_B)) \quad (28)$$

$$= \mathbf{D}(\rho_{AB} || \rho_A \otimes \rho_B) \quad (29)$$

$$= I_1(\rho_{AB}). \quad (30)$$

Next, consider type I_2 . Again, using the invariance under isometric transformations, one has:

$$I_2(\rho_{AB}) = \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\rho_{AB} || \rho_A \otimes \sigma_B) \quad (31)$$

$$= \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\mathcal{V}(\rho_{AB}) || \mathcal{V}_A(\rho_A) \otimes \mathcal{V}_B(\sigma_B)) \quad (32)$$

$$= \inf_{\tilde{\sigma}_B \in \text{im}(\mathcal{V}_B)} \mathbf{D}(\mathcal{V}(\rho_{AB}) || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (33)$$

$$\geq \inf_{\tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\mathcal{V}(\rho_{AB}) || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (34)$$

$$= I_2(\mathcal{V}(\rho_{AB})) \quad (35)$$

Conversely, using the data-processing inequality for $\mathbf{D}$ with the reversal channel $\mathcal{R}_\mathcal{V}$, one has:

$$I_2(\mathcal{V}(\rho_{AB})) = \inf_{\tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\mathcal{V}(\rho_{AB}) || \text{Tr}_B[\mathcal{V}(\rho_{AB})] \otimes \tilde{\sigma}_B) \quad (36)$$

$$= \inf_{\tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\mathcal{V}(\rho_{AB}) || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (37)$$

$$\geq \inf_{\tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\rho_{AB} || \mathcal{R}_\mathcal{V}(\mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B)) \quad (38)$$

Using the reversal channel as in Lemma 2, with arbitrary $\hat{\sigma}_A \in \mathcal{D}(\mathcal{H}_A)$, yields therefore:

$$I_2(\mathcal{V}(\rho_{AB})) \geq \inf_{\tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\rho_{AB} || \rho_A \otimes \text{Tr}_A[\mathcal{R}_\mathcal{V}(\mathcal{V}_A(\hat{\sigma}_A) \otimes \tilde{\sigma}_B)]) \quad (39)$$

$$\geq \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} \mathbf{D}(\rho_{AB} || \rho_A \otimes \sigma_B) \quad (40)$$

$$= I_2(\rho_{AB}). \quad (41)$$

Note, that the second inequality holds because $\text{Tr}_A[\mathcal{R}_\mathcal{V}(\mathcal{V}_A(\hat{\sigma}_A) \otimes \tilde{\sigma}_B)] \in \mathcal{D}(\mathcal{H}_B)$. The inequalities below (31) and (39) together imply the claimed equality.

Finally, consider type I_3^ϵ . Using similar arguments and Lemma 1, one finds:

$$I_3^\epsilon(\rho_{AB}) = \inf_{\substack{\hat{\rho}_{AB} \in B^\epsilon(\rho_{AB}) \\ \sigma_B \in \mathcal{D}(\mathcal{H}_B)}} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \sigma_B) \quad (42)$$

$$= \inf_{\substack{\hat{\rho}_{AB} \in B^\epsilon_\mathcal{V}(\rho_{AB}) \\ \tilde{\sigma}_B \in \text{im}(\mathcal{V}_B)}} \mathbf{D}(\hat{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (43)$$

$$\geq \inf_{\substack{\hat{\rho}_{AB} \in B^\epsilon(\mathcal{V}(\rho_{AB})) \\ \tilde{\sigma}_B \in \mathcal{D}(\mathcal{H}_B)}} \mathbf{D}(\hat{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (44)$$

$$= I_3^\epsilon(\mathcal{V}(\rho_{AB})) \quad (45)$$

Conversely, using the reversal channel again as in Lemma 2 with arbitrary $\hat{\sigma}_A$ together with Lemma 1 and noting the set equality $\mathcal{R}_V[B_{\mathcal{R}_V}^{\varepsilon}(\rho_{AB})] = B^{\varepsilon}(\rho_{AB})$, one concludes:

$$\mathbf{I}_3^{\varepsilon}(\mathcal{V}(\rho_{AB})) = \inf_{\substack{\tilde{\rho}_{AB} \in B^{\varepsilon}(\mathcal{V}(\rho_{AB})) \\ \tilde{\sigma}_B \in \mathcal{D}(\tilde{\mathcal{H}}_B)}} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (46)$$

$$\geq \inf_{\substack{\tilde{\rho}_{AB} \in B_{\mathcal{R}_V}^{\varepsilon}(\mathcal{V}(\rho_{AB})) \\ \tilde{\sigma}_B \in \mathcal{D}(\tilde{\mathcal{H}}_B)}} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B) \quad (47)$$

$$\geq \inf_{\substack{\tilde{\rho}_{AB} \in B_{\mathcal{R}_V}^{\varepsilon}(\mathcal{V}(\rho_{AB})) \\ \tilde{\sigma}_B \in \mathcal{D}(\tilde{\mathcal{H}}_B)}} \mathbf{D}(\mathcal{R}_V(\tilde{\rho}_{AB}) || \mathcal{R}_V(\mathcal{V}_A(\rho_A) \otimes \tilde{\sigma}_B)) \quad (48)$$

$$= \inf_{\substack{\hat{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB}) \\ \hat{\sigma}_B \in \mathcal{D}(\tilde{\mathcal{H}}_B)}} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \text{Tr}_A[\mathcal{R}_V(\mathcal{V}_A(\hat{\sigma}_A) \otimes \hat{\sigma}_B)]) \quad (49)$$

$$\geq \inf_{\substack{\hat{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB}) \\ \sigma_B \in \mathcal{D}(\tilde{\mathcal{H}}_B)}} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \sigma_B) \quad (50)$$

$$= \mathbf{I}_3^{\varepsilon}(\rho_{AB}) \quad (51)$$

□

Proof of Proposition 2. Let $\mathcal{V} = \mathcal{V}_A \otimes \mathcal{V}_B$, $\mathcal{V}_A(\cdot) = U_A(\cdot)U_A^\dagger$, $\mathcal{V}_A : \mathcal{D}(\mathcal{H}_A) \mapsto \mathcal{D}(\mathcal{H}_A)$ with unitary U_A and $\mathcal{V}_B : \mathcal{D}(\mathcal{H}_B) \mapsto \mathcal{D}(\tilde{\mathcal{H}}_B)$ with isometric $\mathcal{V}_B$. Let $\text{im}(\mathcal{V}_B) \subset \mathcal{D}(\tilde{\mathcal{H}}_B)$ denote the image of $\mathcal{V}_B$.

First, consider the type $\mathbf{I}_4^{\varepsilon}$. Using the invariance of $\mathbf{D}$ under isometric transformations, and Lemma 1, one has:

$$\mathbf{I}_4^{\varepsilon}(\rho_{AB}) = \inf_{\hat{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB})} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \hat{\rho}_B) \quad (52)$$

$$= \inf_{\hat{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB})} \mathbf{D}(\mathcal{V}(\hat{\rho}_{AB}) || \mathcal{V}(\rho_A \otimes \hat{\rho}_B)) \quad (53)$$

$$= \inf_{\tilde{\rho}_{AB} \in B_{\mathcal{V}}^{\varepsilon}(\rho_{AB})} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\rho}_B) \quad (54)$$

$$\geq \inf_{\tilde{\rho}_{AB} \in B^{\varepsilon}(\mathcal{V}(\rho_{AB}))} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\rho}_B) \quad (55)$$

$$= \mathbf{I}_4^{\varepsilon}(\mathcal{V}(\rho_{AB})). \quad (56)$$

Conversely, using again Lemma 1 and applying the data-processing inequality for the reversal channel $\mathcal{R}_V$ as in Lemma 4, one finds:

$$\mathbf{I}_4^{\varepsilon}(\mathcal{V}(\rho_{AB})) = \inf_{\tilde{\rho}_{AB} \in B^{\varepsilon}(\mathcal{V}(\rho_{AB}))} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\rho}_B) \quad (57)$$

$$\geq \inf_{\tilde{\rho}_{AB} \in B_{\mathcal{R}_V}^{\varepsilon}(\rho_{AB})} \mathbf{D}(\tilde{\rho}_{AB} || \mathcal{V}_A(\rho_A) \otimes \tilde{\rho}_B) \quad (58)$$

$$\geq \inf_{\tilde{\rho}_{AB} \in B_{\mathcal{R}_V}^{\varepsilon}(\rho_{AB})} \mathbf{D}(\mathcal{R}_V(\tilde{\rho}_{AB}) || \mathcal{R}_V(\mathcal{V}_A(\rho_A) \otimes \tilde{\rho}_B)). \quad (59)$$

Now, since U_A is unitary, Lemma 4 implies together with the set equality $\mathcal{R}_V[B_{\mathcal{R}_V}^{\varepsilon}(\rho_{AB})] = B^{\varepsilon}(\rho_{AB})$,

$$\mathbf{I}_4^{\varepsilon}(\mathcal{V}(\rho_{AB})) \geq \inf_{\tilde{\rho}_{AB} \in B_{\mathcal{R}_V}^{\varepsilon}(\rho_{AB})} \mathbf{D}(\mathcal{R}_V(\tilde{\rho}_{AB}) || \rho_A \otimes \text{Tr}_A[\mathcal{R}_V(\tilde{\rho}_{AB})]) \quad (60)$$

$$= \inf_{\hat{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB})} \mathbf{D}(\hat{\rho}_{AB} || \rho_A \otimes \hat{\rho}_B) \quad (61)$$

$$= \mathbf{I}_4^{\varepsilon}(\rho_{AB}), \quad (62)$$

proving the claimed property.

Finally, consider the conditional entropy types $\mathbf{H}_1, \mathbf{H}_2$ and $\mathbf{H}_3^\varepsilon$. Noting $\mathcal{V}_A(\mathbb{1}_A) = \mathbb{1}_A$ due to the fact that $\mathcal{V}_A$ is unitary, the proofs of local invariance are equivalent to those for $\mathbf{I}_1, \mathbf{I}_2$ and $\mathbf{I}_3^\varepsilon$ given in the proof of Proposition 1, respectively, if ρ_A is replaced by $\mathbb{1}_A$ and Lemma 3 is used instead of Lemma 2. $\square$

4. Summary and Conclusions

In this work, we defined types of quantum information quantities based on generalized divergences and analyzed their properties regarding local isometric and unitary transformations. Leveraging properties of the reversal channel associated with the local transformation, we proved that several types of information quantities imply invariance under any local isometric transformation, while others are invariant if one of the subsystems is transformed by a unitary. Two main technical results are utilized to prove the local invariance. First, the smoothing environment required by some types of information quantities is related to its image and its pre-image under the isometric channel and a corresponding reversal channel, implying the set relation of Lemma 1. Second, the action of the reversal channel for a local isometric channel is characterized in Lemmas 2–4. Combining these insights with the data-processing inequality of any generalized divergence allows us to derive the main results regarding invariance for several types of information quantities under local isometric or unitary transformation as presented in Propositions 1 and 2.

Since methods for proving the invariance do not depend on the specific form of the generalized divergence used to define the information quantity, the invariances hold for numerous key quantities with operational relevance in quantum information processing. These invariances can enable more efficient computation of quantum information quantities by reducing complex states to simpler, equivalent forms with symmetries or smaller dimensions. For instance, consider the setting of quantum cryptography including a third party holding the purifying system of a bipartite state. If the output of a tripartite quantum channel is a pure state, it is equivalent to the purified output of a corresponding bipartite channel up to a local isometry in the purifying system. The dimension of the purification system of the latter state, however, may be smaller than the dimension of the former purification system. Consequently, the evaluation of relevant information quantities like the private information may be significantly improved by calculating it for the smaller but equivalent output state. Such invariances reduce computational overhead and unlock flexibility in protocol design. Given that the protocol performance is measured by an invariant information quantity, local unitary or isometric operations like encoding or correction operations can be added, modified or removed while preserving performance metrics.

Note, that the general proofs for local invariance depend only on the monotonicity by the data-processing inequality and the reversal properties of the isometric channel as given in Lemma 2–4. Therefore, local invariance may be shown similarly for other channels and corresponding reversals, such as the Petz recovery map (cf. [10]).

The computation of generalized information quantities remains a significant challenge, in general. While some quantities like the hypothesis testing mutual information can be calculated efficiently, e.g., by semidefinite programs [23,24], others can only be approximated (see Ref. [25] regarding the smooth max-relative entropy) or lack a general and efficient solution like the smooth max-mutual information. Our results may enable new methods to bound or approximate these quantities by leveraging state equivalences. Finally, we note that the invariance under local isometries, notably local Clifford and unitary operations, plays a pivotal role in the classification and optimization of graph states as used in error-correction [26,27] or measurement-based quantum computation [28,29], allowing for the optimization of codes or resource states with respect to an invariant information measure.

In conclusion, the results of this contribution allow us to derive the behavior of general types of information quantities under local isometric or unitary transformations. Therefore, they improve the capability to characterize and compute information quantities relevant throughout the vast field of quantum information processing.

Author Contributions: Conceptualization, C.P.; validation, C.P., T.C.S. and B.C.H.; formal analysis, C.P.; writing—original draft preparation, C.P.; writing—review and editing, C.P., T.C.S. and B.C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [10.55776/P36102]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Neumann, J.v. Thermodynamik quantenmechanischer Gesamtheiten. *Nachrichten Von Der Ges. Der Wiss. Zu Göttingen Math.-Phys. Kl.* **1927**, *1927*, 273–291.
2. Wilde, M.M. *Quantum Information Theory*; Cambridge University Press: Cambridge, UK, 2013. [\[CrossRef\]](#)
3. Umegaki, H. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math. Semin. Rep.* **1962**, *14*, 59–85. [\[CrossRef\]](#)
4. Hiai, F.; Petz, D. The proper formula for relative entropy and its asymptotics in quantum probability. *Commun. Math. Phys.* **1991**, *143*, 99–114. [\[CrossRef\]](#)
5. Devetak, I.; Winter, A. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2005**, *461*, 207–235. [\[CrossRef\]](#)
6. Tomamichel, M. *Quantum Information Processing with Finite Resources—Mathematical Foundations*; Springer: Cham, Switzerland, 2016; Volume 5. [\[CrossRef\]](#)
7. Renner, R. Security of Quantum Key Distribution. *arXiv* **2006**, arXiv:quant-ph/0512258. [\[CrossRef\]](#)
8. Wang, L.; Renner, R. One-Shot Classical-Quantum Capacity and Hypothesis Testing. *Phys. Rev. Lett.* **2012**, *108*, 200501. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Buscemi, F.; Datta, N. The Quantum Capacity of Channels With Arbitrarily Correlated Noise. *IEEE Trans. Inf. Theory* **2010**, *56*, 1447–1460. [\[CrossRef\]](#)
10. Khatri, S.; Wilde, M.M. Principles of Quantum Communication Theory: A Modern Approach. *arXiv* **2024**, arXiv:2011.04672. [\[CrossRef\]](#)
11. Petz, D. Quasi-entropies for States of a von Neumann Algebra. *Publ. Res. Inst. Math. Sci.* **1985**, *21*, 787–800. [\[CrossRef\]](#)
12. Petz, D. Quasi-entropies for finite quantum systems. *Rep. Math. Phys.* **1986**, *23*, 57–65. [\[CrossRef\]](#)
13. Tomamichel, M.; Colbeck, R.; Renner, R. A Fully Quantum Asymptotic Equipartition Property. *IEEE Trans. Inf. Theory* **2009**, *55*, 5840–5847. [\[CrossRef\]](#)
14. Müller-Lennert, M.; Dupuis, F.; Szehr, O.; Fehr, S.; Tomamichel, M. On quantum Rényi entropies: A new generalization and some properties. *J. Math. Phys.* **2013**, *54*, 122203. [\[CrossRef\]](#)
15. Wilde, M.M.; Winter, A.; Yang, D. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Commun. Math. Phys.* **2014**, *331*, 593–622. [\[CrossRef\]](#)
16. Petz, D.; Ruskai, M.B. Contraction of Generalized Relative Entropy Under Stochastic Mappings on Matrices. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1998**, *01*, 83–89. [\[CrossRef\]](#)
17. Matsumoto, K. A New Quantum Version of f-Divergence. In *Reality and Measurement in Algebraic Quantum Theory*; Ozawa, M., Butterfield, J., Halvorson, H., Rédei, M., Kitajima, Y., Buscemi, F., Eds.; Springer: Singapore, 2018; pp. 229–273. [\[CrossRef\]](#)
18. Polyanskiy, Y.; Verdú, S. Arimoto channel coding converse and Rényi divergence. In Proceedings of the 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, USA, 29 September–1 October 2010; pp. 1327–1333. [\[CrossRef\]](#)
19. Sharma, N.; Warsi, N.A. Fundamental Bound on the Reliability of Quantum Information Transmission. *Phys. Rev. Lett.* **2013**, *110*, 080501. [\[CrossRef\]](#)
20. Rastegin, A.E. Sine distance for quantum states. *arXiv* **2006**, arXiv:quant-ph/0602112. [\[CrossRef\]](#)

21. Uhlmann, A. The “transition probability” in the state space of a $\ast$ -algebra. *Rep. Math. Phys.* **1976**, *9*, 273–279. [[CrossRef](#)]
22. Datta, N. Min- and Max-Relative Entropies and a New Entanglement Monotone. *IEEE Trans. Inf. Theory* **2009**, *55*, 2816–2826. [[CrossRef](#)]
23. Dupuis, F.; Krämer, L.; Faist, P.; Renes, J.M.; Renner, R. Generalized entropies. In *XVIIIth International Congress on Mathematical Physics*; World Scientific: Singapore, 2012; pp. 134–153. [[CrossRef](#)]
24. Datta, N.; Tomamichel, M.; Wilde, M.M. On the second-order asymptotics for entanglement-assisted communication. *Quantum Inf. Process.* **2016**, *15*, 2569–2591. [[CrossRef](#)]
25. Nuradha, T.; Wilde, M.M. Fidelity-Based Smooth Min-Relative Entropy: Properties and Applications. *IEEE Trans. Inf. Theory* **2024**, *70*, 4170–4196. [[CrossRef](#)]
26. Bény, C.; Oreshkov, O. General Conditions for Approximate Quantum Error Correction and Near-Optimal Recovery Channels. *Phys. Rev. Lett.* **2010**, *104*, 120501. [[CrossRef](#)]
27. Sarkar, R.; Yoder, T.J. A graph-based formalism for surface codes and twists. *Quantum* **2024**, *8*, 1416. [[CrossRef](#)]
28. Hein, M.; Eisert, J.; Briegel, H.J. Multiparty entanglement in graph states. *Phys. Rev. A* **2004**, *69*, 062311. [[CrossRef](#)]
29. Briegel, H.J.; Browne, D.E.; Dür, W.; Raussendorf, R.; Van den Nest, M. Measurement-based quantum computation. *Nat. Phys.* **2009**, *5*, 19–26. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

B.9. Computation of the Smooth Max-Mutual Information via Semidefinite Programming

Computation of the Smooth Max-Mutual Information via Semidefinite Programming

Christopher Popp¹, Tobias C. Sutter², and Beatrix C. Hiesmayr³

^{1,2,3}University of Vienna, Faculty of Physics, Währingerstrasse 17, 1090 Vienna.

¹christopher.popp@univie.ac.at

²tobias.christoph.sutter@univie.ac.at

³beatrix.hiesmayr@univie.ac.at

Abstract

We present an iterative algorithm based on semidefinite programming (SDP) for computing the quantum smooth max-mutual information $I_{\max}^{\varepsilon}(\rho_{AB})$ of bipartite quantum states in any dimension. The algorithm is accurate if a rank condition for marginal states within the smoothing environment is satisfied and provides an upper bound otherwise. Central to our method is a novel SDP, for which we establish primal and dual formulations and prove strong duality. With the direct application of bounding the one-shot distillable key of a quantum state, this contribution extends SDP-based techniques in quantum information theory. Thereby it improves the capabilities to compute or estimate information measures with application to various quantum information processing tasks.

1 Introduction

Quantum information theory has extended our understanding of information processing, communication, and cryptography by leveraging the principles of quantum mechanics [1]. Central to this field are entropic quantities that quantify correlations, uncertainties, and operational capabilities of quantum systems (cf. [2] for a comprehensive overview). In particular, considering small deviations from the quantum states at hand, smoothed versions of entropic and information measures, such as the smooth max-mutual information $I_{\max}^{\varepsilon}(\rho_{AB})$ [3, 4] for bipartite quantum states ρ_{AB} are essential for characterizing one-shot operational tasks (cf. Ref. [5]) in noisy environments with finite resources.

The smooth max-mutual information captures the correlation between subsystems A and B , accounting for small errors due to noise or approximations due to limited resources. This quantity is particularly relevant in cryptographic settings, where it can be combined with other measures like the hypothesis testing mutual information [6] to bound the one-shot distillable key [7]. This provides practical estimates for secure key rates achievable from quantum states [2]. However, computing $I_{\max}^{\varepsilon}$ exactly remains a computationally challenging optimization problem. Existing methods in quantum information theory have increasingly relied on semidefinite programming (SDP) to compute, approximate, or bound entropic quantities efficiently [8–12]. For instance, SDP formulations have been successfully applied to compute the hypothesis testing mutual information and conditional smooth min- and max-entropies [13]. Despite these advances, an SDP-based approach for the smooth max-mutual information is lacking. In Ref. [13], the authors recently used a so-called mountain-climbing-type algorithm [14] to use SDPs iteratively to calculate the smooth min-entropy, an information quantity not allowing a direct formulation as an SDP. In this work, we

employ a similar approach by introducing an SDP-based iterative algorithm for computing $I_{max}^\varepsilon(\rho_{AB})$. Crucially, the quantity considered in Ref. [13] and I_{max}^ε differ in their smoothing structure, requiring an alternative approach, including a novel SDP at the core of the algorithm. We characterize this SDP by its prime-dual formulations, show strong duality, and prove that the iterative algorithm computes the smooth max-mutual information accurately, assuming a rank constraint regarding the marginal states in the smoothing neighborhood.

Section 2 establishes the notation and the information quantities and SDP methods applied in this work. In Section 3, we present an SDP-based algorithm to compute the smooth max-mutual information. We characterize the underlying SDP by its primal and dual form and prove the accuracy of the algorithm under the assumption of a rank constraint. Finally, we conclude our results in Section 4.

2 Background

2.1 Notation

Consider the bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ of two parties, A and B . The set of linear operators on $\mathcal{H}$ is denoted by $L(\mathcal{H})$. The set of positive semidefinite operators $X \geq 0$ is denoted as $L_+(\mathcal{H})$ and the set of density operators, or quantum states, satisfying $\rho \in L_+(\mathcal{H})$ and $\text{Tr}(\rho) = 1$ is written as $\mathcal{D}(\mathcal{H})$. The marginal states of a multipartite state are denoted as $\rho_{A/B} := \text{Tr}_{B/A}[\rho_{AB}]$. If X is positive definite, we denote it by $X > 0$. A superoperator Φ is a linear operator mapping $L(\mathcal{H}) \rightarrow L(\mathcal{H})$. Φ is hermiticity preserving, if $\Phi(X)$ is hermitian for every hermitian X . The adjoint superoperator $\Phi^\dagger$ is defined by satisfying $\text{Tr}[\Phi(X)Y] = \text{Tr}[X\Phi^\dagger(Y)]$ for all $X, Y \in L(\mathcal{H})$.

2.2 The max-relative entropy and the smooth max-mutual information

The information measure analyzed in this work involves an optimization over an environment around the quantum state of interest, called smoothing. For that purpose, we use the so-called sine distance [15] based on the fidelity [16]:

Definition 1 (Fidelity, sine distance).

The fidelity of two quantum states ρ and σ is defined as follows:

$$\mathcal{F}(\rho, \sigma) := \left(\text{Tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}] \right)^2 \quad (1)$$

The sine distance is defined as:

$$P(\rho, \sigma) := \sqrt{1 - \mathcal{F}(\rho, \sigma)} \quad (2)$$

Using this distance measure, the smoothing environment is defined as a ball around the quantum state for which the information measure is to be evaluated.

Definition 2 (Smoothing environment).

Using the sine distance, the smoothing environment $B^\varepsilon(\rho)$ around a state ρ is defined as

$$B^\varepsilon(\rho) := \{\tilde{\rho} \in \mathcal{D}(\mathcal{H}) : P(\rho, \tilde{\rho}) \leq \varepsilon\} = \{\tilde{\rho} : \mathcal{F}(\rho, \tilde{\rho}) \geq 1 - \varepsilon^2\}. \quad (3)$$

The quantity of interest in this work is related to the max-relative entropy D_{max} [4]. Here, we analyze a smoothed generalized mutual information I_{max}^ε [3] based on D_{max} , which can be used for quantifying the capability of a quantum state to be used for the task of secret-key distillation (cf., e.g., [2, 7]).

Definition 3 (Max-relative entropy, smooth max-mutual information).

$$D_{max}(\rho||\sigma) := \begin{cases} \log_2 \left\| \sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}} \right\|_{\infty} & \text{if } \text{supp}(\rho) \subseteq \text{supp}(\sigma) \\ +\infty & \text{else} \end{cases} \quad (4)$$

$$I_{max}^{\varepsilon}(\rho_{AB}) := \inf_{\tilde{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB})} D_{max}(\tilde{\rho}_{AB} || \rho_A \otimes \tilde{\rho}_B) \quad (5)$$

where $B^{\varepsilon}(\rho_{AB})$ is as in Definition 2, $\|\cdot\|_{\infty}$ is the spectral norm and $\tilde{\rho}_B := \text{Tr}_A[\tilde{\rho}_{AB}]$.

2.3 Semidefinite Programming

Semidefinite programs are constrained optimization problems with positive semidefinite optimization variables that often occur in quantum information theory.

Definition 4 (Semidefinite program).

Let Φ be a hermiticity preserving superoperator and A and B be hermitian operators. A semidefinite program (SDP) corresponds to the following two optimization problems over positive semidefinite operators X, Y .

The primal SDP:

$$\begin{aligned} & \text{maximize } \text{Tr}[AX] \\ & \text{subject to } \Phi(X) \leq B, X \geq 0. \end{aligned} \quad (6)$$

The dual SDP:

$$\begin{aligned} & \text{minimize } \text{Tr}[BY] \\ & \text{subject to } \Phi^{\dagger}(Y) \geq A, Y \geq 0, \end{aligned} \quad (7)$$

where $\Phi^{\dagger}$ is the adjoint superoperator of Φ .

A variable X satisfying the constraints of the primal SDP is called a feasible point. If X satisfies $\Phi(X) < B$ and $X > 0$, it is called a strictly feasible point. Similarly, Y is feasible if it satisfies the constraints of the dual SDP and strictly feasible if the constraints are strictly satisfied. An SDP is said to satisfy the strong duality property if the optimal values of (6) and (7) of the primal and dual problems are equal.

3 Computing the smooth max-mutual information with an SDP-based algorithm

The max-relative entropy $D_{max}(\rho||\sigma)$ of two quantum states ρ and σ can be characterized by the following optimization problem (cf. Ref. [2]):

$$D_{max}(\rho||\sigma) = \log_2 \inf_{\lambda \geq 0} \{\lambda : \lambda \sigma - \rho \geq 0\}. \quad (8)$$

The inner optimization can be formulated as an SDP in its dual form by setting $Y \equiv \lambda$, $A = \rho$, $B = 1$, $\Phi^{\dagger}(Y) = Y\sigma$ in Definition 4. The same does not hold for the smooth max-mutual information $I_{max}^{\varepsilon}(\rho_{AB})$. This can be seen by writing I_{max}^{ε} using the characterization for the max-relative entropy (8), implying

$$I_{max}^{\varepsilon}(\rho_{AB}) := \log_2 \left(\inf_{\tilde{\rho}_{AB} \in B^{\varepsilon}(\rho_{AB})} \inf_{\lambda \geq 0} \{\lambda : \lambda \rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} \geq 0\} \right) := \log_2(\lambda_{min}). \quad (9)$$

As the constraint involves a bilinear term in the variables $\tilde{\rho}_{AB}$ and λ , this cannot be directly solved by a SDP. We propose an iterative approach applicable for states ρ_{AB} and $\varepsilon > 0$ for which $B^\varepsilon(\rho_{AB})$ contains only full-rank marginal states $\rho_A := \text{Tr}_B[\rho_{AB}]$ and $\tilde{\rho}_B := \text{Tr}_A[\tilde{\rho}_{AB}]$. Alternately fixing one of the variables and solving an SDP for the other, this seesaw or mountain-climbing algorithm converges to the optimal solution after a finite amount of iterations [14], providing an upper bound approximating $I_{max}^\varepsilon(\rho_{AB})$. Define the i -th iteration of the procedure as follows:

Algorithm 1.

1. If $i = 1$, choose $\rho_{AB}^i = \rho_{AB}$.
2. Solve the following SDP, with the solution defining

$$\lambda^i := \inf_{\lambda \geq 0} \{ \lambda : \lambda \rho_A \otimes \rho_B^i - \rho_{AB}^i \geq 0 \}. \quad (10)$$

3. Solve the following SDP, with the solution defining

$$\mu^i := \sup_{\substack{\mu \geq 0, \\ \tilde{\rho}_{AB} \in B^\varepsilon(\rho_{AB})}} \{ \mu : \lambda^i \rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu \mathbb{1} \geq 0 \} \quad (11)$$

4. If $\mu^i = 0$ stop and set $\lambda_{min} = \lambda^i$. Else define ρ_{AB}^{i+1} to be state $\tilde{\rho}_{AB} \in B^\varepsilon(\rho_{AB})$ maximizing the expression in step 3 and use it in step 2 for the next iteration.

Note that for all iterations, $B^\varepsilon(\rho_{AB})$ remains fixed around ρ_{AB} . We now characterize the SDP of (11) by its primal and dual form and show strong duality.

Proposition 2. *The primal and dual SDPs of*

$$\sup_{\substack{\mu \geq 0, \\ \tilde{\rho}_{AB} \in B^\varepsilon(\rho_{AB})}} \{ \mu : \lambda^i \rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu \mathbb{1} \geq 0 \} \quad (12)$$

can be written in the following forms:

$$\sup_{\substack{\mu \geq 0, \\ \tilde{\rho}_{AB} \in \mathcal{D}(\mathcal{H}), \\ X \in L(\mathcal{H})}} \left\{ \mu : \lambda \rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu \mathbb{1} \geq 0, \text{Re}\{\text{Tr}[X]\} \geq \sqrt{1 - \varepsilon^2}, \begin{pmatrix} \rho_{AB} & X \\ X^\dagger & \tilde{\rho}_{AB} \end{pmatrix} \geq 0 \right\} \quad (\text{primal}), \quad (13)$$

$$\inf_{\substack{W, D \geq 0, \\ \eta \in \mathbb{R}, \nu \geq 0}} \left\{ \eta - 2\nu\sqrt{1 - \varepsilon^2} + \text{Tr}[\rho D] : \text{Tr}[W] \geq 1, \begin{pmatrix} D & \nu \mathbb{1} \\ \nu \mathbb{1} & \eta \mathbb{1} - (\lambda M - W) \end{pmatrix} \geq 0 \right\} \quad (\text{dual}), \quad (14)$$

and strong duality holds.

Proof. First note that by (3), the ε -ball based on the sine distance can be characterized by the root fidelity, for which we can use the characterization $\sqrt{\mathcal{F}}(\rho, \tilde{\rho}) := \sqrt{\mathcal{F}(\rho, \tilde{\rho})} = \sup_{X \in L(\mathcal{H})} \left\{ \text{Re}\{\text{Tr}[X]\} : \begin{pmatrix} \rho & X \\ X^\dagger & \tilde{\rho} \end{pmatrix} \geq 0 \right\}$ [11] to write the optimization problem in (12) as:

$$\sup_{\substack{\mu \geq 0, \\ \tilde{\rho}_{AB} \in \mathcal{D}(\mathcal{H}), \\ X \in L(\mathcal{H})}} \left\{ \mu : \lambda \rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu \mathbb{1} \geq 0, \text{Re}\{\text{Tr}[X]\} \geq \sqrt{1 - \varepsilon^2}, \begin{pmatrix} \rho_{AB} & X \\ X^\dagger & \tilde{\rho}_{AB} \end{pmatrix} \geq 0 \right\}. \quad (15)$$

For the primal problem in the standard form $\sup_{Z \geq 0} \{\text{Tr}[AZ] : \Phi(Z) \leq B\}$, where A and B are hermitian operators with suitable block sizes so that all matrix products are well defined and Φ is a hermiticity-preserving superoperator, one finds:

$$Z = \begin{pmatrix} \mu & 0 & 0 \\ 0 & D & X \\ 0 & X^\dagger & \tilde{\rho}_{AB} \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -\sqrt{1-\varepsilon^2} & 0 \\ 0 & 0 & 0 & 0 & \begin{bmatrix} \rho_{AB} & 0 \\ 0 & 0 \end{bmatrix} \end{pmatrix} \quad (16)$$

$$\Phi(Z) = \begin{pmatrix} -(\lambda\rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu\mathbb{1}) & 0 & 0 & 0 & 0 \\ 0 & \text{Tr}[\tilde{\rho}_{AB}] & 0 & 0 & 0 \\ 0 & 0 & -\text{Tr}[\tilde{\rho}_{AB}] & 0 & 0 \\ 0 & 0 & 0 & -\text{Re}\{\text{Tr}[X]\} & 0 \\ 0 & 0 & 0 & 0 & \begin{bmatrix} 0 & -X \\ -X^\dagger & -\tilde{\rho}_{AB} \end{bmatrix} \end{pmatrix}. \quad (17)$$

We use this to derive the dual problem and show strong duality. The dual problem is $\inf_{Y \geq 0} \{\text{Tr}[BY] : \Phi^\dagger(Y) \geq A\}$, where $\text{Tr}[\Phi(Z)Y] = \text{Tr}[Z\Phi^\dagger(Y)]$ defines the adjoint superoperator. To derive $\Phi^\dagger$, set

$$Y = \begin{pmatrix} W & 0 & 0 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 & 0 \\ 0 & 0 & \mu_2 & 0 & 0 \\ 0 & 0 & 0 & \nu & 0 \\ 0 & 0 & 0 & 0 & \begin{bmatrix} D_1 & V \\ V^\dagger & D_2 \end{bmatrix} \end{pmatrix}, \quad (18)$$

with $W \in L_+(\mathcal{H})$, $D_1, D_2, V \in L(\mathcal{H})$ such that $Y \geq 0$ and $\text{Tr}[BY]$ is well-defined. We then have

$$\begin{aligned} \text{Tr}[\Phi(Z)Y] &= \text{Tr}[-(\lambda\rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB} - \mu\mathbb{1})W] + \text{Tr}[\tilde{\rho}_{AB}](\mu_1 - \mu_2) - \text{Re}\{\text{Tr}[X]\}\nu \\ &\quad - \text{Tr} \left[\begin{pmatrix} 0 & X \\ X^\dagger & \tilde{\rho}_{AB} \end{pmatrix} \begin{pmatrix} D_1 & V \\ V^\dagger & D_2 \end{pmatrix} \right] \end{aligned} \quad (19)$$

$$\begin{aligned} &= \mu \text{Tr}[W] - \text{Tr}[(\lambda\rho_A \otimes \tilde{\rho}_B - \tilde{\rho}_{AB})W] + \text{Tr}[(\mu_1 - \mu_2)\tilde{\rho}_{AB}] - \text{Re}\{\text{Tr}[X]\}\nu \\ &\quad - \text{Tr}[XV^\dagger + X^\dagger V + \tilde{\rho}_{AB}D_2] \end{aligned} \quad (20)$$

$$= \mu \text{Tr}[W] + \text{Tr}\{[(\mu_1 - \mu_2)\mathbb{1} - D_2 - (\lambda M - W)]\tilde{\rho}_{AB}\} - \text{Re}\{\text{Tr}[X(\nu\mathbb{1} + 2V^\dagger)]\}. \quad (21)$$

In the last equation, we have defined the operator M , using an ancillary space A' of the same size as A :

$$M := \mathbb{1}_A \otimes \{\text{Tr}_{A'}[W_{A'B}(\rho_{A'} \otimes \mathbb{1}_B)]\} \implies \text{Tr}_{AB}[M\tilde{\rho}_{AB}] = \text{Tr}_{A'B}[(\rho_{A'} \otimes \tilde{\rho}_B)W]. \quad (22)$$

This implies the following three relations:

1.

$$\text{Tr}[\Phi(Z)Y] = \mu \text{Tr}[W] + \text{Tr} \left[\begin{pmatrix} 0 & \frac{\nu}{2}\mathbb{1} + V \\ \frac{\nu}{2}\mathbb{1} + V^\dagger & (\mu_1 - \mu_2)\mathbb{1} - D_2 - (\lambda M - W) \end{pmatrix} \begin{pmatrix} D & X \\ X^\dagger & \tilde{\rho}_{AB} \end{pmatrix} \right] \quad (23)$$

$$\Rightarrow \Phi^\dagger(Y) = \begin{pmatrix} \text{Tr}[W] & 0 \\ 0 & \begin{bmatrix} 0 & \frac{\nu}{2}\mathbb{1} + V \\ \frac{\nu}{2}\mathbb{1} + V^\dagger & (\mu_1 - \mu_2)\mathbb{1} - D_2 - (\lambda M - W) \end{bmatrix} \end{pmatrix} \quad (24)$$

2.

$$\Phi^\dagger(Y) \geq A \Leftrightarrow \text{Tr}[W] \geq 1, \begin{pmatrix} 0 & \frac{\nu}{2}\mathbb{1} + V \\ \frac{\nu}{2}\mathbb{1} + V^\dagger & (\mu_1 - \mu_2)\mathbb{1} - D_2 - (\lambda M - W) \end{pmatrix} \geq 0 \quad (25)$$

3.

$$\text{Tr}[BY] = \mu_1 - \mu_2 - \nu\sqrt{1 - \varepsilon^2} + \text{Tr}[\rho D_1]. \quad (26)$$

Hence, the dual program can be written as follows:

$$\inf_{\substack{W \geq 0 \\ \mu_1, \mu_2, \nu \geq 0}} \left\{ \mu_1 - \mu_2 - \nu\sqrt{1 - \varepsilon^2} + \text{Tr}[\rho D_1] : \right. \\ \left. \text{Tr}[W] \geq 1, \begin{pmatrix} 0 & \frac{\nu}{2}\mathbb{1} + V \\ \frac{\nu}{2}\mathbb{1} + V^\dagger & (\mu_1 - \mu_2)\mathbb{1} - D_2 - (\lambda M - W) \end{pmatrix} \geq 0, \begin{pmatrix} D_1 & V \\ V^\dagger & D_2 \end{pmatrix} \geq 0 \right\}. \quad (27)$$

We can simplify the program by defining $\eta := (\mu_1 - \mu_2) \in \mathbb{R}$ and setting $\nu/2 \rightarrow \nu \geq 0$ in (27), yielding

$$\begin{pmatrix} 0 & \nu\mathbb{1} + V \\ \nu\mathbb{1} + V^\dagger & \eta\mathbb{1} - D_2 - (\lambda M - W) \end{pmatrix} \geq 0, \begin{pmatrix} D_1 & V \\ V^\dagger & D_2 \end{pmatrix} \geq 0. \quad (28)$$

We can simplify this further by noting that the first condition requires $V = -\nu\mathbb{1}$, which implies $D_2 \leq \eta\mathbb{1} - (\lambda M - W)$. Setting $D_2 = \eta\mathbb{1} - (\lambda M - W)$ and $V = -\nu\mathbb{1}$, one finds with the second constraint

$$\begin{pmatrix} D_1 & \nu\mathbb{1} \\ \nu\mathbb{1} & \eta - (\lambda M - W) \end{pmatrix} \geq 0, \quad (29)$$

which holds if and only if (28) holds. For invertible D_1 or $\nu = 0$, this is implied by the Schur complement, and in case D_1 is singular and $\nu > 0$, neither (28) nor (29) can be fulfilled. Hence, we can write the final program using (29) and renaming $D_1 \rightarrow D$ as:

$$\inf_{\substack{W, D \geq 0 \\ \eta \in \mathbb{R}, \nu \geq 0}} \left\{ \eta - 2\nu\sqrt{1 - \varepsilon^2} + \text{Tr}[\rho D] : \text{Tr}[W] \geq 1, \begin{pmatrix} D & \nu\mathbb{1} \\ \nu\mathbb{1} & \eta\mathbb{1} - (\lambda M - W) \end{pmatrix} \geq 0 \right\}. \quad (30)$$

Finally, we conclude strong duality for the primal and dual problem. Set $\tilde{\rho}_{AB} = \rho_{AB}$, μ such that $\lambda\rho_A \otimes \rho_B - \rho_{AB} - \mu\mathbb{1} \geq 0$, $X = \sqrt{1 - \varepsilon^2}$, implying $\begin{pmatrix} \rho_{AB} & X \\ X^\dagger & \tilde{\rho}_{AB} \end{pmatrix} \geq 0$ and thus being a feasible point of the primal SDP. Let $\alpha > 1$, $\nu > 0$, $\beta > 0$, $\eta = \beta + \nu + (\lambda - 1)\alpha$, $D = (\nu + \beta)\mathbb{1}$ and $W = \alpha\mathbb{1}$. Then, $\text{Tr}[W] > 1$, $W > 0$, $D > 0$ and by (22) $M = \alpha\mathbb{1}$ and $\begin{pmatrix} D & \nu\mathbb{1} \\ \nu\mathbb{1} & \eta\mathbb{1} - (\lambda M - W) \end{pmatrix} = \begin{pmatrix} (\nu + \beta)\mathbb{1} & \nu\mathbb{1} \\ \nu\mathbb{1} & (\nu + \beta)\mathbb{1} \end{pmatrix} > 0$. Consequently, this is a strictly feasible

point and strong duality is implied by Slater's condition (cf. [17]). $\square$

Proposition 3. *If for ρ_{AB} and all $\tilde{\rho}_{AB} \in B^\varepsilon(\rho_{AB})$, $\rho_A \otimes \tilde{\rho}_B$ is positive definite, the procedure above converges to λ_{min} of (9).*

Proof. Let λ^i be as in (10). Assume that $\lambda^i > \lambda_{min}$ as in (9). Then, there exists a $\hat{\rho}_{AB}$ and $\hat{\lambda} = \lambda^i - \hat{\delta}$ with $\hat{\delta} > 0$ and $\hat{\lambda}\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} \geq 0$. This implies $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} - \hat{\delta}\rho_A \otimes \hat{\rho}_B \geq 0$ and consequently, because $\hat{\delta} > 0$ and by assumption $\rho_A \otimes \hat{\rho}_B > 0$, we must have $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} > 0$. This indicates that there exist $\hat{\rho}_{AB} \in B^\varepsilon(\rho_{AB})$ and $\hat{\mu} > 0$ such that $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} - \hat{\mu}\mathbb{1} \geq 0$. This implies that the SDP (11) has a feasible solution (μ^i, ρ_{AB}^{i+1}) and consequently $\lambda^i\rho_A \otimes \rho_B^{i+1} - \rho_{AB}^{i+1} > 0$. Thus, using ρ_{AB}^{i+1} for the next iteration in (10), will return a $\lambda^{i+1} < \lambda^i$. In summary, we have shown that if $\lambda^i > \lambda_{min}$, then there exists a feasible solution μ^i of the SDP (11), provided that all states have full rank. This solution implies a feasible improved solution $\lambda^{i+1} < \lambda^i$ of (10) for the next iteration. Due to the strong duality of the SDP involved (see Proposition 2), if it exists, a feasible solution will be found by the program. Failing to do so in step 3 in Algorithm 1 consequently implies that no feasible $\mu^i > 0$ exists and therefore $\lambda^i = \lambda_{min}$. Proposition 2.3. in [14] implies that this seesaw/mountain-climbing-type algorithm converges to the optimal solution after a finite number of steps. $\square$

Before concluding, we briefly discuss the role of the rank constraint $\rho_A \otimes \hat{\rho}_B > 0$: If this constraint is violated, then $\lambda^i > \lambda_{min}$ does not necessarily imply $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} > 0$. Instead, $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB}$ and $\rho_A \otimes \hat{\rho}_B$ may have nontrivial intersecting kernels, such that $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} - \hat{\delta}\rho_A \otimes \hat{\rho}_B \geq 0$ holds for some $\hat{\delta} > 0$, but $\lambda^i\rho_A \otimes \hat{\rho}_B - \hat{\rho}_{AB} - \mu\mathbb{1} < 0$ for all $\mu > 0$. In this case, the SDP in (11) results in $\mu = 0$ and the algorithm stops without necessarily having found the optimal solution λ_{min} .

4 Conclusion

In this work, we developed an SDP-based iterative algorithm to compute the smooth max-mutual information $I_{\max}^\varepsilon(\rho_{AB})$ of bipartite quantum states. The algorithm was proven to be accurate under the assumption that for all states $\tilde{\rho}_{AB}$ in the smoothing environment around ρ_{AB} , the marginal states $\rho_A \otimes \tilde{\rho}_B$ have full rank.

A key contribution of this work is the characterization of a previously unexplored semidefinite program, which forms a core component of the iterative procedure. We established its primal and dual formulations and proved strong duality, thereby laying a rigorous foundation for its use in quantum information theory. Our results extend the landscape of SDP-based techniques for evaluating operational quantities, complementing existing methods as used for the computation of, e.g., the hypothesis testing mutual information [12, 18], or the conditional smooth min- and max-entropy [13]. Notably, since the smooth max-mutual information can be combined with the hypothesis testing mutual information to lower-bound the one-shot distillable key (cf. Ref. [2]), our algorithm offers a tool for estimating cryptographic capacities of quantum states that satisfy the rank constraints.

If the constraints are not fulfilled, the algorithm may not yield optimal solutions. Future work may focus on relaxing the rank assumptions, which would broaden the applicability of our method to general quantum states. In Ref. [19], it is shown that the smooth max-mutual information of ρ_{AB} is invariant under local unitary transformations U_A acting on A and general local isometric transformations $V_{B' \rightarrow B}$, mapping a smaller system B' to B . If for a state ρ_{AB} the rank condition is violated, but there exists a state $\hat{\rho}_{AB'}$ with $\rho_{AB} = U_A \otimes V_{B' \rightarrow B} \hat{\rho}_{AB'} U_A^\dagger \otimes V_{B' \rightarrow B}^\dagger$ that satisfies the condition, then the smooth max-mutual information can be computed with the presented algorithm for $\hat{\rho}_{AB'}$, instead. One potential way to lift the rank conditions is to analyze for which states such a pre-image under local transformations exists. Moreover, the fully characterized SDP presented in Proposition 2 may find utility beyond this context, potentially serving as a building block for evaluating other information-theoretic quantities or optimization problems in quantum theory and related fields.

References

- ¹M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
- ²S. Khatri and M. M. Wilde, *Principles of Quantum Communication Theory: A Modern Approach*, arXiv:2011.04672 [cond-mat, physics:hep-th, physics:math-ph, physics:quant-ph], Feb. 2024.
- ³F. Buscemi and N. Datta, “The Quantum Capacity of Channels With Arbitrarily Correlated Noise”, *IEEE Transactions on Information Theory* **56**, 1447–1460 (2010).
- ⁴N. Datta, “Min- and Max-Relative Entropies and a New Entanglement Monotone”, *IEEE Transactions on Information Theory* **55**, 2816–2826 (2009).
- ⁵M. Tomamichel, *Quantum Information Processing with Finite Resources – Mathematical Foundations*, Vol. 5, arXiv:1504.00233 [quant-ph] (2016).
- ⁶F. Hiai and D. Petz, “The proper formula for relative entropy and its asymptotics in quantum probability”, en, *Communications in Mathematical Physics* **143**, 99–114 (1991).
- ⁷S. Khatri, E. Kaur, S. Guha, and M. M. Wilde, *Second-order coding rates for key distillation in quantum key distribution*, arXiv:1910.03883 [quant-ph], July 2021.
- ⁸P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*, en (IOP Publishing, Mar. 2023).
- ⁹A. Tavakoli, A. Pozas-Kerstjens, P. Brown, and M. Araújo, “Semidefinite programming relaxations for quantum correlations”, *Reviews of Modern Physics* **96**, Publisher: American Physical Society, 045006 (2024).
- ¹⁰J. Watrous, *Semidefinite programs for completely bounded norms*, arXiv:0901.4709 [quant-ph], Apr. 2009.
- ¹¹J. Watrous, *Simpler semidefinite programs for completely bounded norms*, arXiv:1207.5726 [quant-ph], Aug. 2012.
- ¹²F. Dupuis, L. Krämer, P. Faist, J. M. Renes, and R. Renner, “Generalized entropies”, in *XVIIth International Congress on Mathematical Physics* (WORLD SCIENTIFIC, Dec. 2012), pp. 134–153.
- ¹³T. Nuradha and M. M. Wilde, “Fidelity-Based Smooth Min-Relative Entropy: Properties and Applications”, *IEEE Transactions on Information Theory* **70**, arXiv:2305.05859 [quant-ph], 4170–4196 (2024).
- ¹⁴H. Konno, “A cutting plane algorithm for solving bilinear programs”, en, *Mathematical Programming* **11**, 14–27 (1976).
- ¹⁵A. E. Rastegin, *Sine distance for quantum states*, arXiv:quant-ph/0602112, Feb. 2006.
- ¹⁶A. Uhlmann, “The “transition probability” in the state space of a *-algebra”, *Reports on Mathematical Physics* **9**, 273–279 (1976).
- ¹⁷S. Boyd, L. Vandenberghe, and L. Faybusovich, “Convex Optimization”, *IEEE Transactions on Automatic Control* **51**, 1859–1859 (2006).
- ¹⁸N. Datta, M. Tomamichel, and M. M. Wilde, “On the second-order asymptotics for entanglement-assisted communication”, en, *Quantum Information Processing* **15**, 2569–2591 (2016).
- ¹⁹C. Popp, T. C. Sutter, and B. C. Hiesmayr, *Local Invariance of Divergence-based Quantum Information Measures*, arXiv:2509.04079 [quant-ph], Sept. 2025.

Author Contributions

Conceptualization, C.P.; validation, C.P., T.C.S. and B.C.H.; formal analysis, C.P.; writing—original draft preparation, C.P.; writing—review and editing, C.P., T.C.S. and B.C.H. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [10.55776/P36102]. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

B.10. Bipartite bound entanglement

Bipartite bound entanglement

Beatrix C. Hiesmayr ^{*}, Christopher Popp [†] and Tobias C. Sutter [‡]

*Faculty of Physics, University of Vienna,
Währingerstraße 17, 1090 Vienna, Austria*

**Beatrix.Hiesmayr@univie.ac.at*

†Christopher.Popp@univie.ac.at

‡Tobias.Christoph.Sutter@univie.ac.at

Received 28 April 2025

Accepted 13 May 2025

Published 14 July 2025

Bound entanglement is a special form of quantum entanglement that cannot be used for distillation, i.e. the local transformation of copies of arbitrarily entangled states into a smaller number of approximately maximally entangled states. Implying an inherent irreversibility of quantum resources, this phenomenon highlights the gaps in our current theory of entanglement. This review provides a comprehensive exploration of the key findings on bipartite bound entanglement. We focus on systems of finite dimensions, an area of high relevance for many quantum information processing tasks. We elucidate the properties of bound entanglement and its interconnections with various facets of quantum information theory and quantum information processing. The paper highlights areas where our understanding of bound entangled states, particularly in their detection and characterization, remains underdeveloped. By highlighting the need for further research into this phenomenon and underscoring relevant open questions, this paper invites researchers to unravel its relevance for our understanding of entanglement in Nature and how this resource can most effectively be used for applications in quantum technology.

Keywords: Separability problem; entanglement distillation; bound entanglement; PPT entanglement.

Contents

1	Introduction	3
2	Basics of Quantum States and Operations	6
2.1	Definition of quantum states	6
2.1.1	Bipartite quantum states, separability and entanglement . . .	7
2.2	Quantum operations	8
2.2.1	Positive and completely positive maps	8
2.2.2	Positive operator valued measures	9

*Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

2.2.3	Entanglement distillation	10
2.3	Bell bases, Bell-diagonal states and the magic simplex	13
2.3.1	Standard Bell-diagonal system – The magic simplex	13
2.3.2	General Bell-diagonal systems	16
3	Free Vs. Bound Entanglement	17
3.1	PNCP maps for entanglement detection	17
3.2	PPT criterion	18
3.3	Existence of PPT entangled and bound entangled states	19
4	Detection of Bound Entanglement	21
4.1	Entanglement witnesses	21
4.1.1	Decomposable and nondecomposable entanglement witnesses	22
4.1.2	Mirrored entanglement witnesses	22
4.1.3	Nonlinear entanglement witnesses	24
4.2	Entanglement witnesses based on MUBs and SICs and generalizations	25
4.2.1	A Complete MUB set is not needed to detect bound entanglement	25
4.2.2	On the universality of MUBs for nondecomposable witnesses	28
4.2.3	Entanglement witnesses based on sets of SICs	29
4.3	Criteria based on a correlation tensor	32
4.4	Realignment criterion	32
4.5	Entanglement criteria based on SICs and GSICs	33
4.6	Concurrence and quasi-pure approximation	34
4.7	Range criterion	35
4.8	Reduction criterion and more criteria from PNCP maps	35
5	Characteristics of Bound Entanglement	36
5.1	Rank of bound entangled states	36
5.2	Frequency of bound entangled states	37
5.2.1	General bipartite bound entangled states	38
5.2.2	Bell-diagonal states	39
5.3	Bound entanglement and information processing	41
5.3.1	Quantum teleportation and distillation	42
5.3.2	Dense coding	44
5.3.3	Activation of bound entanglement	45
5.3.4	Pure state conversion assisted by bound entanglement	46
5.3.5	Bound entangled states and quantum channels	47
5.3.6	Quantum key distribution	47
5.4	Bound entanglement and entanglement measures	49
5.4.1	Irreversibility and PPT bound entanglement	52
5.4.2	Additivity of distillable entanglement and NPT bound entanglement	52

5.4.3	Inequalities for entanglement measures	52
5.5	Bound entanglement and quantum metrology	53
5.6	Bound entanglement and Bell inequalities	54
5.7	Bound entanglement and Lorentz invariance	56
6	Experiments Verifying Bound Entanglement	57
6.1	First experiment verifying the existence of bound entangled states for two qutrits	58
7	Construction of Bound Entangled States	59
7.1	Construction from unextendible product bases	60
8	Conclusion and Outlook	61
	Appendix A Minimal Fidelity of Teleportation	66
	Appendix B Protocol for Activation of Bound Entanglement	67
	References	70

1. Introduction

According to Erwin Schrödinger,¹ entanglement is not only “*one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought”. Almost 90 years after this statement was made, we still lack a full physical and mathematical understanding of this nonclassical phenomenon. Today, one of the main challenges of modern quantum information theory lies in determining whether a composite system is separable or entangled, a problem known as the *separability problem*.² The problem is solved only for qubit–qubit and qubit–qutrit systems since, for these dimensions, there exists a necessary and sufficient criterion to distinguish between separability and entanglement. For general higher-dimensional bipartite systems, $d > 2$, the separability problem is known to be NP-hard,³ i.e. no polynomial-time algorithm is known that is capable of solving it deterministically. A curious form of entanglement arises in those higher dimensions, so-called *bound entanglement*. It is extraordinary as it requires maximally entangled quantum states to generate it, yet this process is irreversible. Hence, once a bound entangled state is produced, one can no longer extract the consumed resources in the form of pure entanglement. This is in stark contrast to *free entanglement* for which this extraction procedure, in the following named entanglement distillation, is partly possible.

Let us briefly mention how the idea of labeling this new type of entanglement with the word “bound” came about by quoting from Ref. 4:

As a matter of fact, we have revealed a kind of entanglement that cannot be used for sending reliably quantum information via teleportation. Using an analogy with thermodynamics, we can consider entanglement as a counterpart of energy, and sending of quantum information as a kind of “informational work”. Consequently, we can consider “free entanglement” (E_{free}) which can be distilled, and “bound entanglement” (E_{bound}). In particular, the free entanglement is naturally identified with distillable

entanglement as the latter asks us how many qubits can we reliably teleport via the mixed state. This kind of entanglement can always be converted via distillation protocol to the “active” singlet form.

To complete the analogy, one could consider the asymptotic number of singlets which are needed to produce a given mixed state as “internal entanglement” E_{int} (the counterpart of internal energy). Then the bound entanglement can be quantitatively defined by the following equation:

$$E_{\text{int}} = E_{\text{free}} + E_{\text{bound}}.$$

So the initial association was with thermodynamics (cf. also Refs. 5 and 6) and especially the second law of thermodynamics, connecting bound entanglement to the arrow of time. The latter connection was taken by an identification with the mathematical property PPT (positive under partial transposition, Definition 13) of the discovered bound entangled state. However, this argument has not proved successful since the (partial) transpose has no one-to-one correspondence to the time arrow of thermodynamics. And yet, the PPT property is central to our understanding of bound entangled states because it remains the only sufficient criterion for undistillability known until today.

Since the discovery of an exemplary state by the Horodecki family at the end of the last millennium,^{4,7} many researchers have put effort into the study of bound entanglement. However, a full characterization of the set of bound entangled states or merely their construction is still lacking. On the other hand, it is shown that they can violate a Bell inequality, i.e. exhibit nonclassical correlations. Recent research suggests that, for specific dimensions, the volume of such bound entangled states in Hilbert space is of nonzero measure, and hence it cannot be neglected in general. These results suggest many similarities, like nonlocality, but also distinctive operational differences, like distillability, between bound and free entanglement. Over the past thirty years, various quantum technologies have been devised that require entangled quantum states for their realization. It turns out, however, that for several of these applications, bound entangled states do not provide an additional resource compared to separable ones. However, some applications exist for which bound entanglement can improve the performance when added to available free entangled states. Current research aims to determine how bound entangled states’ nonclassical features can be further employed in specialized quantum information applications. All of this shows that a full understanding of the theory of entanglement is still not achieved, and with that, an answer to the question of how entanglement can be utilized as a resource for advanced quantum technologies such as, e.g. quantum computing, quantum machine learning, quantum cryptography, or quantum communication.

This review provides a comprehensive summary of the current status of knowledge of *finite-dimensional* and *bipartite* bound entanglement, which is of particular interest for a wide range of quantum technologies. The bipartite scenario also offers

the minimal system size for which bound entanglement appears. The main focus is on contemporary research and the many open questions to be answered. To this end, this paper gives a solid foundation that can pave the way for future developments in the theory of bound entanglement.

The review paper is designed so that it can be read in a different order than provided. However, the order is chosen so readers without prior knowledge can follow the arguments. In particular, Sec. 2 introduces the necessary mathematical concepts and notations to define bound entanglement clearly. The notion of entanglement distillation, central to the definition of bound entanglement, is explained in detail. It further defines Bell-diagonal states, a subset of quantum states with high relevance in many entanglement distillation schemes and a comparatively high volume of bound entanglement. Equipped with the basic mathematical tools, Sec. 3 introduces the PPT criterion and defines bound entanglement as entangled states that cannot be distilled. We establish the crucial connection between the PPT property and bound entangled states by showing that all distillable states violate the PPT criterion. Section 4 reviews the main approaches to detect bound entangled states and differentiate them from free entangled and separable states. Multiple methods are discussed, and it is emphasized that there is no strict hierarchy between them, i.e. no entanglement criterion is strictly stronger than any other in terms of bound entangled states it detects. The logical next step following the detection of some bound entangled states is to investigate their characteristics, which is the focus of Sec. 5. Properties of the set of bound entangled states and individual bound entangled states are discussed. Furthermore, the limitations and usefulness of bound entanglement for quantum information processing tasks are summarized. We also introduce various entanglement measures and point out their hierarchical dependence and the resulting consequences. In this regard, the difficulty of finding a good entanglement measure is connected to the existence of bound entangled states. Concluding the characteristics of bound entanglement, we argue that bound entanglement can be Bell-nonlocal and is not Lorentz invariant. In addition to the fundamental difficulties of detecting, characterizing and classifying bound entanglement, Sec. 6 deals with how this curious phenomenon was uncovered in experiments. In particular, we present experimental realizations using entanglement witnesses based on SICs and MUBs, which are formally introduced in Sec. 4.2. Even though we can detect and characterize bound entangled states to some degree, it is still not fully understood how to construct the set of bound entangled states. The dual problem is finding a complete set of nondecomposable entanglement witnesses (cf. Sec. 4.1.1). Some approaches to constructing bound entangled states are presented in Sec. 7, focusing on the analytic construction via unextendible product bases. The paper concludes in Sec. 8 with a summary of fundamental questions that have been answered and a broad outlook on open questions. This includes the problem of why Nature provides us with this special kind of entanglement and for what quantum technologies it can be effectively leveraged. Also, the potential existence of NPT

bound entangled states is discussed. Solving these issues will be challenging as the detection of bound entanglement is intimately linked to the NP-hard separability problem. At the same time, their characterization can be traced back to our incomplete understanding of entanglement in general. Nonetheless, closing these open gaps in our knowledge will certainly lead to a more complete comprehension of one of our fundamental physical theories.

2. Basics of Quantum States and Operations

This section contains the basic definitions of quantum states and how they can be manipulated by positive operator-valued measures and completely positive trace-preserving maps. It thus lays the mathematical foundations needed for understanding bound entanglement. Importantly, in Sec. 2.2.3, the concept of entanglement distillation is introduced, which is at the heart of the definition of bound entanglement.

2.1. Definition of quantum states

The mathematical structure of quantum physics is determined by linear algebra on finite- or infinite-dimensional Hilbert spaces. In the following, we restrict ourselves to Hilbert spaces $\mathcal{H}$ of finite dimension $d > 1$. For clarity, we will sometimes choose $\mathcal{H} = \mathbb{C}^d$ when $\dim(\mathcal{H}) = d$. We denote the set of linear operators mapping $\mathcal{H}$ into itself by $\mathcal{L}(\mathcal{H})$, in detail:

Definition 1 (Quantum state). A linear operator $\rho \in \mathcal{L}(\mathcal{H})$ acting on a Hilbert space $\mathcal{H}$ is called quantum state or density operator if it is positive semi-definite ($\rho \geq 0$) and has unit trace ($\text{Tr}(\rho) = 1$).

The set of all quantum states on $\mathcal{H}$ is denoted $\mathcal{D}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$.

Note that the condition $\rho \geq 0$ implies hermiticity, i.e. $\rho = \rho^\dagger$, with $\rho^\dagger$ being the Hermitian conjugate of ρ . For general Hilbert space dimension $\dim(\mathcal{H}) = d$, the quantum system is also called *qudit*. In the cases $d = 2$, $d = 3$ and $d = 4$, they are known as a qubit, qutrit and ququart, respectively. A quantum state $\rho \in \mathcal{D}(\mathcal{H})$ is called *pure* if $\text{Tr}(\rho^2) = 1$ and *mixed* if $\text{Tr}(\rho^2) < 1$. Due to the spectral theorem (cf. Ref. 8) we can write any $\rho \in \mathcal{D}(\mathcal{H})$ in its eigenbasis as

$$\rho = \sum_{i=0}^{d-1} \lambda_i |\psi_i\rangle\langle\psi_i|, \quad (1)$$

where $|\psi_i\rangle$ are eigenvectors of ρ with respective eigenvalues λ_i , and the set $\{|\psi_i\rangle\}_{i=0}^{d-1}$ of eigenvectors forms an orthonormal basis of $\mathcal{H}$. Consequently, any pure state can be written as $\rho = |\psi\rangle\langle\psi|$, and thus directly associated with the corresponding Hilbert space element $|\psi\rangle \in \mathcal{H}$. We call ρ a rank- r state if it has r nonzero eigenvalues.

2.1.1. Bipartite quantum states, separability and entanglement

Bipartite quantum states are identified with linear operators acting on the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of two Hilbert spaces. Such density matrices can be expanded in the computational product basis as

$$\rho_{AB} = \sum_{i,j,k,l=0}^{d-1} \rho_{ik,jl} |i\rangle\langle j| \otimes |k\rangle\langle l|. \quad (2)$$

In this case, the *partial trace* of only one subsystem is defined by linearity as

$$\text{Tr}_A(\rho_{AB}) := (\text{Tr} \otimes \mathbb{1}_B)(\rho_{AB}) = \sum_{i,k,l=0}^{d-1} \rho_{ik,il} |k\rangle\langle l|, \quad (3)$$

where $\mathbb{1}_B$ denotes the identity map on $\mathcal{L}(\mathcal{H}_B)$.^a This operation results in a state on $\mathcal{D}(\mathcal{H}_B)$ called the *reduced state* with respect to subsystem B . The operation Tr_B is defined analogously.

For bipartite systems, nonclassical features can arise, such as quantum entanglement (an extensive review of this broad topic can be found in Ref. 2). We only reproduce the facts necessary for understanding bound entanglement, formally introduced in Sec. 3.

Definition 2 (Separability and entanglement). A bipartite quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called separable if it can be written as a convex combination of product states, i.e. $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ with probabilities $p_i \geq 0$, $\sum_i p_i = 1$, $\rho_A^i \in \mathcal{D}(\mathcal{H}_A)$ and $\rho_B^i \in \mathcal{D}(\mathcal{H}_B)$. If ρ_{AB} is not separable, it is called inseparable or entangled. We denote the set of all separable (entangled) states by SEP (ENT).

Note that the representation of a separable state as a mixture of product states is generally not unique. Furthermore, the above definition induces a partition of

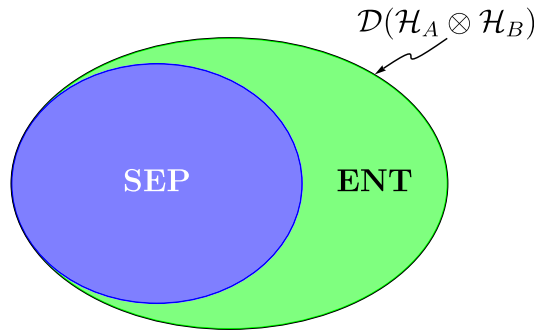


Fig. 1. Partition of bipartite state space $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ into the sets SEP (blue) and ENT (green). The diagram does not faithfully represent the relative volumes of SEP and ENT, which is known to depend on the dimensions of $\mathcal{H}_A$ and $\mathcal{H}_B$.

^aThis means $\mathbb{1}_B$ should be seen as a linear map from $\mathcal{L}(\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_B)$. Below, we also denote the identity matrix by $\mathbb{1}_B \in \mathcal{L}(\mathcal{H}_B)$. The context unambiguously establishes the meaning for each appearance.

the set of density operators $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. This means it splits up the set into two proper disjoint subsets SEP and ENT satisfying $\text{SEP} \neq \emptyset$, $\text{ENT} \neq \emptyset$, $\text{SEP} \cup \text{ENT} \equiv \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\text{SEP} \cap \text{ENT} \equiv \emptyset$. Fig. 1 visualizes this partition.

Definition 3 (Maximally entangled states, locally maximally mixed states). A pure bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$ is said to be maximally entangled if the partial trace on either system yields the maximally mixed state, i.e. $\text{Tr}_A(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_B$ and $\text{Tr}_B(|\psi\rangle\langle\psi|) = d^{-1} \mathbb{1}_A$. Any bipartite quantum state with the latter property is called locally maximally mixed.

This definition is equivalent to $|\psi\rangle$ having full Schmidt rank with Schmidt coefficients $1/\sqrt{d}$. One important maximally entangled state discussed further in Sec. 2.3 is given by

$$|\Omega_{0,0}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d. \quad (4)$$

Determining whether a given quantum state is separable or entangled is a natural problem of interest that can be formalized as follows:

Definition 4 (Separability problem). Given $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, decide whether or not $\rho_{AB} \in \text{SEP}$.

This is a well-posed question because SEP, together with ENT, forms a partition of the state space. Generally, the separability problem is known to be NP-hard.³ This means that there is no known polynomial time algorithm that deterministically decides whether or not $\rho_{AB} \in \text{SEP}$ for all $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and arbitrary Hilbert space dimensions $\dim(\mathcal{H}_A)$ and $\dim(\mathcal{H}_B)$ unless the complexity classes P and NP are equal (they are commonly assumed not to be⁹).

2.2. Quantum operations

Any general operation that modifies a quantum state should naturally result in another quantum state. This motivates the concepts of completely positive trace-preserving maps and positive operator-valued measures.

2.2.1. Positive and completely positive maps

The following definition captures the central object of interest in this section.¹⁰

Definition 5 (Completely Positive (CP) map). A linear map

$$\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_{A'}) \quad (5)$$

is called positive if $\rho_A \geq 0$ implies $\mathcal{E}(\rho_A) \geq 0$ for any $\rho_A \in \mathcal{L}(\mathcal{H}_A)$. Additionally, $\mathcal{E}$ is called completely positive (CP) if $\mathcal{E} \otimes \text{id}_B : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_{A'} \otimes \mathcal{H}_B)$ is a positive map for every Hilbert space $\mathcal{H}_B$. Here, the action of $\mathcal{E} \otimes \text{id}_B$ on a composite system is defined by linearity. If, for a positive map $\mathcal{E}$, there exists a Hilbert space

$\mathcal{H}_B$ such that $\mathcal{E} \otimes \text{id}_B$ is not a positive map, $\mathcal{E}$ is called positive but not completely positive (PNCP).

From a physical point of view, applying a positive map to a quantum state ensures that the measurement probabilities remain nonnegative. Furthermore, complete positivity captures the intuition that the probabilities remain nonnegative even if the map is only applied to part of a larger system. It is clear that every completely positive map is also positive.

Any physically implementable map must be completely positive and trace-preserving (CPTP) to map density operators to density operators in accord with Definition 1. Such maps are also called *quantum channel*. To check whether a given linear map $\mathcal{E}$ is CP, it is sufficient to check the positive semi-definiteness of its Choi operator

$$J(\mathcal{E}) := \sum_{i,j=0}^{d_A-1} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \in \mathcal{L}(\mathcal{H}_{A'} \otimes \mathcal{H}_A). \quad (6)$$

One can also show the reverse statement to find:¹⁰

Theorem 1 (Choi's Theorem). *A linear map $\mathcal{E}$ is CP if and only if $J(\mathcal{E}) \geq 0$.*

Although they cannot be realized experimentally, *positive but not completely positive* (PNCP) maps are valuable in theoretical considerations, in particular with regards to the separability problem (see Sec. 3).

2.2.2. Positive operator valued measures

Positive operator-valued measures (POVMs) generalize projection-valued measures (also known as projective or von Neumann measurements). They describe generalized measurements on quantum states and are defined as follows:¹¹

Definition 6 (Positive operator valued measure). A positive operator-valued measure (POVM) on a Hilbert space $\mathcal{H}$ is a set of n positive semi-definite operators $E_m \in \mathcal{L}(\mathcal{H})$ that sum up to the identity on $L(\mathcal{H})$,

$$\sum_{m=1}^n E_m = \mathbb{1}_{\mathcal{L}(\mathcal{H})}. \quad (7)$$

The individual E_m are called *POVM elements*. Equation (7) establishes that a POVM on $\mathcal{H}$ is a partition of the identity on $\mathcal{L}(\mathcal{H})$. The probability of obtaining outcome m after conducting a POVM measurement on a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ is given by

$$p_m = \text{Tr}(E_m \rho). \quad (8)$$

That the probabilities sum up to one, i.e. $\sum_{m=1}^n p_m = 1$, is ensured by (7).

Projective measurements are special cases of POVMs for which $E_m^2 = E_m$ for all $m = 1, \dots, n$. In this case, each POVM element E_m is a rank-1 projector and can be written as $E_m = |\pi_m\rangle\langle\pi_m|$ for some normalized $|\pi_m\rangle \in \mathcal{H}$. The completeness

relation (7) enforces that the set $\{|\pi_m\rangle\}_m$ constitutes an orthonormal basis of $\mathcal{H}$, i.e. we need to have exactly $n = \dim(\mathcal{H})$ POVM elements.

Alternatively, a POVM can be defined as a set of n linear operators M_m acting on $\mathcal{L}(\mathcal{H})$ and satisfying $\sum_{m=1}^n M_m^\dagger M_m = \mathbb{1}$. These two statements are equivalent as any positive semi-definite operator E_m can be written as

$$E_m = M_m^\dagger M_m, \quad (9)$$

e.g. by writing $M_m = \sqrt{E_m}$ (which is unique up to an isometry). Here, the probability for outcome m is given by

$$p_m = \text{Tr}(M_m^\dagger M_m \rho) = \text{Tr}(M_m \rho M_m^\dagger). \quad (10)$$

The advantage of this alternative definition is that the post-measurement state is given by

$$\rho_m = \frac{1}{p_m} M_m \rho M_m^\dagger. \quad (11)$$

POVMs are crucial for theoretical applications such as the detection of bound entanglement via mutually unbiased bases (MUBs) or symmetric informationally complete POVMs (SICs) (see Sec. 4.2).

2.2.3. Entanglement distillation

The defining feature of bound entanglement is given by the notion of *entanglement distillation*.^{b,12,15} Heuristically speaking, distillation describes the process of two parties, A and B , converting multiple “weakly” entangled shared quantum states into less but “more strongly” entangled states using only local quantum operations and classical communication (LOCC).

More precisely, A and B share n identical copies of some (generally mixed) bipartite resource state ρ ,

$$\rho^{\otimes n} = \rho_{A_1 B_1} \otimes \rho_{A_2 B_2} \otimes \cdots \otimes \rho_{A_n B_n} \in \mathcal{D}\left(\bigotimes_{i=1}^n \mathcal{H}_{A_i} \otimes \mathcal{H}_{B_i}\right), \quad (12)$$

where A (B) is in possession of the subsystems indexed by A_i (B_i), $i \in \{1, 2, \dots, n\}$. Here, we only consider $\dim(\mathcal{H}_{A_i}) = \dim(\mathcal{H}_{B_i}) = d$ for all i . They are allowed to perform *local* quantum operations, i.e. A can implement any CPTP map $\Phi : \mathcal{L}(\bigotimes_i \mathcal{H}_{A_i}) \rightarrow \mathcal{L}(\bigotimes_i \mathcal{H}_{A_i})$ or conduct any POVM measurement on $\bigotimes_i \mathcal{H}_{A_i}$, and similar for B . Furthermore, they can communicate *classically*, e.g. send an obtained measurement outcome to the other party, who may adjust their subsequent quantum operations accordingly. A general distillation protocol can consist of any

^bEntanglement distillation is also sometimes called entanglement purification.^{12–14} This should not be confused with the purification of a mixed quantum state $\rho \in \mathcal{D}(\mathcal{H}_A)$ by a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ such that $\rho = \text{Tr}_B(|\psi\rangle\langle\psi|)$. We use the term entanglement distillation, or simply distillation, throughout.

number of consecutive LOCC operations. *Not* allowed are joint quantum operations in the form of CPTP maps or POVMs involving A 's and B 's qudits together or any exchange of qudits. These operations are considered global and require resources unavailable in a LOCC setting.

Given n copies of a resource state ρ , the challenge in any practical approach is to develop an optimal strategy for CPTP maps and POVM measurements. The goal of distillation is to produce multiple copies of a maximally entangled *target state*, e.g. the maximally entangled qubit state $|\Phi^+\rangle = (1/\sqrt{2}) \sum_{i=0}^1 |i\rangle \otimes |i\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. We define entanglement distillation as any LOCC protocol that transforms multiple copies of a resource state into (generally fewer) copies of this target state:

Definition 7 (Entanglement distillation, distillability, distillation rate).

Entanglement distillation describes any LOCC protocol that transforms n copies of a resource state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ to m copies of the maximally entangled qubit state $|\Phi^+\rangle\langle\Phi^+| \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with nonzero probability:

$$\rho^{\otimes n} \xrightarrow{\text{LOCC}} |\Phi^+\rangle\langle\Phi^+|^{\otimes m}. \quad (13)$$

A state ρ is distillable if in the asymptotic limit $n \rightarrow \infty$, a distillation rate $r := \frac{m}{n}$ larger than zero can be achieved via entanglement distillation. If no such operations exist, ρ is called undistillable.

Other protocols producing different target states are possible and interesting for practical purposes and theoretical analyses. However, this is not a limitation of the notion of distillability. Analyzing distillability with respect to a different target state is fully equivalent to the given definition if it is itself distillable. In this case, multiple copies of this target state can be transformed via LOCC to some number of maximally entangled qubit states. In this case, the distillation rate generally changes. Still, any state for which a positive distillation rate with respect to such a target state can be achieved (in the asymptotic limit) will also be distillable according to Definition 7.

In particular, the d -dimensional maximally entangled state $|\psi\rangle = |\Omega_{0,0}\rangle = (1/\sqrt{d}) \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ (cf. Refs. 13, 16 and 17) provides straightforward applications in quantum information processing tasks in d -dimensional state space (see, e.g. Chap. 6 in Ref. 17) and therefore is a natural target state. Indeed, this target state can be equivalently used as the target state for distillation. This can be shown by using the concept of majorization on the ordered vectors consisting of the squared Schmidt coefficients of the two states (the interested reader may consult Sec. 12.5 of Ref. 8 for details). Using this method, one finds that $|\Phi^+\rangle$ can be obtained by LOCC from a single copy of $|\Omega_{0,0}\rangle$. Hence, $|\Omega_{0,0}\rangle$ is equally suited as a target state when we are only interested in the (asymptotic) distillability of the resource state. The only quantitative difference when considering $|\Omega_{0,0}\rangle$ instead of $|\Phi^+\rangle$ is the change in the distillation rate $m/n \rightarrow m/(n \lceil \log_2(d) \rceil)$. To see this, note that one can distill $|\Omega_{0,0}\rangle$ using $\lceil \log_2(d) \rceil$ copies of $|\Phi^+\rangle$ (here $\lceil \cdot \rceil$ denotes the ceiling function).

It is clear from the definition that only entangled states are potentially distillable because LOCC operations cannot convert separable states to entangled ones.¹⁶ In light of Sec. 3, this can also be considered a corollary of Theorem 5 together with Theorem 7. A formal statement about the quantification of entanglement of a quantum state via distillation can be found in Definition 25, introducing distillable entanglement as an entanglement measure.

Depending on whether the resource state ρ is pure or mixed, there are considerable differences in the methods used to determine its distillability. For pure states, the Schmidt coefficients completely characterize the state's distillation properties, both in the finite and in the asymptotic setting.⁸ However, for general mixed states, the issue is more complicated and an active area of research. It is shown in Ref. 16 that if the overlap of the resource state ρ with the maximally entangled state $|\Omega_{0,0}\rangle$ is greater than $1/d$, the state is distillable. In the case of mixed qubit resource states ($d = 2$), the following theorem holds:¹⁸

Theorem 2. *A bipartite qubit state $\rho \in \mathcal{D}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is distillable if and only if $\rho \in \text{ENT}$.*

This no longer holds for qudits of Hilbert space dimension $d > 2$, meaning undistillable entangled states exist. These are termed bound entangled⁴ and are treated more formally in Sec. 3. The theorem implies that any entangled qubit state can be used as a target state, as it can be transformed to maximally entangled qubits with some nonzero distillation rate. Using such a target state can significantly simplify analyses of distillability for a given state and motivates the following definition.

Definition 8 (n -distillability). A bipartite state ρ is called n -distillable if n copies of ρ can be transformed via LOCC into a bipartite entangled pair of qubits. Otherwise, it is called n -undistillable.

Note that no restriction is placed upon the amount of entanglement of the pair of qubits. Clearly, any n -distillable state is also distillable. However, the following Theorem of Ref. 19 shows that establishing the undistillability of a quantum state via n -undistillability is difficult. The reason is that one needs to check all possible numbers of copies n , as there exist states that are n_0 -undistillable for some large n_0 , but are $(n_0 + 1)$ -distillable.

Theorem 3. *For any choice of integers $d \geq 3$ and $n \geq 1$, there exists a $d^2 \times d^2$ -dimensional bipartite mixed quantum state that is distillable but not n -distillable.*

Consequently, to show the undistillability of an entangled bipartite state, it is generally not sufficient to show n -undistillability for some finite n . This is one manifestation of the fact that bound entangled states are generally hard to detect and characterize (see Secs. 4 and 5).

2.3. Bell bases, Bell-diagonal states and the magic simplex

This section describes the properties of special quantum states called Bell states that are highly relevant to quantum information processing tasks.

Consider a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$ composed of two subsystems $\mathcal{H}_A$ and $\mathcal{H}_B$ and corresponding quantum states $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Let $d_A(d_B)$ be the dimension of the subsystem $\mathcal{H}_A(\mathcal{H}_B)$ and $d_{AB} = d_A d_B$ the dimension of the product system. If not explicitly specified otherwise, we assume $d_A = d_B = d$. Pure, maximally entangled states on $\mathcal{H}_A \otimes \mathcal{H}_B$ are called *Bell states*. Their relevance stems from the fact that Bell states contain a maximum amount of entanglement that can be leveraged to process quantum information. Per definition, these states are locally maximally mixed (cf. Definition 3) as the reduced state for any subsystem is the maximally mixed state. Hence, all information lies in the correlation between the two systems. A *Bell basis* consists of d_{AB} linear independent Bell states that span $\mathcal{H}_A \otimes \mathcal{H}_B$. These bases are frequently used in applications like quantum teleportation,²⁰ dense coding,²¹ or error correction.²² *Bell-diagonal states* are mixtures of Bell basis states and arise naturally when errors due to imperfections in the experimental setup or other unintentional interactions are considered. The entanglement structure of Bell basis states strongly influences the properties of corresponding Bell-diagonal states. Therefore, understanding the entanglement structure of Bell states is crucial for effectively applying entanglement as a resource.

A frequently used “standard” Bell basis generalizing the Pauli-basis of bipartite qubits to higher dimensions shows special algebraic and geometric properties (see Sec. 2.3.1). Due to these properties, the corresponding set of Bell-diagonal states is called “Magic Simplex” and its special entanglement structure, including a high relative share of bound entanglement, can be analyzed with unique approaches (see Sec. 5). Alternatively, more general Bell-diagonal systems are defined (see Sec. 2.3.2) and compared with a focus on the frequency of bound entanglement in Sec. 5.2.2.

2.3.1. Standard Bell-diagonal system – The magic simplex

A special Bell basis can be generated by the *Weyl–Heisenberg* operators.²³ This generalization of the Pauli operators to general dimension d is defined as follows:

Definition 9 (Weyl–Heisenberg operators).

$$W_{k,l} := \sum_{j=0}^{d-1} \omega^{jk} |j\rangle \langle j+l|, \quad k, l = 0, \dots, d-1, \quad (14)$$

where here and in the following, $\omega := \exp(\frac{2\pi i}{d})$, and indices and variables taking the values $0, 1, \dots, d-1$ are elements of the ring $\mathbb{Z}_d := \mathbb{Z}/d\mathbb{Z}$ with addition and multiplication modulo d . From $W_{k,l} = Z(k)X(l)$ with $Z(k) = \sum_j \omega^{jk} |j\rangle \langle j|$ and $X(l) = \sum_j |j-l\rangle \langle j|$ it is clear that $W_{k,l}$ acts as phase and shift operations depending on the phase (k) and shift (l) index. These unitary, but generally not Hermitian operators obey the *Weyl-relations* (with $*$ and T denoting complex conjugation and

transposition with respect to the computational basis, respectively):

$$\begin{aligned} W_{k_1, l_1} W_{k_2, l_2} &= \omega^{l_1 k_2} W_{k_1 + k_2, l_1 + l_2}, \\ W_{k, l}^\dagger &= \omega^{kl} W_{-k, -l} = W_{k, l}^{-1}, \\ W_{k, l}^* &= W_{-k, l}, \\ W_{k, l}^T &= \omega^{-kl} W_{k, -l}. \end{aligned} \quad (15)$$

The algebraic relations imply that the elements $\omega^m W_{k, l}$ ($m, k, l \in \mathbb{Z}_d$) form a finite discrete group under multiplication, named *Weyl–Heisenberg group*.

A d^2 -dimensional Bell basis $\{|\Omega_{k, l}\rangle\}$ can be constructed by local application of the Weyl–Heisenberg operators to a specified Bell state, for which we choose $|\Omega_{0, 0}\rangle$:

Definition 10 (Bell states).

$$\begin{aligned} |\Omega_{k, l}\rangle &:= (W_{k, l} \otimes \mathbb{1}_d) |\Omega_{0, 0}\rangle, \quad k, l \in \mathbb{Z}_d, \\ P_{k, l} &:= |\Omega_{k, l}\rangle \langle \Omega_{k, l}|. \end{aligned} \quad (16)$$

The corresponding density matrices $P_{k, l}$, called *Bell projectors*, have the reduced states $\text{Tr}_A(P_{k, l}) = \text{Tr}_B(P_{k, l}) = \frac{1}{d} \mathbb{1}_d$, so $|\Omega_{k, l}\rangle$ are indeed maximally entangled. This basis represents a direct generalization of the qubit Bell basis to higher dimensions. In this work, it is referred to as *standard Bell basis* to differentiate this specific but relevant basis from other, generally unitarily inequivalent, Bell bases that exist for $d \geq 3$ (see Sec. 2.3.2). This Bell basis inherits a linear structure from the Weyl–relations (15), as any element $P_{k, l}$ corresponds to one Weyl–Heisenberg operator $W_{k, l}$. The operations $P_{k, l} \rightarrow (W_{i, j} \otimes \mathbb{1}_d) P_{k, l} (W_{i, j} \otimes \mathbb{1}_d)^\dagger = P_{k+i, l+j}$ ($i, j \in \mathbb{Z}_d$) map the set of Bell projectors onto itself. This structure allows associating states $P_{k, l}$ with elements of a discrete phase space $\mathbb{Z}_d^2$ and operations of the above form with translations in that phase space (see Fig. 2).

Mixtures of standard Bell states with mixing probabilities $\{c_{k, l}\}$ form the set of *standard Bell-diagonal states*, the so-called *Magic Simplex* $\mathcal{M}_d$:

Definition 11 (Magic simplex).

$$\mathcal{M}_d := \left\{ \rho = \sum_{k, l=0}^{d-1} c_{k, l} P_{k, l} \left| \sum_{k, l=0}^{d-1} c_{k, l} = 1, c_{k, l} \geq 0 \right. \right\}. \quad (17)$$

Following the choice of words for the qubit Bell basis of Wootters and Hill,²⁴ “magic” reflects in the very special geometric and algebraic properties inherited from the Weyl–Heisenberg operators (14) and their relations (15). These properties have been used to characterize the entanglement structure of the Magic Simplex and to detect a significant amount of bound entangled states both analytically and numerically.^{23, 25–28} Identifying a state in $\mathcal{M}_d$ with a point in $\mathbb{R}^{d^2}$ given by its mixing probabilities $c_{k, l}$, the set of Bell-diagonal states form a $(d^2 - 1)$ -dimensional simplex. The geometry of this simplex strongly reflects its entanglement structure. Two

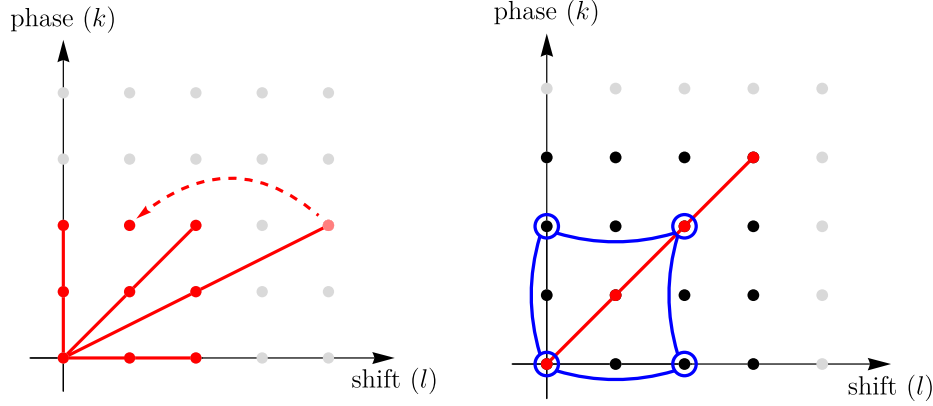


Fig. 2. Visualization of the discrete phase space $\mathbb{Z}_d^2$ for $d = 3$ (left) and $d = 4$ (right). Each vertex represents a Bell state $P_{k,l}$. Connected vertices correspond to Bell states $P_{k,l}$ in some exemplary d -element subgroups. For prime d , subgroups form so-called “lines” (red), while for composite d , other subgroups exist (blue).

relevant subsets with direct relations to the entanglement properties of contained states are the *enclosure polytope* ($\mathcal{E}_d$) and the *kernel polytope* ($\mathcal{K}_d$). $\mathcal{E}_d$ contains all states with positive partial transposition (PPT; see Definition 13)²³ and other states, while $\mathcal{K}_d$ contains only, but not all, separable states. The former is defined as

$$\mathcal{E}_d := \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l} \left| \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \in \left[0, \frac{1}{d}\right] \right. \right\}. \quad (18)$$

The kernel polytope $\mathcal{K}_d$ relates to mixed states $\rho_{S_d} \in \mathcal{M}_d$, which are based on d -element subgroups of the Weyl–Heisenberg group.²³ More precisely, let S_d be a subgroup of $\mathbb{Z}_d^2$ with d elements. The *subgroup state*

$$\rho_{S_d} := \frac{1}{d} \sum_{(k,l) \in S_d} P_{k,l} \quad (19)$$

is a separable state.²³ If d is prime, all such subgroups are generated by one generator each. In this case, the Bell states corresponding to the elements of that subgroup form “lines” in discrete phase space (see Fig. 2), and the subgroup state is also called *line state*. $\mathcal{K}_d$ is defined as the set of convex combinations of all d -element subgroup states:

$$\mathcal{K}_d := \left\{ \rho = \sum_{S_d} q_{S_d} \rho_{S_d} \left| \sum_{S_d} q_{S_d} = 1, q_{S_d} \geq 0 \right. \right\}, \quad (20)$$

where the sum is over all d -element subgroups S_d . Since the states ρ_{S_d} are separable, all states in the kernel polytope are separable by definition. Compare Fig. 3 for an illustration of $\mathcal{K}_d$ for $d = 2$.

The Weyl relations (15) further define symmetry transformations for states in $\mathcal{M}_d$. These linear transformations form a group and act as permutations of the Bell

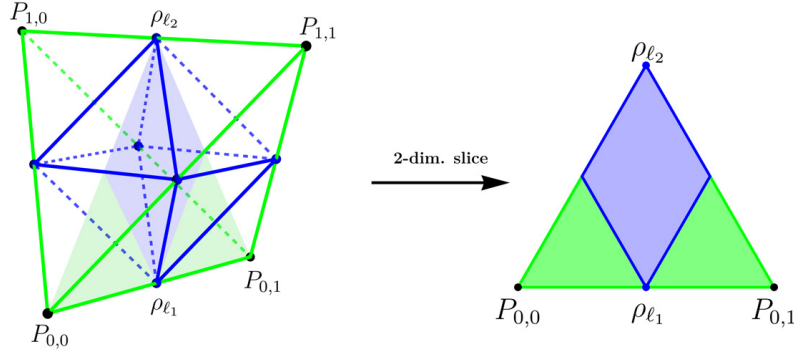


Fig. 3. Visualization of the 3-dimensional Magic Simplex corresponding to $\mathcal{M}_2$ and a 2-dimensional slice thereof. The convex hull of all four maximally entangled states $P_{k,l}$ (black vertices) is a mathematical simplex in the Euclidean space with dimension $d = 3$. The convex hull of the blue vertices representing “line states” form the kernel polytope $\mathcal{K}_2$ and is equivalent to the set of separable states for $d = 2$, i.e. $\text{SEP} \cap \mathcal{M}_2$. Points outside the blue double pyramid represent entangled states.

basis. They are known to preserve entanglement and the PPT/NPT property (see Definition 13 and Ref. 23). Consequently, they conserve the *entanglement class*, i.e. they map the sets of separable, PPT entangled, and NPT entangled states onto themselves.²⁷ The symmetry group is generated by the following generators, which can be defined by their action on the Bell basis projectors:

- $t_{p,q} : P_{k,l} \mapsto P_{k+p,l+q}$ ($p, q \in \mathbb{Z}_d$) (“translation”).
- $m : P_{k,l} \mapsto P_{-k,l}$ (“momentum inversion”).
- $r : P_{k,l} \mapsto P_{l,-k}$ (“quarter rotation”).
- $v : P_{k,l} \mapsto P_{k+l,l}$ (“shear”).

These symmetries are highly relevant for the entanglement structure and the detection of bound entanglement in $\mathcal{M}_d$ (see Sec. 4).

2.3.2. General Bell-diagonal systems

For $d \geq 3$, there exist other Bell bases of the Hilbert space that are not unitarily equivalent to the basis of standard Bell states (16). In Ref. 29, it is shown that the entanglement structure of Bell-diagonal states strongly depends on the particular choice of Bell basis.

A family of *generalized Bell bases* can be constructed as follows: Define the complex matrix $(\alpha_{s,t})_{s,t \in \mathbb{Z}_d}$ with $|\alpha_{s,t}| = 1 \forall s, t \in \mathbb{Z}_d$ and corresponding *generalized Weyl–Heisenberg operators*

$$V_{k,l}^\alpha := \sum_j w^{jk} \alpha_{j+l,l} |j\rangle\langle j+l|, \quad k, l \in \mathbb{Z}_d. \quad (21)$$

Note that these operators are unitary, but generally do not satisfy relations of the form (15) and consequently do not exhibit similar group properties. Nonetheless,

by local application of those operators to the maximally entangled state $|\Omega_{0,0}\rangle$, a family of generalized orthonormal Bell states forming a basis can be obtained:

Definition 12 (Generalized Bell states).

$$\begin{aligned} |\Phi_{k,l}^\alpha\rangle &:= (V_{k,l}^\alpha \otimes \mathbb{1}_d) |\Omega_{0,0}\rangle, \quad k, l \in \mathbb{Z}_d, \\ P_{k,l}^\alpha &:= |\Phi_{k,l}^\alpha\rangle \langle \Phi_{k,l}^\alpha|. \end{aligned} \quad (22)$$

For $\alpha_{s,t} = 1 \forall s, t \in \mathbb{Z}_d$ the generalized Bell states are equal to the standard Bell states (16). In general, however, the (generalized) Bell states and the corresponding bases are not unitarily equivalent. The set of *generalized Bell-diagonal states* is defined as mixtures of generalized Bell basis states:

$$\mathcal{M}_d^\alpha := \left\{ \rho = \sum_{k,l=0}^{d-1} c_{k,l} P_{k,l}^\alpha \left| \sum_{k,l=0}^{d-1} c_{k,l} = 1, c_{k,l} \geq 0 \right. \right\}. \quad (23)$$

Due to the lack of a group structure of the generalized Weyl–Heisenberg operators $V_{k,l}^\alpha$, the generalized Bell-diagonal states do not show equally “magical” properties as the Magic Simplex (Eq. (17)). The algebraic properties of the generalized Bell states and the geometry of the corresponding simplex are not equally strongly related to the entanglement structure of its Bell-diagonal states.²⁹

3. Free Vs. Bound Entanglement

This section discusses the fact that there are two disjoint subsets of entangled states with genuinely different properties. To derive this fact, we revisit the concepts of Sec. 2.2.1. There, it is argued that only completely positive (CP) maps are physically realizable. Hence, for experimental purposes, only the CP maps can be implemented. However, it can nonetheless be beneficial to consider positive but not completely positive (PNCP) maps for theoretical considerations. This section looks into those maps in Sec. 3.1 and presents an important example, the transposition map, in Sec. 3.2. The resulting PPT criterion for separability is closely related to entanglement distillability and is a prime tool for detecting undistillable states. A surprising twist is the emergence of entangled yet undistillable states for which no examples were known for a long time. These so-called “bound” entangled states are formally introduced in Sec. 3.3.

3.1. PNCP maps for entanglement detection

From a theoretical point of view, positive but not completely positive (PNCP) maps (cf. Definition 5) are useful as they give an alternative perspective on the separability of quantum states. The starting point of this consideration is the convexity of the set of separable states SEP. This is easily shown by taking the convex combination of two arbitrary separable states and observing that the result again satisfies the separability condition of Definition 2. By Minkowski’s theorem, any convex set

is fully characterized by the convex hull of its extreme (boundary) points.¹¹ As any boundary point of a convex set can be viewed as the intersection of the set itself with a tangent hyperplane, one may use the set of all tangent hyperplanes of SEP for its complete characterization. In light of Sec. 4.1, SEP is thus characterized by a complete set of optimal entanglement witnesses. Due to the Choi–Jamiołkowski isomorphism,^{30,31} there is a correspondence between entanglement witnesses (viewed as linear operators) and positive maps. This can be utilized to formulate the following characterization of separability:³²

Theorem 4. *A state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable if and only if $(\mathcal{E} \otimes \mathbb{1}_B)(\rho) \geq 0$ for all positive maps $\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_{A'})$.*

Let us shortly outline the practical usage of this theorem for entanglement detection. The contrapositive of the above statement reads: A state ρ is entangled if and only if there exists a positive map $\mathcal{E}$ such that $(\mathcal{E} \otimes \mathbb{1}_B)(\rho) \not\geq 0$. By definition, we have $(\mathcal{E}_{CP} \otimes \mathbb{1}_B)(\rho) \geq 0$ for all CP maps $\mathcal{E}_{CP}$, meaning the search for such a map can be restricted to the smaller set of PNCP maps. By Theorem 1, the Choi operator of a PNCP map satisfies $J(\mathcal{E}_{PNCP}) \not\geq 0$, while for a CP map we have $J(\mathcal{E}_{CP}) \geq 0$. Thus, given a positive map, checking whether it is CP or PNCP is straightforward. Next, to detect the entanglement of a given quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ it is sufficient to show that for a PNCP map $\mathcal{E}_{PNCP} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_{A'})$ at least one eigenvalue of $(\mathcal{E}_{PNCP} \otimes \mathbb{1}_B)(\rho)$ is negative. The nontrivial aspect of this procedure lies in finding a suitable PNCP map. Ideally, this map allows the detection of a large subset of ENT. This sketches one basic strategy of entanglement detection via PNCP maps. It is straightforwardly applied in the next section and with minor adjustments, also for further entanglement detection criteria (see Sec. 4.8).

3.2. PPT criterion

The prime example of a PNCP map is the transposition map $T : L(\mathcal{H}) \rightarrow L(\mathcal{H})$ with respect to a specific basis, acting as $\rho \mapsto \rho^T$. The map is positive since the spectrum of a matrix is invariant under transposition. Furthermore, it is easy to check that the Choi operator $J(T) \not\geq 0$ (cf. Theorem 1). Hence, T is indeed PNCP. The *partial transposition* of only one part of a larger system is denoted by $\Gamma : L(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow L(\mathcal{H}_A \otimes \mathcal{H}_B)$ and its action on a bipartite quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined as

$$\rho^\Gamma := (T \otimes \mathbb{1}_B)(\rho). \quad (24)$$

Here, we arbitrarily choose the transposition map to act on the first subsystem rather than the second. For the spectrum of ρ^Γ this choice is indifferent.^c

^cTo see this, we define the alternative choice as $\rho^{TB} := (\mathbb{1}_A \otimes T)(\rho)$. From $(\rho^T)^T = \rho$ we get $(\rho^{TB})^T = \rho^\Gamma$, and thus $\rho^{TB} \geq 0$ if and only if $\rho^\Gamma \geq 0$ because the eigenvalues are invariant under transposition.

Definition 13 (PPT property). A bipartite quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ satisfying $\rho^\Gamma \geq 0$ is said to be positive under partial transposition (denoted $\rho \in \text{PPT}$), else nonpositive under partial transposition ($\rho \in \text{NPT}$).

By definition, the subsets PPT and NPT form a partition of the space of bipartite quantum states. A direct consequence of Theorem 4 is the *PPT criterion* or *Peres-Horodecki criterion*.^{32,33}

Theorem 5 (PPT criterion). For a bipartite quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ it holds that $\rho \in \text{SEP} \Rightarrow \rho \in \text{PPT}$, or equivalently $\rho \in \text{NPT} \Rightarrow \rho \in \text{ENT}$.

A quantum state $\rho \in \text{PPT}$ is said to satisfy the PPT criterion. Hence, a violation thereof indicates entanglement. Furthermore, for small Hilbert space dimensions, a stronger statement can be derived.³²

Theorem 6. For qubit-qubit ($d_A = d_B = 2$) and qubit-qutrit ($d_A = 2, d_B = 3$) systems, the PPT criterion is necessary and sufficient for separability, i.e. $\rho \in \text{PPT} \Leftrightarrow \rho \in \text{SEP}$.

As elaborated in the upcoming section, this no longer holds for composite systems with greater Hilbert space dimensions.

3.3. Existence of PPT entangled and bound entangled states

The violation of the PPT criterion is only sufficient but not necessary for entanglement in higher-dimensional systems. This indicates the possibility of different entanglement classes characterized by their properties regarding certain quantum operations (cf. Sec. 5.4).

Note that the PPT criterion is only one among many criteria to detect, in general, different subsets of entangled states (see Sec. 4). Nowadays, several PPT entangled states, including qutrit-qutrit ($d_A = d_B = 3$) and qubit-quqart ($d_A = 2, d_B = 4$) systems, are known (cf. Ref. 7). Consequently, Theorem 6 cannot be extended to higher dimensions.

Entanglement distillation (cf. Sec. 2.2.3) is one quantum information theoretic task that is strongly affected by the PPT property. Considering general LOCC protocols, one can show the following.⁴

Theorem 7. If a bipartite quantum state ρ is distillable, then $\rho \in \text{NPT}$. Equivalently, if $\rho \in \text{PPT}$, then ρ is undistillable.

Proof. Consider the d -dimensional Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B \simeq \mathbb{C}^d$. A bipartite quantum state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is distillable if and only if, for some $n \geq 1$, one can transform by LOCC n copies of ρ into the maximally entangled qubit state $|\Phi^+\rangle = (1/\sqrt{2}) \sum_{i=0}^1 |i\rangle \otimes |i\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ embedded in $\mathbb{C}^d \otimes \mathbb{C}^d$ (cf. Definition 7). The target state $|\Phi^+\rangle$ is NPT (i.e. $|\Phi^+\rangle\langle\Phi^+|^\Gamma \not\geq 0$, where Γ denotes the partial transposition on $\mathcal{H}_B$).

The crucial observation in the proof is that LOCC transformations preserve the PPT property of a state. To see this, consider the most general LOCC transformation $\mathcal{E}_{\text{LOCC}}$ of n copies of the state ρ given by³⁴

$$\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n}) = \frac{1}{N} \sum_i (A_i \otimes B_i) \rho^{\otimes n} (A_i^\dagger \otimes B_i^\dagger), \quad (25)$$

where $N = \text{Tr}(\sum_i (A_i \otimes B_i) \rho^{\otimes n} (A_i^\dagger \otimes B_i^\dagger))$ is a normalization factor, and $A_i, B_i \in \mathcal{L}(\mathbb{C}^{d^n})$. If ρ is distillable, then $\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n}) = |\Phi^+\rangle\langle\Phi^+|$ for some LOCC channel. Next, for any local operators $K, L, M, N \in \mathcal{L}(\mathbb{C}^d)$, one can verify $((K \otimes L)\rho(M \otimes N))^\Gamma = (K \otimes L^T)\rho^\Gamma(M \otimes N^T)$, which, by linearity, extends to the general LOCC transformation given in (25) to yield

$$\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n})^\Gamma = \frac{1}{N} \sum_i (A_i \otimes B_i^T) (\rho^\Gamma)^{\otimes n} (A_i^\dagger \otimes (B_i^T)^\dagger), \quad (26)$$

defining again a CP map. Thus, $\rho^\Gamma \geq 0$ implies $\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n})^\Gamma \geq 0$. Consequently, $\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n}) \neq |\Phi^+\rangle\langle\Phi^+|$ for any $\rho \in \text{PPT}$ which concludes the proof. $\square$

The theorem implies that also entangled states can be undistillable, particularly if they are PPT entangled. This leads to the following definition:

Definition 14 (Free and bound entanglement). A bipartite quantum state ρ is called free entangled (denoted $\rho \in \text{FE}$) if it is distillable. If ρ is entangled and undistillable, it is called bound entangled ($\rho \in \text{BE}$).

The term “bound entanglement” originates from the fact that the entanglement is bound to the state in some sense. More specifically, it is inaccessible for

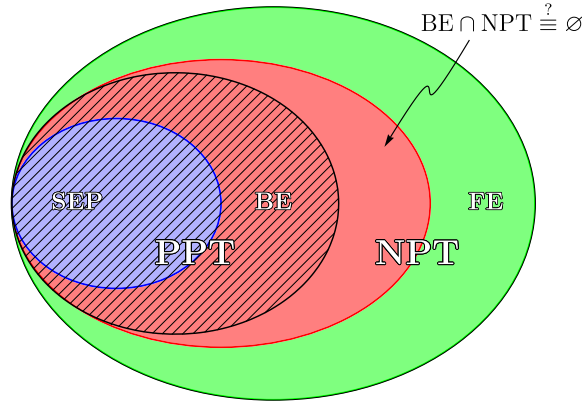


Fig. 4. Set relations of the sets SEP (blue), ENT (green and red), PPT (hatched), NPT (non-hatched), FE (green) and BE (red). We have $\text{FE} \subset \text{NPT}$ or equivalently $\text{PPT} \subset \text{SEP} \cup \text{BE}$. The question is whether it holds that $\text{PPT} \stackrel{?}{=} \text{SEP} \cup \text{BE}$. To settle the case, one would need to prove or disprove $\text{BE} \stackrel{?}{\subset} \text{PPT}$. Disproving this could be accomplished by finding a specific NPT bound entangled state, which, up to today, has not been successful. See Sec. 8 for a discussion of this problem.

various quantum information protocols that involve entanglement distillation (cf. Sec. 2.2.3). This can be formalized using entanglement measures where the bound nature is reflected in the difference between two measures with related operational interpretations (cf. Sec. 5.4).

It is important to emphasize that the sets of PPT entangled and bound entangled states are not necessarily equivalent. Even though PPT entanglement implies bound entanglement by Theorem 7, the converse might not be true. This becomes more transparent when viewing Theorem 7 in set-theoretic language, illustrated in Fig. 4. Up to today, no NPT bound entangled state has been found. Therefore, the question whether $\text{PPT} \stackrel{?}{=} \text{SEP} \cup \text{BE}$ is still open.

4. Detection of Bound Entanglement

In this section, we describe mathematical methods to verify whether a particular state is separable, bound entangled, or free entangled. The first two sections introduce the powerful toolbox of entanglement witnesses, whereas the rest of the section focuses on further entanglement criteria. To verify that a quantum state is bound entangled, one needs to prove that the state is entangled and undistillable. Up to date, the only known strategy to show undistillability is given by the PPT criterion (cf. Theorem 5). Therefore, other entanglement criteria must be combined with the PPT criterion to classify a state as PPT bound entangled. Note that the reduction criterion presented in Sec. 4.8 cannot be used in this way because it is weaker than the PPT criterion. However, it may allow the distinction between potential NPT bound entangled and NPT free entangled states.

4.1. Entanglement witnesses

The concept of entanglement witnesses is of fundamental importance for characterizing the separability problem (Definition 4). Optimal entanglement witnesses characterize the convex set of separable states, and decomposable/nondecomposable entanglement witnesses classify the sets of NPT and PPT entangled states, respectively. Moreover, entanglement witnesses are also a powerful tool to verify entanglement in experiments since they are Hermitian operators and therefore, in principle, measurable observables. In Sec. 6, we present more details on realizations of those observables for detecting entanglement.

Definition 15 (Entanglement witness). A Hermitian operator W is called an entanglement witness (EW) if

$$\begin{aligned} 0 &\leq \text{Tr}(W\sigma) \quad \forall \sigma \in \text{SEP} \quad \text{and} \\ 0 &> \text{Tr}(W\rho) \quad \text{for some } \rho \in \text{ENT}, \end{aligned} \tag{27}$$

where $\sigma, \rho \in \mathcal{D}(\mathcal{H})$ and $W \in \mathcal{L}(\mathcal{H})$. In this case, we say that W witnesses/detects the entanglement of ρ . Furthermore, an entanglement witness is called optimal if there exists some $\sigma \in \text{SEP}$ such that $\text{Tr}(W\sigma) = 0$.

Any entangled state is detectable by an entanglement witness, i.e. given a particular entangled state there always exists a Hermitian operator W with the above properties. The knowledge of all optimal entanglement witnesses $\tilde{W}$ fully characterizes the set SEP. This means that the separability problem is dual to the knowledge of the set of all optimal entanglement witnesses. The proven NP-hardness of the separability problem is reflected in the equally hard problem of finding a complete set of optimal entanglement witnesses.

Since an entanglement witness is a Hermitian operator, it can be realized experimentally in order to verify, i.e. detect, the entanglement of a given state. This also implies a significant advantage from an experimental point of view because, generally, significantly fewer measurement settings are necessary to detect entanglement of an unknown state than via state tomography. More details can be found in Sec. 6.

In the following sections, we review important properties of entanglement witnesses such as decomposability, mirrored entanglement witnesses and nonlinear witnesses, with a focus on detecting and characterizing bound entangled states.

4.1.1. Decomposable and nondecomposable entanglement witnesses

Definition 16 (Decomposable entanglement witness). W is called a decomposable entanglement witness if it can be written as $W = A + B^\Gamma$ with $A, B \geq 0$. Here, B^Γ denotes the partial transposition of B .

From that, the following theorem can be deduced.

Theorem 8. *Decomposable entanglement witnesses cannot detect PPT entangled states.*

This is true because we have $\text{Tr}(W\rho) = \text{Tr}(A\rho) + \text{Tr}(B^\Gamma\rho) = \text{Tr}(A\rho) + \text{Tr}(B\rho^\Gamma) \geq 0$, where we used that the trace is invariant under transposition and that $\rho^\Gamma \geq 0$ for all $\rho \in \text{PPT}$.

Equivalently, one can define an entanglement witness W as decomposable if $\text{Tr}(W\rho) \geq 0$ holds for all PPT states ρ (see, e.g. Ref. 35). In summary, nondecomposable witnesses have to be constructed to detect bound entangled states. Consequently, finding nondecomposable witnesses or bound entangled states is a dual problem and as such, a generally difficult problem.

4.1.2. Mirrored entanglement witnesses

Surprisingly, the authors of Ref. 36 found that the very definition of an entanglement witness (Definition 15) does often also have a nontrivial second bound U_W , namely

$$0 \leq \text{Tr}(W\sigma) \leq U_W, \quad \forall \sigma \in \text{SEP}, \quad (28)$$

such that one can also find entangled states that violate this new bound, i.e. $\text{Tr}(W\rho) > U_W$ for some $\rho \in \text{ENT}$. This can be formalized by the following definition.³⁶

Definition 17 (Mirrored entanglement witness). Given an entanglement witness W , one defines a mirrored operator by

$$W_M = \mu \mathbb{1}_{d_A} \otimes \mathbb{1}_{d_B} - W, \quad (29)$$

with the smallest $\mu > 0$ such that $\text{tr}(W_M \sigma) \geq 0$ for all $\sigma \in \text{SEP}$. Moreover, if the maximal eigenvalue of W satisfies $\lambda_{\max} > \mu$, then W_M is an entanglement witness, and hence one has a pair (W, W_M) of mirrored entanglement witnesses.

The concept is summarized in Fig. 5. The idea comes from the observation that for any positive semi-definite matrix $K \geq 0$ with $\text{Tr}(K) = 1$ and $K \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$, which can be interpreted as a state on a composite system, we can find an upper bound U and a lower bound L defined by $U(K) := \max_{\sigma} (\text{Tr}(K \sigma))$ and $L(K) := \min_{\sigma} (\text{Tr}(K \sigma))$ with $\sigma \in \text{SEP}$. Obviously, it follows that

$$L(K) \leq \text{Tr}(K \sigma) \leq U(K) \quad \forall \sigma \in \text{SEP}. \quad (30)$$

Let us denote by $\lambda_{\min}(K)$ and $\lambda_{\max}(K)$ the minimal and maximal eigenvalues of K , respectively. Then, if $\lambda_{\min}(K) < L(K)$ and/or $U(K) < \lambda_{\max}(K)$ the bounds can be used to detect entangled states. Namely, one finds

$$\text{Tr}(K \rho) \notin [L(K), U(K)] \Rightarrow \rho \in \text{ENT}. \quad (31)$$

The interval $[L(K), U(K)]$ is called the *separability window* of the observable K .

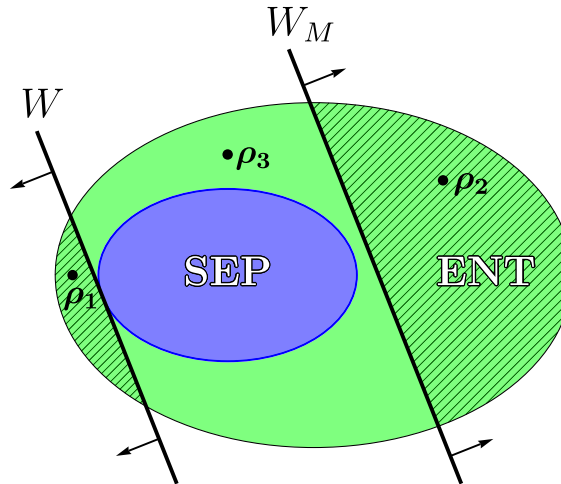


Fig. 5. This picture depicts the quantum state space, separated into the sets SEP and ENT. The witness W detects the entanglement of the state ρ_1 , but neither of ρ_2 nor ρ_3 . On the other hand, the mirrored witness W_M based on the same local observables can also witness the entanglement of ρ_2 . Therefore, the pair of witnesses $\{W, W_M\}$ can detect, in general, more entangled states than only considering one of the two witnesses.

Due to linearity, each bound can be associated with an entanglement witness, i.e.

$$W^{(+)} = \frac{1}{1 - d_1 d_2 L(K)} (K - L(K) \mathbb{1}_{d_A} \otimes \mathbb{1}_{d_B}), \quad (32)$$

$$W^{(-)} = \frac{1}{d_1 d_2 U(K) - 1} (-K + U(K) \mathbb{1}_{d_A} \otimes \mathbb{1}_{d_B}). \quad (33)$$

Thus, we have

$$\begin{aligned} 0 \leq \text{Tr}(W^{(+)} \rho) &\Leftrightarrow L(K) \leq \text{Tr}(K \rho), \\ 0 \leq \text{Tr}(W^{(-)} \rho) &\Leftrightarrow \text{Tr}(K \rho) \leq U(K). \end{aligned} \quad (34)$$

This means that from one observable K with a nontrivial separability window, we constructed a pair of witnesses $\{W^{(+)}, W^{(-)}\}$. Two witnesses can potentially detect more entangled states than one. Moreover, note that those witnesses are constructed from the same experimental settings.

As proven in detail in Ref. 36, there is a universal relation between the pair $\{W^{(+)}, W^{(-)}\}$, so that mirrored entanglement witnesses can directly be constructed from each other, which coincides with Definition 17.

Furthermore, the authors of Ref. 37 showed that there exist mirrored pairs of entanglement witnesses that are both optimal and nondecomposable. Consequently, the concept of mirrored entanglement witness can aid in detecting PPT-entangled states in an optimal way. This also refutes a recent conjecture posed in Ref. 38.

4.1.3. Nonlinear entanglement witnesses

Entanglement witnesses are linear hyperplanes that distinguish entangled states from the set of separable states in the space of Hermitian operators. Since the set of separable states is convex, one may improve a given entanglement witness by introducing nonlinearities such that the convex set is better approximated and a larger set of entangled states can be detected.

The first ideas in this direction came from uncertainty relations which are inequalities in terms of variances.^{39,40} The variance of an observable K for a state ρ is defined as $\Delta(K)_\rho = \sqrt{\text{Tr}(K^2 \rho) - \text{Tr}(K \rho)^2}$, which is clearly nonlinear in ρ , and is shown to be capable to detect entanglement in Ref. 40.

Another, more systematic method of constructing nonlinear entanglement witnesses is introduced in Ref. 40. Suppose that some entangled state ρ is detected by a positive but not completely positive map Λ , that is, $(\mathbb{1} \otimes \Lambda)(\rho)$ possesses an eigenvector $|\varphi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ corresponding to a negative eigenvalue. Then the formula

$$W = (\mathbb{1} \otimes \Lambda^*)(|\varphi\rangle\langle\varphi|), \quad (35)$$

where Λ^* denotes a dual map, defines an entanglement witness detecting ρ , i.e. $\text{Tr}(W \rho) < 0$. A nonlinear functional is then constructed by subtracting polynomials from an entanglement witness. To be specific, let us consider subtracting a single

nonlinear polynomial as follows:⁴⁰

$$\mathcal{F}[\rho] = \text{Tr}(W\rho) - \frac{1}{s^2(\chi)} \text{Tr}(\tilde{X}\rho) \cdot \text{Tr}(\tilde{X}^\dagger \rho), \quad (36)$$

where one can choose $\tilde{X} = (\mathbb{1} \otimes \Lambda^*)(|\varphi\rangle\langle\chi|)$ for some state $|\chi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and $s(\chi)$ is the largest Schmidt coefficient of $|\chi\rangle$. The parameters are chosen such that $\mathcal{F}[\sigma] \geq 0$ for all $\sigma \in \text{SEP}$. It is clear that a witness W does not detect a larger set of entangled states than its nonlinear functional $\mathcal{F}$ since $\mathcal{F}[\rho] \leq \text{Tr}(W\rho)$.

4.2. Entanglement witnesses based on MUBs and SICs and generalizations

Any entanglement witnesses can be decomposed into local measurements $A_i \otimes B_j \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$:

$$W = \sum_{i,j} w_{i,j} A_i \otimes B_j. \quad (37)$$

This is the important connection of the theory to any experimental realizations, where it is of practical importance to minimize the number of local observables.

An interesting class of entanglement witnesses arises from sets of mutually unbiased bases (MUBs) that are experimentally very feasible (see also Sec. 6). They also gradually approach a set that can be used for quantum state tomography as they converge to a complete set of MUBs.^d Therefore, it is clear that a well-chosen complete set of MUBs should, in principle, be able to detect bound entanglement, at least indirectly. Surprisingly, as presented in Sec. 4.2.1, a complete set is not necessary to witness bound entanglement. A smaller set of experimental settings can sometimes be found, which is a sign of a substructure of the set of bound entangled states.

A complete set of MUBs is a 2-design⁴¹ such as SICs (symmetrical informationally complete), Sec. 4.2.3. SIC witnesses can be constructed by applying the same method as for a set of MUBs. As shown in Sec. 4.1.2, every entanglement witness has a mirrored one. For MUB- and SIC-witnesses the corresponding nontrivial upper bounds become strongly dependent on the dimension d because different choices exist of complete MUB or SIC bases. This reveals that the entanglement structure is involved when the dimension d increases and by that, the detection of bound entanglement.

4.2.1. A Complete MUB set is not needed to detect bound entanglement

Mutually unbiased bases are sets of different bases with the property that a measurement in one basis reveals no information about the measurement statistics for a measurement in another. Mathematically, this can be defined as follows:

^dA set of MUBs is called complete if it contains the maximal number of MUBs. For prime-power dimensions, this is known to be $d + 1$.

Definition 18 (Mutually unbiased bases). A pair of orthonormal bases $\mathcal{B}_k$ and $\mathcal{B}_l$ of the Hilbert space $\mathbb{C}^d$ are mutually unbiased if the basis vectors satisfy $|\langle i_k | j_l \rangle|^2 = \frac{1}{d}(1 - \delta_{k,l}) + \delta_{k,l}\delta_{i,j}$ for all $|i_k\rangle \in \mathcal{B}_k$ and $|j_l\rangle \in \mathcal{B}_l$. A set of pairwise mutually unbiased bases $\{\mathcal{B}_k\}_k$ is referred to as mutually unbiased bases (MUBs).

This condition formalizes mathematically the concept of complementarity, i.e. if the system resides in an eigenstate of one observable, the outcome of measuring a complementary observable is maximally uncertain. Generalizing to larger collections of measurements, we say a set of bases is mutually unbiased if each pair of bases is mutually unbiased. For prime-power dimensions, it has been shown that there exists a complete set of exactly $d + 1$ MUBs^{24,42} and explicit constructions thereof, e.g. based on the Weyl–Heisenberg operators (14) (see Ref. 43 for a review). However, it is conjectured that fewer MUBs exist in all other cases.⁴⁴ In particular, for dimension $d = 6$, so far, only a set of 3 MUBs has been found.

The first direct connection between MUBs and entanglement witnesses is made in Ref. 45, which we show below. The one for dimension $d = 3$ was used for the first experiment witnessing bound entanglement (see Sec. 6). Generalizations of the witnesses to a wider class of measurements, including 2-designs and mutually unbiased measurements (MUMs), for which MUBs are an example, are investigated in Ref. 35,41,46.

We derive entanglement witnesses from a collection of MUBs. Let us consider m MUBs in a d -dimensional Hilbert space,

$$\mathcal{M}_m = \{\mathcal{B}_0, \mathcal{B}_1, \dots, \mathcal{B}_{m-1}\}, \quad (38)$$

where $\mathcal{B}_k = \{|i_k\rangle\}_{i=0}^{d-1}$ and $\mathcal{B}_0$ is chosen as the computational basis $\{|i_0\rangle\} = \{|0\rangle, \dots, |d-1\rangle\}$. From these MUBs, we construct the operators

$$\mathbf{B}(\mathcal{M}_m, s) = \sum_{l=0}^{d-1} |l\rangle\langle l| \otimes |l+s\rangle\langle l+s| + \sum_{k=1}^{m-1} \sum_{i=0}^{d-1} |i_k\rangle\langle i_k| \otimes |i_k\rangle\langle i_k| \quad (39)$$

and its partial transpose

$$\mathbf{B}^\Gamma(\mathcal{M}_m, s) = \sum_{l=0}^{d-1} |l\rangle\langle l| \otimes |l+s\rangle\langle l+s| + \sum_{k=1}^{m-1} \sum_{i=0}^{d-1} |i_k\rangle\langle i_k| \otimes |i_k^*\rangle\langle i_k^*|, \quad (40)$$

where the complex conjugation (and transposition) is performed with respect to the canonical basis $\mathcal{B}_0$. It is proven in Ref. 45 that for any separable state $\sigma \in \text{SEP}$, the two operators are bounded from above by the same quantity, depending on the number of MUBs, m , and the dimension d . In particular,

$$\text{Tr}(\mathbf{B}(\mathcal{M}_m, s)\sigma) \leq \frac{d + m - 1}{d} \quad (41)$$

and

$$\text{Tr}(\mathbf{B}^\Gamma(\mathcal{M}_m, s)\sigma) \leq \frac{d + m - 1}{d}. \quad (42)$$

Hence, if the operators

$$W(\mathcal{M}_m, s) = \frac{d+m-1}{d} \mathbb{1}_d \otimes \mathbb{1}_d - \mathbf{B}(\mathcal{M}_m, s) \quad (43)$$

and $W^\Gamma(\mathcal{M}_m, s)$ possess a strictly negative eigenvalue, they satisfy the very definition of an entanglement witness (Definition 15) with

$$\begin{aligned} 0 &\leq \text{Tr}(W(\mathcal{M}_m, s)\sigma) \quad \text{for all } \sigma \in \text{SEP}, \\ 0 &\leq \text{Tr}(W^\Gamma(\mathcal{M}_m, s)\sigma) \quad \text{for all } \sigma \in \text{SEP}. \end{aligned} \quad (44)$$

Only nondecomposable entanglement witnesses can detect PPT entangled states (see Theorem 8). This property depends on the choice s :⁴⁵

Theorem 9. *The entanglement witnesses $W(\mathcal{M}_{m=d+1}, s=0)$ and $W^\Gamma(\mathcal{M}_{m=d+1}, s=0)$ are decomposable.*

Consequently, they cannot detect PPT entangled states. But note that the entanglement witnesses $W^\Gamma(\mathcal{M}_m, s=0)$ are anyhow powerful witnesses as they detect all possible entangled states of the family of isotropic states $\{\rho_{iso}(p) = pP_{k,l} + (1-p)\frac{1}{9}\mathbb{1}_9\}_{k,l=0}^{d-1}$ (where $P_{k,l}$ is one of the Bell states (16)). Via this family of states, it is also proven that the following theorem holds:⁴⁵

Theorem 10. *No more than $d+1$ MUBs can exist. Otherwise, $W^\Gamma(\mathcal{M}_m, s=0)$ would witness separable states as entangled states.*

We are interested in nondecomposable entanglement witnesses since only those can detect bound entangled states. As already mentioned above, the nondecomposability depends on the choice of s . In particular, it has been found that:⁴⁷

Theorem 11. *Let $s > 0$ and $d > 2$. Then $W^\Gamma(\mathcal{M}_{m=d+1}, s)$ defines a nondecomposable entanglement witness if and only if $s \neq \frac{d}{2}$.*

Finally, in strong contrast to SICs (see Sec. 4.2.3), a fascinating fact for entanglement witnesses based on MUBs has been proven, namely a complete set of MUBs is not needed to detect bound entangled states. In detail:

Theorem 12. *Let $s \neq \frac{d}{2}$ with $s > 0$ and $d > 2$. If $m > \frac{d}{2} + 1$, then $W^\Gamma(\mathcal{M}_m, s)$ defines a nondecomposable entanglement witness, i.e. it can detect bound entangled states.*

Particularly, for $d=3$, exactly 4 mutually unbiased bases exist. However, bound entangled states can also be detected by choosing only 3 MUBs, considerably reducing the number of different experimental setups. In Fig. 6, we present two slices through the Magic Simplex (17) visualizing the effect of a MUB witness based on 4 versus 3 MUBs.

This brings us to the next question: Can a different subset of MUBs be more or less efficient than another subset in detecting entanglement? Indeed, this is the case⁴⁶ and is discussed in the next section.

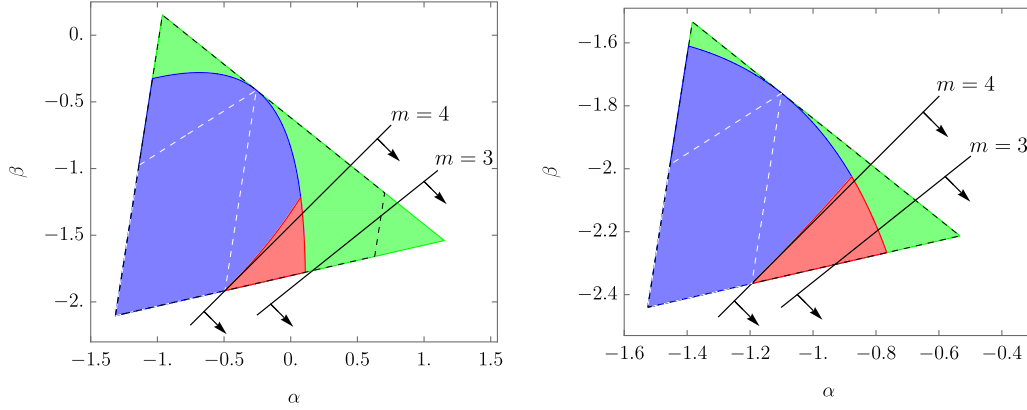


Fig. 6. These graphics depict two slices, B_1 (left; $\gamma = -1/\sqrt{3}$) and B_2 (right; $\gamma = -0.83$), through the Magic Simplex for bipartite qutrits in the state family $\rho_B = \frac{1-\frac{\alpha}{5}-\frac{\beta}{4}-\gamma}{9} \mathbb{1}_9 + \frac{\alpha}{5} P_{0,0} + \frac{\beta}{8} \sum_{i=1}^2 P_{i,0} + \frac{\gamma}{3} \sum_{i=0}^2 P_{i,1}$. NPT states are colored green. The red area corresponds to bound entangled states detected by the realignment criterion (Theorem 14). The MUB witnesses $W^\Gamma(\mathcal{M}_m, 1)$, solid lines with arrows, are drawn for the maximal number of MUBs $m = 4$ and $m = 3$. Only for family B_2 can it detect entanglement in the PPT region (blue and red). The white dashed lines correspond to the kernel polytope (20); all states inside must be separable. The black dashed lines correspond to the enclosure polytope (18); all states outside must be NPT entangled.

4.2.2. On the universality of MUBs for nondecomposable witnesses

It has been shown that certain sets of MUBs are more useful than others in detecting entanglement⁴⁶ — a consequence of the unitary inequivalence of different classes of MUBs. In this sense, MUBs are not universal. Thus, it is expected that MUB witnesses do not only depend on the number of MUBs m (see Theorem 12) but also on the particular choice of the set of MUBs. In Ref. 46, it is shown that for $d \geq 4$, inequivalent sets differ in their efficiency of detecting entanglement. Moreover, there are not only unitarily inequivalent sets of complete MUBs but also unextendible MUB sets, i.e. sets that cannot be extended to a complete set of MUBs, which are more effective in detecting entanglement.

The next section discusses the mirrored entanglement witnesses of MUB witnesses, where the lower bounds show their dependence not only on the number of MUBs but the particular choice of MUBs within the MUB set.

In summary, Theorem 12 shows that, since *any* set of m MUBs ($m > d/2 + 1$) can be used to construct a nondecomposable entanglement witness that detects bound entangled states, the nondecomposability of such a witness is a universal property of mutually unbiased bases. However, since different equivalence classes of MUBs give rise to different entanglement witnesses, an interesting interplay is revealed between the universality of MUBs for nondecomposable witnesses and the dependence on the local choice of MUBs in detecting a given bound entangled state. The nature of such connections would be worth investigating further since they may shed light on the structure of the set of bound entangled states.

4.2.3. Entanglement witnesses based on sets of SICs

Let us consider a special set of POVMs (see also Sec. 2.2.2) called *symmetric informationally complete* POVMs (abbreviated SICs).

Definition 19 (Symmetric informationally complete POVMs). Let $\{\Pi_s\}_{s=1}^{d^2}$ be a set of rank-1 projectors with

$$\text{Tr}(\Pi_s \Pi_r) = \frac{1}{d+1}(1 + d\delta_{sr}), \quad (45)$$

such that the set $\{\frac{1}{d}\Pi_s\}_{s=1}^{d^2}$ forms an informationally complete POVM, i.e. a set of d^2 POVM elements that span the whole Hilbert space. Then, the set $\{\Pi_s\}_{s=1}^{d^2}$ is called symmetric informationally complete (SIC) and referred to as SICs.

SICs are a special case of generalized measurements on a Hilbert space like MUBs, therefore considering the full set of MUBs or SICs needs to be equivalent for entanglement detection because both are so-called 2-designs.⁴¹ In contrast to MUBs, the existence of SICs is conjectured for any d , the largest dimension for which an exact solution has been found is $d = 323$.⁴⁸

Following the idea behind entanglement witnesses based on MUBs (presented in Sec. 4.2.1), we can construct a witness based on a set of complete basis-POVMs in the same manner.

In an experiment, the following correlation functions C are bounded from below and from above:

$$L_{m,d}^{\text{MUB}} \leq C_m := \sum_{k=1}^m \sum_{i=0}^{d-1} \text{Tr}(|i_k\rangle\langle i_k| \otimes |i_k\rangle\langle i_k| \rho) \leq U_{m,d}^{\text{MUB}}, \quad (46)$$

$$L_{\tilde{m},d}^{\text{SIC}} \leq C_{\tilde{m}} := \sum_{k=1}^{\tilde{m}} \text{Tr}(|s_k\rangle\langle s_k| \otimes |s_k\rangle\langle s_k| \rho) \leq U_{\tilde{m},d}^{\text{SIC}}, \quad (47)$$

with $|\langle i_k | j_l \rangle|^2 = \frac{1}{d}(1 - \delta_{k,l}) + \delta_{k,l}\delta_{i,j}$ and $|\langle s_k | s_j \rangle|^2 = \frac{1}{d+1}(1 - \delta_{k,j}) + \delta_{k,j}$. Here U and L denote the bounds over all separable states, i.e. a violation of those inequalities detects entanglement. The bounds depend on the number of MUBs m and on the number of SICs $\tilde{m}$ that are applied, respectively. Note that $[L, U]$ is the separability window (31) of the pair of mirrored witnesses which can be constructed from C_m or $C_{\tilde{m}}$ (see Sec. 4.1.2).

The upper bound for the MUBs has been derived for all dimensions to be $U_{m,d}^M = 1 + \frac{m-1}{d}$ and is independent of the choice of the set of MUBs.⁴⁵ A particular reordering of the MUB or SIC projectors for one of the two subsystems in C_m or $C_{\tilde{m}}$ changes, in general, the upper and lower bounds, which can be exploited to detect bound entanglement (cf. Sec. 4.2.1). In Tables 1 and 2, taken from Ref. 41, we present results of the lower and upper bounds on the expression C_m and $C_{\tilde{m}}$ as defined above. In Sec. 6, we present an experiment that utilizes the above expression C_m based on MUBs to detect bound entanglement (by applying a particular reordering in one subsystem).

Table 1. [Table taken from Ref. 41.] This table shows the results of the optimization over all separable states for the observable (46), which is based on a set of m MUBs. The upper bounds U only depend on the number m of MUBs; however, the lower bounds L are starting to depend on the choice of the MUBs for $d \geq 4$. Results for higher dimensions can be found in Ref. 46.

m	Lower bounds L				Upper bounds U		
	$d = 2$	$d = 3$	$d = 4$		$d = 2$	$d = 3$	$d = 4$
	$L_{m,2}^{(\text{MUB})}$	$L_{m,3}^{(\text{MUB})}$	$L_{m,4}^{-(\text{MUB})}$	$L_{m,4}^{+(\text{MUB})}$	$U_{m,2}^{(\text{MUB})}$	$U_{m,3}^{(\text{MUB})}$	$U_{m,4}^{(\text{MUB})}$
2	0.5	0.211	0	0	1.5	1.333	1.125
3	1	0.5	0.25	0.5	2	1.666	1.5
4		1	0.5	0.5		2	1.725
5			1	1			2

Table 2. [Table taken from Ref. 41.] The lower and upper bounds, L and U , respectively, for entanglement witnesses based on $\tilde{m}$ SICs are shown for $d = 2, 3$. We use the Hesse SICs for $d = 3$ defined in Ref. 50. Note that for $d = 2$, the smallest and largest lower/upper bounds, $L^\pm/U^\pm$, are equal, whereas for $d > 2$, they may differ (marked in bold). This strongly contrasts MUBs, where only the lower bounds can vary.

$\tilde{m}$	Lower bounds L			Upper bounds U		
	$d = 2$	$d = 3$		$d = 2$	$d = 3$	
	$L_{\tilde{m},2}^{(\text{SIC})}$	$L_{\tilde{m},3}^{-(\text{SIC})}$	$L_{\tilde{m},3}^{+(\text{SIC})}$	$U_{\tilde{m},2}^{(\text{SIC})}$	$U_{\tilde{m},3}^{+(\text{SIC})}$	$U_{\tilde{m},3}^{-(\text{SIC})}$
2	0	—	—	1.244	—	—
3	0.266	0	0	1.333	1.254	1.125
4	0.666	0	0	1.333	1.4	1.25
5	—	0	0	—	1.463	1.400
6	—	0	0.112	—	1.5	1.482
7	—	0.15	0.15	—	1.5	1.5
8	—	0.375	0.375	—	1.5	1.5
9	—	0.75	0.75	—	1.5	1.5

For the MUB case, the lower bounds depend on the choice of a set of MUBs for dimensions greater than $d = 4$. Here, inequivalent sets of m MUBs exist.⁴⁹ Optimizing C_m over all separable states, one finds the smallest and the largest lower bound $L^\pm$ if $m = 3$ (Table 1). Thus, an experimenter needs to optimize the MUB observables with respect to the source's state or try out all options. Otherwise, they may not be able to detect the entanglement.

In the SIC case, in strong contrast to MUBs, different sets of SICs affect lower and upper bounds, $L^\pm, U^\pm$, for some $\tilde{m}$ numbers of SICs (Table 2), and it starts already with the first nontrivial dimension, i.e. $d = 3$. In practice, this means that witnesses based on SIC observables need local information about the state to be optimized.

In Fig. 7, we visualize the entanglement witnesses depending on different numbers of MUBs or SICs for a slice in the Magic Simplex for two qutrits, (17), of the

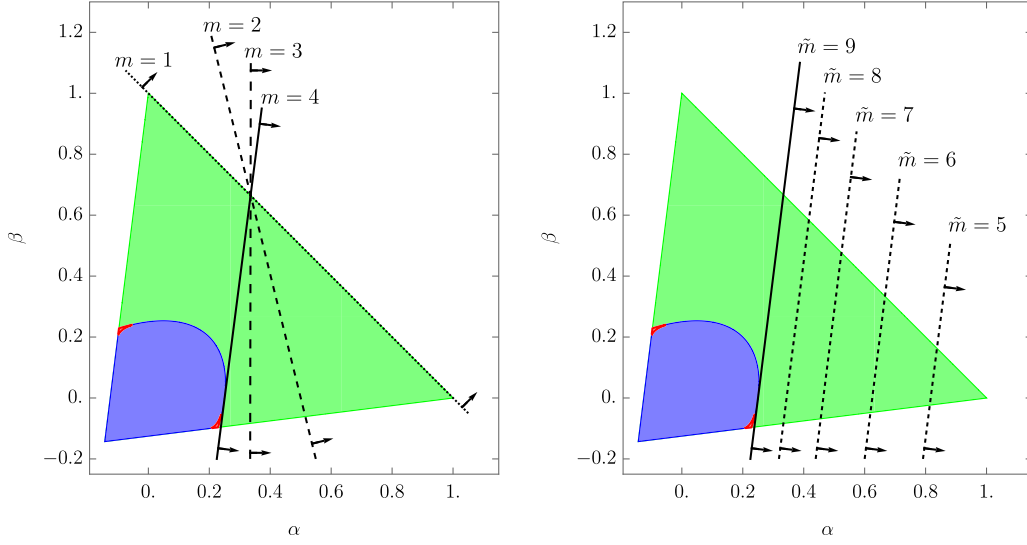


Fig. 7. These graphics show a 2-dimensional slice through the 8-dimensional Magic Simplex for two qutrits for the state family $\rho_A = \frac{1-\alpha-\beta}{9} \mathbb{1}_9 + \alpha P_{0,0} + \beta P_{1,0}$. The green-colored region corresponds to NPT states. The red area corresponds to PPT states for which the quasi-pure approximation (cf. Sec. 4.6) detects bound entanglement. On the left, the MUB witnesses for different numbers m of MUBs are shown, and on the right, the SIC witnesses for different numbers $\tilde{m}$ of SICs. None of the witnesses can detect bound entanglement; however, these graphics illustrate that those observables detect different sets of entangled states.

mixture of the totally mixed states and two Bell states. Both MUBs and SICs are equivalent over a 2-design if they form a complete set, and therefore, the corresponding witnesses are also equivalent.⁴¹ In this slice, they represent the optimal witnesses for the family of isotropic states $\{\rho_{iso}(p) = pP_{k,l} + (1-p)\frac{1}{9}\mathbb{1}_9\}_{k,l=0}^{d-1}$. As can be deduced from Fig. 7, fewer MUBs or SICs detect a smaller region of the entangled states; however, which region is detected is very different for MUBs and SICs. Thus, incomplete sets of those two POVMs can reveal different entangled states in the Hilbert space. In this particular case, the MUB witness “rotates” in this slice, whereas for the SIC witness, we find a “parallel shift” for this slice in the Magic Simplex.

Less is known for higher dimensions since there is little knowledge about the constructions of the MUBs and SICs in general dimensions. For dimensions up to $d = 9$, results for MUBs are presented in Ref. 46.

Due to the mentioned equivalence over a 2-design, all bound entangled states that are detected by entanglement witnesses based on MUBs are also detected by the corresponding SIC witnesses. In contrast to MUBs (see Theorem 12), using an incomplete set of SICs, no example of a bound entangled state that is detected by the corresponding witness has been found so far. This also holds for different generalizations of SICs (see also Sec. 4.5).

In summary, sets of MUBs or SICs are typically easily constructed observables in experiments (see Sec. 6) and as shown, capable of detecting bound entangled states

since, in general, nondecomposable entanglement witnesses can be constructed. Moreover, those entanglement witnesses can be mirrored, giving rise to two bounds, L and U . Generally, these bounds depend on the specific set of MUBs. Interestingly, the upper bounds for the MUB witnesses depend solely on the number of MUBs.

4.3. Criteria based on a correlation tensor

In Ref. 51, a family of entanglement criteria is introduced that is based on the Bloch representation of bipartite quantum states $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ of dimension $d = d_A d_B$.

Let G_α^A be an orthonormal basis of $\mathcal{L}(H_A)$ with respect to the Hilbert–Schmidt inner product $\text{Tr}(G_\alpha^{A\dagger} G_\beta^A) = \delta_{\alpha,\beta}$ and $G_0^A = \frac{1}{d_A} \mathbb{1}_A$ (similar for system B). The Bloch representation of a state ρ is then

$$\begin{aligned} \rho = & \frac{1}{d_A d_B} \mathbb{1}_A \otimes \mathbb{1}_B + \sum_{i=1}^{d_A^2-1} \frac{r_i^A}{d_B} G_i^A \otimes \mathbb{1}_B \\ & + \sum_{j=1}^{d_B^2-1} \frac{r_j^B}{d_A} \mathbb{1}_A \otimes G_j^B + \sum_{i=1}^{d_A^2-1} \sum_{j=1}^{d_B^2-1} t_{i,j} G_i^A \otimes G_j^B \end{aligned} \quad (48)$$

$$=: \sum_{\alpha=0}^{d_A^2-1} \sum_{\beta=0}^{d_B^2-1} C_{\alpha,\beta} G_\alpha^A \otimes G_\beta^B. \quad (49)$$

Let us define two diagonal matrices $D_x^A := \text{diag}(x, 1, \dots, 1)$ and $D_y^B := \text{diag}(y, 1, \dots, 1)$ of dimension d_A^2 and d_B^2 , respectively. With the matrix $C = (C_{\alpha,\beta})$, the following entanglement criterion can be formulated.⁵¹

Theorem 13 (Correlation tensor criterion). *A state $\rho \in \text{ENT}$ if there exist $x, y \geq 0$ such that*

$$\text{Tr} \left(\sqrt{D_x^A C D_y^B (D_x^A C D_y^B)^\dagger} \right) > \sqrt{\frac{d_A - 1 + x^2}{d_A}} \sqrt{\frac{d_B - 1 + y^2}{d_B}}. \quad (50)$$

This family of criteria reproduces several other known entanglement criteria. For $x = y = 1$, it reproduces the realignment criterion (cf. Sec. 4.4), while for $x = y = 0$ it is equivalent to the so-called *de Vicente criterion*.⁵² It is shown in Ref. 51 that the correlation tensor criterion is also equivalent to the ESIC criterion (cf. Sec. 4.5) for $(x, y) = (\sqrt{d_A + 1}, \sqrt{d_B + 1})$. Moreover, criteria based on this correlation tensor allow deriving a corresponding class of entanglement witnesses (cf. Sec. 4.1; Ref. 51).

4.4. Realignment criterion

The realignment criterion (also named cross norm criterion) is presented in Ref. 53 in matrix notation and in Ref. 54 in Dirac notation. It is based on the singular values of a realigned bipartite density matrix.

The realignment operator R can be defined via its action on the basis states of a bipartite system

$$R : |i\rangle\langle j| \otimes |k\rangle\langle l| \mapsto (|i\rangle\langle j| \otimes |k\rangle\langle l|)_R := |i\rangle\langle k| \otimes |j\rangle\langle l| \quad (51)$$

and extends to general density matrices $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by linearity:

$$R : \rho \mapsto \rho_R. \quad (52)$$

In Refs. 53 and 54 it is shown that if $\rho \in \text{SEP}$, then $\text{Tr}(\sqrt{\rho_R^\dagger \rho_R}) \leq 1$. Conversely, if the sum of singular values of a realigned density matrix is greater than one, then the state must be entangled. Thus, the realignment or cross norm criterion reads:

Theorem 14 (Realignment criterion/cross norm criterion).

$$\text{Tr}(\sqrt{\rho_R^\dagger \rho_R}) > 1 \Rightarrow \rho \in \text{ENT}. \quad (53)$$

The realignment criterion is neither stronger nor weaker than the PPT criterion (Theorem 5), i.e. it does not detect all NPT entangled states but can detect PPT entanglement.

Considering Bell-diagonal states (cf. Sec. 2.3), the realignment criterion seems very effective for the detection of PPT entanglement for $d = 3$ and $d = 4$ (cf. Refs. 28 and 29).

4.5. Entanglement criteria based on SICs and GSICs

SICs, i.e. symmetric informationally complete POVMs, provide another way to derive entanglement criteria. Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, Π_α^A and Π_β^B be SICs (Definition 19) for $\mathcal{H}_A$ and $\mathcal{H}_B$ and define the correlation tensor $\mathcal{P}$ with elements

$$\mathcal{P}_{\alpha,\beta} := \text{Tr}(\Pi_\alpha^A \otimes \Pi_\beta^B \rho). \quad (54)$$

In Ref. 55, the following entanglement criterion is derived.

Theorem 15 (ESIC criterion).

$$\text{Tr}(\sqrt{\mathcal{P}^\dagger \mathcal{P}}) > \frac{2}{\sqrt{d_A(d_A+1)d_B(d_B+1)}} \Rightarrow \rho \in \text{ENT}. \quad (55)$$

A similar criterion can be stated for *general symmetric informationally complete POVMs* (GSICs). Compared to SICs, for GSICs the POVM elements $\{\Pi_\alpha^A\}$ are not required to be of rank-1, but instead satisfy

$$\text{Tr}(\Pi_\alpha^A \Pi_\alpha^A) = a_A, \quad (56)$$

$$\text{Tr}(\Pi_\alpha^A \Pi_\gamma^A) = \frac{1 - d_A a_A}{d_A(d_A^2 - 1)}, \quad \text{for } \alpha \neq \gamma \quad (57)$$

for system $\mathcal{H}_A$ and similarly for POVM elements $\{\Pi_\beta^B\}$ for $\mathcal{H}_B$. With the corresponding correlation tensor (54), the GSIC criterion reads:⁵⁶

Theorem 16 (GSIC criterion).

$$\text{Tr}(\sqrt{\mathcal{P}\mathcal{P}^\dagger}) > \sqrt{\frac{a_A d_A^2 + 1}{d_A(d_A + 1)}} \sqrt{\frac{a_B d_B^2 + 1}{d_B(d_B + 1)}} \Rightarrow \rho \in \text{ENT}. \quad (58)$$

4.6. Concurrence and quasi-pure approximation

The quasi-pure approximation $\mathcal{C}_{qp}$ ⁵⁷ of the concurrence $\mathcal{C}$, presented in Ref. 58, is another useful sufficient criterion to detect entanglement in systems of finite dimension.

The concurrence, introduced in Ref. 58, is an entanglement monotone closely related to the entanglement of formation (cf. Definition 23). It can only be explicitly calculated for a pair of qubits, while for systems of larger size, only numerical solutions or estimates exist.^{58–61} The quasi-pure approximation provides a lower bound for the concurrence of a quantum state

$$\mathcal{C}_{qp}(\rho) \leq \mathcal{C}(\rho). \quad (59)$$

Given the spectral decomposition $\rho = \sum_{i=0}^{d-1} \lambda_i |\psi_i\rangle\langle\psi_i|$, a matrix $\mathcal{T}$ is constructed, whose singular values determine the quasi-pure approximation. Let $A := 4 \sum_{i < j, k < l} (|ikjl\rangle - |jkil\rangle - |iljk\rangle + |jljk\rangle) \times (h.c.)$ ($h.c.$ being shorthand for Hermitian conjugate) and let $|\psi_0\rangle$ be the dominant eigenvector. With $|\chi\rangle \propto A|\psi_0\rangle \otimes |\psi_0\rangle$, $\mathcal{T}$ is defined as

$$\mathcal{T}_{ij} := \sqrt{\lambda_i \lambda_j} = (\langle\psi_i| \otimes \langle\psi_j|) |\chi\rangle. \quad (60)$$

Let $\mathcal{S}_i$ be the singular values of $\mathcal{T}$ and denote the largest one by $\mathcal{S}_0$. The quasi-pure approximation of the concurrence is defined as

$$\mathcal{C}_{qp}(\rho) := \max \left(0, \mathcal{S}_0 - \sum_{i>0} \mathcal{S}_i \right). \quad (61)$$

For any state in the Magic Simplex, i.e. standard Bell-diagonal states (cf. Sec. 2.3.1), there exists an explicit form for the singular values of $\mathcal{T}$ (cf. Ref. 57). Given a state of the form of Eq. (17), let $P_{n,m}$ be the Bell projector with the largest weight $c_{k,l}$. The singular values $\mathcal{S}_{k,l}$ are then explicitly given by

$$\mathcal{S}_{k,l} = \sqrt{\frac{d}{2(d-1)} c_{k,l} \left[\left(1 - \frac{2}{d} \right) c_{n,m} \delta_{k,n} \delta_{l,m} + \frac{1}{d^2} c_{2n-k, 2m-l} \right]}. \quad (62)$$

Any state satisfying $\mathcal{C}(\rho) > 0$ is entangled (cf. Ref. 58). Providing a lower bound, the quasi-pure approximation can therefore be directly used to detect entanglement.

Theorem 17 (Quasi-pure concurrence criterion).

$$\mathcal{C}_{qp}(\rho) > 0 \Rightarrow \rho \in \text{ENT}. \quad (63)$$

Like the realignment criterion (Theorem 14), this entanglement criterion can detect PPT entangled states but generally not all NPT entangled states. In Ref. 28, it is shown that the overlap of PPT entangled Bell-diagonal states detected by both the quasi-pure concurrence criterion and the realignment criterion is relatively small compared to the joint number of detected bound entangled states. This indicates a complementary behavior for these criteria, as a state detected by one criterion is likely not detected by the other (cf. Sec. 5.2.2).

4.7. Range criterion

The range criterion is of historical importance as it led to the discovery of the first known PPT entangled state. It is formalized as follows.⁷

Theorem 18 (Range criterion). *Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be separable with $\dim(\mathcal{H}_A \otimes \mathcal{H}_B) = d^2$. Then there exists a set $\{|\psi_i\rangle \otimes |\phi_j\rangle\}_{i,j=0}^{n-1} \subset \mathcal{H}_A \otimes \mathcal{H}_B$ of $n \leq d^2$ product vectors such that*

- (a) *the vectors $\{|\psi_i\rangle \otimes |\phi_j\rangle\}_{i,j=0}^{n-1}$ span the range of ρ and*
- (b) *the vectors $\{|\psi_i\rangle \otimes |\phi_j^*\rangle\}_{i,j=0}^{n-1}$ span the range of ρ^Γ .*

Conversely, if no such vectors exist, ρ is entangled.

Here ρ^Γ is the partial transposition of ρ with respect to $\mathcal{H}_B$ and $|\phi_j^*\rangle$ denotes the complex conjugate of $|\phi_j\rangle$ in the basis where the partial transposition is taken. An analogous version of the theorem also holds for the partial transposition ρ^{T_A} of ρ with respect to $\mathcal{H}_A$. In this case, the vectors $\{|\psi_i^*\rangle \otimes |\phi_j\rangle\}$ span the range of ρ^{T_A} if ρ is separable.

4.8. Reduction criterion and more criteria from PNCP maps

Like the PPT criterion (Theorem 5), the reduction criterion (Ref. 16) is based on a PNCP map (cf. Sec. 3.1). Consider the PNCP map $\Lambda : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$, $\dim(\mathcal{H}) = d$, acting as

$$\Lambda(\sigma) := \mathbb{1}_d \text{Tr}(\sigma) - \sigma. \quad (64)$$

Λ is a PNCP map, so it provides a necessary condition for separability for density matrices:

$$\rho \in \text{SEP} \Rightarrow (\Lambda \otimes \mathbb{1}_B)(\rho) \geq 0. \quad (65)$$

The contrapositive of this statement can be formulated using the reduced density matrix $\rho_B := \text{Tr}_A(\rho)$.

Theorem 19 (Reduction criterion).

$$\mathbb{1}_A \otimes \rho_B - \rho \not\geq 0 \Rightarrow \rho \in \text{ENT}. \quad (66)$$

Reference 16 shows that the reduction criterion is weaker than the PPT criterion. Any state detected as entangled by the reduction criterion is also detected by the PPT criterion. Consequently, NPT entangled states exist that are not detected by the reduction criterion. Additionally, it is shown that all entangled states detected by the reduction criterion (Theorem 19) can be *distilled*, so only states that are detected by the PPT criterion and not by the reduction criterion are candidates for NPT bound entangled states.

However, there exist structurally similar criteria based on PNCP maps that can detect bound entanglement. In Ref. 62, a family of PNCP maps is presented that allows to detect two PPT entangled qutrits due to Eq. (65).

5. Characteristics of Bound Entanglement

This section reviews several key properties and characteristics of bound entangled states. The goal of this investigation is twofold. The first is a logical follow-up to the last section, which summarizes our primary methods of detecting bound entanglement and shows that there are still many challenges. In particular, a hierarchical structure between the presented entanglement criteria is absent. Solving these issues requires further insight into the entanglement structure and the properties of bound entangled states. In this regard, the following sections clarify the role of impurity for bound entangled states and present results on the (nonzero) volume of the set BE, showing that bound entanglement cannot be neglected in general. Moreover, Sec. 5.4 introduces commonly used entanglement measures, and by showing the inequivalence of two such measures, the “bound nature” of the entanglement in bound entangled states is elucidated. Furthermore, the relation of bound entanglement to Bell inequalities and special relativity is discussed. The second goal is to address the role of bound entangled states in known quantum information processing tasks such as quantum teleportation or quantum key distribution. The advantage of bound entangled resource states over fully separable ones is limited to some of these schemes. Further insight into the structure of bound entanglement and its characteristics may contribute to a better understanding of entanglement as a resource and consequently, to novel applications in quantum technologies.

5.1. Rank of bound entangled states

From the discussion in Sec. 2.2.3, it is clear that all bound entangled states must be mixed. This follows from the fact that every entangled pure state has a Schmidt

rank greater than one and can thus be distilled with a nonzero distillation rate. Consequently, there are no pure, i.e. rank-1 bound entangled states. The logical follow-up question is: what are the general restrictions on the rank for bipartite bound entangled states?

A first advance to answering this question is presented in Ref. 63. The authors prove that if the rank of a bipartite quantum state ρ is strictly less than the rank of either of its reduced density matrices $\text{Tr}_A(\rho)$ and $\text{Tr}_B(\rho)$, then ρ is distillable. The contrapositive of this statement is formalized by

Theorem 20. *Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a bipartite quantum state and let $\rho_A = \text{Tr}_B(\rho)$ and $\rho_B = \text{Tr}_A(\rho)$ be its reduced states with respect to the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. If ρ is undistillable, then $\text{rank}(\rho) \geq \max\{\text{rank}(\rho_A), \text{rank}(\rho_B)\}$.*

In specific cases even the strict inequality $\text{rank}(\rho) > \max\{\text{rank}(\rho_A), \text{rank}(\rho_B)\}$ holds for undistillable states. In particular, it is shown in Ref. 64 that this includes all states violating the PPT criterion (Theorem 5). Therefore, no NPT bound entangled states with $\text{rank}(\rho) = \max\{\text{rank}(\rho_A), \text{rank}(\rho_B)\}$ exist.

As observed in Ref. 63, a consequence of Theorem 20 is that undistillable quantum states with rank n have support on at most an $n \times n$ -dimensional subspace of the underlying bipartite Hilbert space. Consequently, an undistillable rank-2 state can be viewed as a qubit state, and in this case, Theorem 2 implies that the state is separable. Thus, there are no rank-2 bound entangled states. Additionally, it is demonstrated in Ref. 65 that no rank-3 bound entangled states satisfying the PPT criterion exist. The authors of Ref. 66 show that this also holds for potential rank-3 NPT bound entangled states.

These findings leave open the possibility of bound entangled states with rank 4 or greater. From explicit constructions (cf. Sec. 7 and Refs. 67 and 68), we know that such rank-4 states do exist. The authors of Ref. 69 and 70 moreover argue that one can use these states of minimum rank to construct bound entangled states of rank 5 up to the full dimension $d_A d_B$ of the bipartite Hilbert space. The above results can be summarized by

Theorem 21. *No bound entangled states exist with rank 1, 2, or 3. Moreover, bound entangled states with rank 4 and greater can be explicitly constructed.*

5.2. Frequency of bound entangled states

A natural question to ask about the characteristics of entanglement classes (NPT, PPT, SEP, BE, FE) is: Given a certain quantum system and a randomly chosen quantum state, what are the probabilities of that state belonging to a certain entanglement class? Or equivalently: How big are the relative volumes of NPT, PPT, SEP, BE and FE within the quantum state space?

Even for the Hilbert space of bipartite quantum states $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, this question is not straightforward to answer as the results strongly depend on the properties

of the subsystems. In fact, it is shown in Ref. 71 that for infinite-dimensional (i.e. continuous-variable) systems, bipartite quantum states are generically distillable. At the same time, Ref. 72 demonstrates that the volume of undistillable states is finite for finite-dimensional systems. A fundamental difficulty in quantifying the volumes within the set of all quantum states is that calculations generally depend on the chosen probability measure. The choice of a “natural” measure on $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is not clear (cf. Ref. 73). In Sec. 5.2.1, we present results related to the frequency of bound entangled states that can be obtained despite this challenge.

The geometric representation of Bell-diagonal states as a $(d^2 - 1)$ -simplex in real space (cf. Sec. 2.3) allows determining frequencies of states in the entanglement classes SEP, BE and FE in this subset of states. These results are compared for the standard and generalized Bell-diagonal states in Sec. 5.2.2.

Since only PPT bound entangled states are detected, reported frequencies of states in BE correspond to the relative volume of $\text{PPT} \cap \text{ENT}$. This provides a lower bound for the volume of BE because the volume of NPT bound entangled states might be nonzero in certain systems. So far, however, no bound entangled state with negative partial transposition is known.

5.2.1. General bipartite bound entangled states

For general systems, the relative amount of bound entangled states in the bipartite Hilbert space strongly depends on the dimensions of the subsystems.

As mentioned above, a unique, “natural” measure for the space $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is not clear, *a priori*. Let $d_{AB} = d_A d_B$ be the dimension of the bipartite system. In Ref. 72,74, a family of product measures is proposed, which follows from the spectral decomposition of any density matrix $\rho = UCU^\dagger$. The unitary matrix $U \in \mathcal{U}(d_{AB})$ ($\mathcal{U}(d_{AB})$ denotes the set of unitary operators on $\mathcal{H}_A \otimes \mathcal{H}_B$) is not unique. $C = \text{diag}(c_1, \dots, c_{d_{AB}})$ is the diagonal matrix of eigenvalues of ρ with the trace condition $\sum_i c_i = 1$. A measure μ on the space $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ can then be defined as $\mu = \nu_C \times \nu_H$. Here, ν_C is a measure on the $(d_{AB} - 1)$ -dimensional simplex formed by the eigenvalues c_i . ν_H is the Haar measure defined on $\mathcal{U}(d_{AB})$. While the Haar measure ν_H can indeed be motivated as a “natural” measure on $\mathcal{U}(d_{AB})$ due to its invariance under unitary transformations, the choice of ν_C is more arbitrary. The normalized Lebesgue measure $\mathcal{L}_{d_{AB}-1}$ corresponds to a uniform and rotationally invariant distribution on the simplex as a submanifold of $\mathbb{R}^{d_{AB}}$. However, based on information-theoretic arguments, other distributions and induced measures can also be considered (cf. Refs. 73 and 74).

Using the product measure $\mu = \mathcal{L}_{d_{AB}-1} \times \nu_H$, it is shown in Ref. 72 that there exists a neighborhood of the maximally mixed state $d_{AB}^{-1} \mathbb{1}_{d_{AB}}$ which only contains separable states and has nonzero measure (less than one). Moreover, it is shown that the measure of PPT entangled states is larger than zero if $d_{AB} > 6$. As those states are bound entangled, the same holds for the set BE. In Ref. 74, it is argued that this is also true for any nonsingular measure μ for finite dimension d_{AB} , and

we can state the following:

Theorem 22. *The measures $\mu(\text{SEP})$ of separable states in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ for finite d_{AB} -dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is nonzero and less than one. If $d_{AB} > 6$, the same holds true for the measure $\mu(\text{BE})$ of bound entangled states.*

Investigations in Ref. 73 show that the volumes depend on the chosen measure. Nonetheless, numerical results of Ref. 74 indicate that there exists a dominant dependence of the volume of states with positive partial transposition (PPT) on the dimension d_{AB} for all nonsingular measures. In particular, it is shown that the volume of PPT decreases approximately exponentially with the dimension: $\mu(\text{PPT}) \propto K_\mu \exp(-t_\mu d_{AB})$ for some constants K_μ and t_μ depending on the chosen measure μ . From $\mu(\text{PPT}) = \mu(\text{SEP}) + \mu(\text{BE} \cap \text{PPT})$ it follows that also the volumes of separable and PPT entangled states decrease approximately exponentially with the dimension.

Finally, and exceeding the primary scope of this paper, we mention a result showing that bipartite bound entanglement occupies a finite volume only if the Hilbert space is finite-dimensional. Consider infinite-dimensional bipartite Hilbert spaces of square-integrable functions ($L^2(\mathbb{R})$), i.e. continuous-variable systems. In Ref. 71, the following result is shown:

Theorem 23. *The subset of undistillable states is nowhere dense in the set of all continuous-variable states in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ for $\mathcal{H}_A \cong \mathcal{H}_B \cong L^2(\mathbb{R})$.*

The theorem directly implies that quantum states are generically distillable in continuous-variable systems. Because the set of undistillable states does not contain any open ball, it is of measure zero independent of the chosen measure. As subsets of undistillable states, the sets SEP, PPT and BE are nowhere dense for infinite-dimensional systems. Note that this result is not limited to $\text{PPT} \cap \text{BE}$ but holds for possibly existing NPT bound entangled states as well.

5.2.2. Bell-diagonal states

In this section, results of Refs. 27–29 regarding the frequency of bound entangled states of Bell-diagonal systems (Sec. 2.3) are summarized. It is shown that a significant share of states is bound entangled for small dimensions of the subsystems ($d = d_A = d_B \geq 3$). Interestingly, the standard Bell-diagonal system, the Magic Simplex (17), stands out among all generalized Bell-diagonal systems (23). This indicates that the entanglement structure, including the frequency of bound entanglement, strongly depends on the choice of the Bell basis.

As discussed in the previous section, the choice of a measure to determine volumes within the set of quantum states is not unique since it depends on the designation of a “natural” distribution of states. Following a geometric approach to entanglement via the identification of d^2 -dimensional Bell-diagonal states with points in the standard $(d^2 - 1)$ -simplex, specific volumes of entanglement classes can be well

approximated by Monte Carlo simulations. More precisely, the normalized Lebesgue measure $\mathcal{L}_{d^2-1}$ can be approximated by the frequencies of a uniformly distributed sampling of points in the $(d^2 - 1)$ -simplex. As described in Ref. 27, the volumes (regarding the induced uniform measure) of certain subsets of Bell-diagonal states can be determined as follows. First, a sufficiently large set of uniformly distributed points in the simplex is generated and associated with states in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. These states are then classified as separable (SEP), PPT entangled (PPT-ENT), or NPT entangled (NPT) using the methods presented in Sec. 4. The relative frequencies approximate the volumes of these subsets in the state space. Note that due to the challenge of detecting bound entangled states, not all PPT states are classified as SEP or PPT-ENT.

In Ref. 27, this method is used to analyze bipartite standard Bell-diagonal qutrits ($d = 3$) for the standard Bell basis, i.e. in the Magic Simplex, and to determine the entanglement class of more than 95% of states. In Ref. 28, similar methods are applied for higher dimensions. As for general systems (cf. Sec. 5.2.1) the relative volume of PPT states decreases with the dimension (cf. Fig. 8). For the standard Bell basis, the relative volume of PPT decreases from 50% for $d = 2$ to less than 1% for $d \geq 6$. For $d = 3$ and $d = 4$, also a significant amount of PPT Bell-diagonal states can be detected as bound entangled. Including the uncertainty of PPT states that cannot be identified as separable or entangled, the relative volume of PPT-ENT lies in the interval $[5.4\%, 7.4\%]$ for qutrits and in $[0.2\%, 2.8\%]$ for ququarts. The results for the obtained frequencies of states in NPT, PPT, SEP and PPT-ENT are summarized in Table 3 and visualized in Fig. 8.

Reference 29 investigates generalized Bell-diagonal systems (cf. Sec. 2.3.2) and shows that the standard Bell basis has some very special properties that distinguish it from other generalized Bell bases. This is also reflected in the entanglement structure, particularly in the frequency of PPT bound entangled states among all generalized Bell-diagonal states. For thousands of generalized Bell bases for $d = 3$, the standard Bell-diagonal system, i.e. the Magic Simplex, has the smallest relative volume of NPT states among the corresponding Bell-diagonal states. In particular, the relative volumes of PPT and PPT-ENT are generally smaller for generalized

Table 3. Relative volumes of entanglement classes in the Magic Simplex. For the dimensions $d = 2, 3, 4, 5$, the relative frequencies of NPT, PPT, SEP and PPT-ENT states among all standard Bell-diagonal states are shown. For $d = 3$ and $d = 4$, a significant amount of PPT entangled states are detected, but not all PPT states can be classified as SEP or PPT-ENT and the relative volumes of SEP and PPT-ENT lies in the given interval.

	$d = 2$	$d = 3$	$d = 4$	$d = 5$
rel. NPT	0.50	0.61	0.884	0.93
rel. PPT	0.50	0.39	0.116	0.07
rel. SEP	0.50	0.33 ± 0.01	0.101 ± 0.013	—
rel. PPT-ENT	0.0	0.06 ± 0.01	0.015 ± 0.013	—

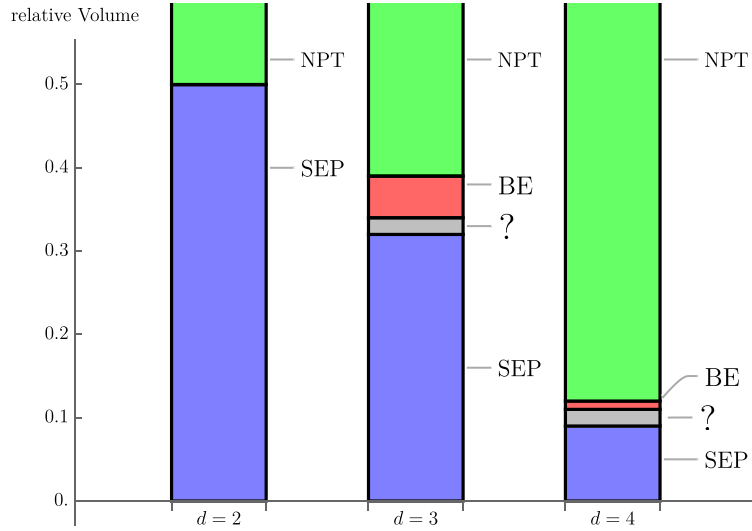


Fig. 8. The frequencies of PPT and NPT states for Bell-diagonal systems of dimension $d = 2, 3, 4$ are illustrated. The (not completely shown) green bar denotes the proportion of NPT states, while the blue, red and gray bars denote PPT states. The fraction of separable and bound entangled states is depicted in blue and red, respectively. Unclassified PPT states are denoted by “?” (gray). One observes a significant decrease in PPT states for increasing dimension.

Bell-diagonal states than for the Magic Simplex. The influence of the chosen Bell basis on the sets PPT, NPT, PPT-ENT and SEP can also be visualized by considering certain subsets of Bell-diagonal states. In Fig. 9, it is shown how the sets of the entanglement classes within a two-dimensional family of states significantly depend on the Bell basis.

Interestingly, when comparing different generalized Bell-diagonal systems, the relative frequency of PPT entangled states among all PPT states strongly correlates positively with the volume of PPT itself. This means that the frequency of PPT entangled states increases disproportionately with the PPT volume. The more PPT states exist, the higher the frequency of entangled states among them.

These observations indicate that the relations of basis states in a Bell basis strongly affect the entanglement structure of corresponding Bell-diagonal states, including the frequency of PPT bound entangled states. The underlying group structure, characterizing the standard Bell basis and the Magic Simplex via the Weyl relations (15), is connected with the frequency of PPT states and PPT entanglement for Bell-diagonal states. If generalized Bell bases are not related by such structure, then fewer Bell-diagonal PPT states and PPT entangled states can be observed, despite the fact that all states represent mixtures of maximally entangled Bell states.

5.3. Bound entanglement and information processing

Many quantum information protocols require some entangled states shared between two parties as a resource. Most of them operate best if the entanglement is

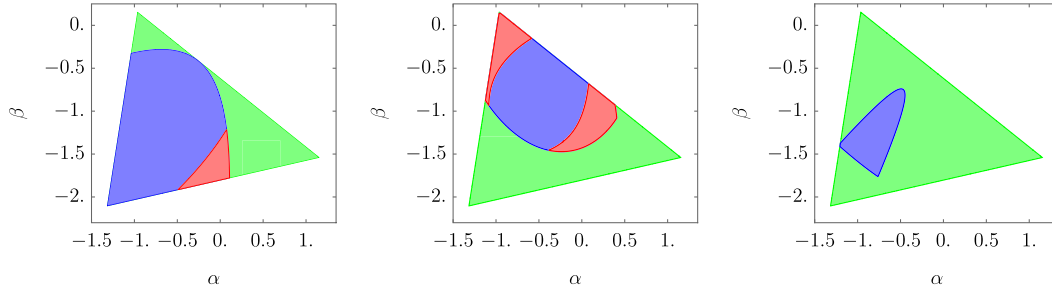


Fig. 9. These graphics illustrate a particular slice (state family B_1 ; see caption of Fig. 6 for the definition) through the Magic Simplex of bipartite qutrits for the standard Bell basis on the left. The middle and the right panel show the same parameterized family for two generalized Bell bases. The green area corresponds to NPT entangled states. The red area depicts PPT bound entangled states that are detected by the realignment criterion (Theorem 14). These three graphics illustrate the dependence of the entanglement structure of Bell-diagonal states on the Bell basis.

concentrated in maximally entangled pure states. However, the resource states in experiments are generally mixed, so the standard protocols do not apply unconditionally. In principle, if one has a close to unrestricted supply of free entangled states, any protocol can be preceded by entanglement distillation. In this case, the protocol for maximally entangled states is applicable after the distillation process. Consequently, all free entangled states can be utilized in specific quantum information processing tasks to outperform classical schemes. This strategy is, by definition, impossible for bound entangled states, yet they are still inseparable and thus display characteristics different from separable states. Therefore, the question arises whether they can serve as a resource for quantum information protocols. This section presents results that answer this question for most currently known applications in the negative.

5.3.1. Quantum teleportation and distillation

The most general setup of quantum teleportation is as follows.⁷⁵ Two parties, Alice and Bob, share a single resource state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d \geq 2$. Additionally, Alice is in possession of an unknown d -dimensional pure^e quantum state $|\phi\rangle \in \mathcal{H}_{A'}$. Teleportation aims to transmit $|\phi\rangle$ from Alice to Bob by some LOCC scheme, i.e. by only using local quantum operations and classical communication (cf. Sec. 2.2.3). Thus, teleportation through the resource state ρ_{AB} constitutes an LOCC-restricted quantum channel $\mathcal{E}_{\rho_{AB}} : \mathcal{D}(\mathcal{H}_{A'}) \rightarrow \mathcal{D}(\mathcal{H}_B)$. In the following, we denote the (generally mixed) teleported state by $\sigma_\phi = \mathcal{E}_{\rho_{AB}}(|\phi\rangle\langle\phi|) \in \mathcal{D}(\mathcal{H}_B)$, thereby omitting its explicit dependence on ρ_{AB} and the particular LOCC protocol.

^eTeleportation of an unknown mixed state $\sigma = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ is readily obtained by linearity of the protocol.

A customary figure of merit for a given teleportation scheme is the *fidelity*⁷⁶ between the teleported state σ_ϕ and the input state $|\phi\rangle$,

$$F(\sigma_\phi, |\phi\rangle) = \text{Tr} \left(\sqrt{(|\phi\rangle\langle\phi|)^{1/2} \sigma_\phi (|\phi\rangle\langle\phi|)^{1/2}} \right)^2 = \langle\phi| \sigma_\phi |\phi\rangle. \quad (67)$$

As the state $|\phi\rangle$ is unknown, it is necessary to integrate over all potential input states to obtain the averaged teleportation fidelity. We thus arrive at the *fidelity of transmission*⁷⁵

$$f(\rho_{AB}) := \int d\phi F(\sigma_\phi, |\phi\rangle), \quad (68)$$

where the integral is taken with respect to the Haar measure on the space of input states $|\phi\rangle$. Since $F(\sigma_\phi, |\phi\rangle) \in [0, 1]$, we also have $f(\rho_{AB}) \in [0, 1]$ for any ρ_{AB} and any teleportation procedure. The upper bound $f_{\max} = 1$ is achieved by the original teleportation protocol of Ref. 20 when a maximally entangled pure resource state is used. On the contrary, assuming a scenario with a separable resource state $\rho_{AB} = \rho_A \otimes \rho_B$, a nonzero bound of the teleportation fidelity can be obtained. In this case, one can apply a measure-and-prepare protocol⁷⁷ and obtain $f_{SEP} = 2/(d+1)$.^f

For what follows, we also need to define the *singlet fraction* of a state ρ_{AB} by

$$\mathcal{F}(\rho_{AB}) := F(\rho_{AB}, |\Omega_{0,0}\rangle) = \text{Tr}(P_{0,0} \rho_{AB}), \quad (69)$$

where $|\Omega_{0,0}\rangle$ is a maximally entangled state^g and $P_{0,0}$ is its projector, defined in (4) and (16), respectively. Note that distillation aims to maximize the singlet fraction of the initial state by possibly using up multiple additional copies of ρ_{AB} . A state is distillable if $\mathcal{F} = 1$ is reached with nonzero probability after some number of LOCC rounds (cf. Sec. 2.2.3).

Reference 75 establishes an important connection between the singlet fraction of a state and the fidelity of transmission.

Theorem 24. *For a given resource state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$, the maximal fidelity of transmission $f_{\max}$ achievable by any teleportation protocol is given by*

$$f_{\max} = \frac{d \mathcal{F}_{\max} + 1}{d + 1}. \quad (70)$$

Here, $\mathcal{F}_{\max}$ denotes the maximal singlet fraction attainable by LOCC from a single copy of ρ_{AB} .

Applied to bound entangled states, we get the following corollary.

^fThis can be shown by direct calculation (see Appendix A). Alternatively, it follows from Theorem 24. For this, note that there are separable states σ with $\mathcal{F}_{\max} = 1/d$, e.g. any subgroup state (19) in the Magic Simplex. For such a state, (70) applies and yields the desired bound.

^gInconsequentially, for dimension $d = 2$, $|\Omega_{0,0}\rangle$ is not the singlet state but can be transformed into the true singlet by local operations.

Corollary 1. *Bound entangled resource states do not permit better-than-classical quantum teleportation since $\mathcal{F}_{\max} \leq 1/d$.*

Otherwise, they would be distillable by the protocol of Ref. 16, contradicting the definition of bound entanglement. As a consequence, the maximal fidelity of transmission for bound entangled states satisfies $f_{\max}^{\text{BE}} \leq 2/(d+1)$. This is exactly the bound of the measure-and-prepare protocol mentioned above. Thus, bound entangled states do not allow for better transmission fidelity than any purely classical channel. Hence, the best teleportation strategy for Alice and Bob sharing a bound entangled state is to discard it and perform classical teleportation.

Finally, Ref. 78 recasts quantum teleportation into a state discrimination problem. Doing so allows them to characterize the resource states that offer no advantage over separable resource states in one-way teleportation protocols. Their novel approach allows for an alternative proof of Corollary 1.

5.3.2. Dense coding

The concept of *dense coding*²¹ (also called super-dense coding) is intimately connected to teleportation. Indeed, in a certain sense, the two problems are dual to one another.^{17,78} Dense coding aims to transmit a certain number of bits (i.e. classical information) from Alice to Bob by exchanging fewer qudits. We assume a noiseless quantum channel through which the qudits are sent, whereas the shared states are generally mixed and noisy. The procedure starts with Alice and Bob sharing n identical bipartite quantum states $\rho_{AB}^{\otimes n}$. Alice can then perform any local quantum operation on her part of the shared states to encode some classical bits and then send her altered n qudits to Bob. Now Bob tries to retrieve $m > n$ classical bits of information encoded by Alice.

The problem as presented above can be restated in terms of the capacity of the (noisy) entanglement-assisted classical channel.^{79–81} A channel capacity greater than 1 indicates that more than one bit of classical information is transmitted for every qudit sent through the noiseless quantum channel. A quantum channel's maximum capacity is $2 \log_2(d)$, where d is the dimension of the qudit (see Ref. 17). One of the results of Ref. 81 is

Theorem 25. *Bound entanglement-assisted classical channels do not admit a capacity greater than 1.*

Consequently, they do not outperform a fully classical protocol, i.e. one where the bits are sent via a classical channel from Alice to Bob.

A different approach to dense coding with bound entangled states is presented in Ref. 78. Therein, the authors reexamine the duality between dense coding and teleportation. As a consequence of Corollary 1, they likewise prove that bound entanglement is not useful in general dense coding schemes beyond the classical level.

5.3.3. Activation of bound entanglement

In Refs. 82 and 83, it is proven that for generic mixed states, it is impossible to distill a maximally entangled state with nonzero probability from only a single copy. This means that no LOCC operations applied to a generic bipartite mixed state (of equal dimensions $d_A = d_B$) can bring the singlet fraction (69) arbitrarily close to 1 with nonvanishing probability. Consequently, by Theorem 24, it is generally impossible to teleport faithfully using only a single entangled mixed state, even if it is free entangled.

An example^h of such a free entangled qutrit-qutrit state ρ^{FE} is given in Ref. 75. Its initial singlet fraction is $\mathcal{F}_0$, and this cannot be increased beyond some $\mathcal{F}_{\text{max}} < 1$ by LOCC and using only a single copy of the state. They further show that this upper bound even holds for LOCC protocols that have a probability of success converging to zero. According to Theorem 24 the fidelity of teleportation through ρ^{FE} as the sole resource state is thus upper bounded by $f_{\text{max}} = \frac{1}{4}(3\mathcal{F}_{\text{max}} + 1) < 1$. Hence, faithful teleportation through ρ^{FE} is not guaranteed. Surprisingly, an additional supply of bound entangled states can considerably improve teleportation fidelity. This is unexpected because bound entangled states can neither be distilled nor allow faithful teleportation on their own (cf. Corollary 1).

A concrete LOCC protocol that achieves this is presented in Ref. 84. It consumes a suitable bound entangled state ρ^{BE} and transforms ρ^{FE} to ρ_1^{FE} with singlet fraction $\mathcal{F}_1 > \mathcal{F}_0$. Heuristically speaking, the entanglement from the bound entangled state is transferred to the free entangled pair. As the states ρ^{FE} and ρ_1^{FE} are structurally identical, the protocol can be repeated. After n repetitions, the singlet fraction of ρ_n^{FE} is $\mathcal{F}_n$ which approaches 1 as n goes to infinity. Importantly, for some $n = n_0$, one eventually obtains $\mathcal{F}_{n_0} > \mathcal{F}_{\text{max}}$. Therefore, the teleportation fidelity of $\rho_{n_0}^{\text{FE}}$ exceeds f_{max} and ultimately approaches 1 for large n_0 , thus making faithful teleportation possible.

The recent work in Ref. 85 investigates a different protocol of bound entanglement activation compared to the one in Ref. 84. The authors utilize quantum Zeno dynamics to similarly increase the singlet fraction (69) of a free entangled state by consuming one or more bound entangled states.

In scenarios like this, where bound entangled states improve the performance of some quantum information theoretic task that requires free entanglement, we say that the bound entanglement has been *activated*. On the other hand, some authors interpret this the other way around: In the above example, they would say that the teleportation capability of the state ρ^{FE} has been activated by the bound entanglement in ρ^{BE} .

A different kind of activation of bound entanglement is presented in Ref. 86. The authors show that bound entanglement can more generally activate a quantum state's distillability. In particular, bound entangled states can make any NPT state

^hThis example and the specific activation protocol of Ref. 84 mentioned below are reproduced in Appendix B.

1-distillable (cf. Sec. 2.2.3). To show this, they use bound entangled states constructed from two highly symmetric families of states, Werner and isotropic states. It is worth noting that the resulting bound entangled activator states have full rank. Moreover, some of them are universal activators, i.e. they can activate the distillability of any given NPT state. Some bound entangled universal activator states lie arbitrarily close to the set of separable states.

5.3.4. Pure state conversion assisted by bound entanglement

A similar effect to the activation of bound entanglement can be achieved for pure state conversion. In particular, it is known that if only one copy of a pure bipartite quantum state is available, its Schmidt rank cannot be increased by any (probabilistic) LOCC scheme.^{87,88} However, when additional specific bound entangled states are available as a resource, this restriction can be circumvented.⁸⁹

Theorem 26. *A single copy of any bipartite entangled pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ can be converted with nonzero success probability to any other pure bipartite state $|\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ by LOCC with the help of suitable bound entangled states.*

The crucial part of the conversion protocol involves the state

$$|\Omega_{0,0}^m\rangle = \frac{1}{\sqrt{m}} \sum_{i=1}^m |i\rangle \otimes |i\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d, \quad (71)$$

where $2 \leq m \leq d$. It is shown that this state can be probabilistically converted to the d -dimensional maximally entangled state $|\Omega_{0,0}^d\rangle$ by LOCC with the assistance of the (unnormalized) bound entangled state

$$\begin{aligned} \sigma_{m \rightarrow d} = & \frac{1}{m^2(d-1)} \left((m-1)(P_{0,0}^d)_{A_1 B_1} \otimes (P_{0,0}^m)_{A_2 B_2} \right. \\ & \left. + \frac{1}{d+1} (\mathbb{1}_{d^2} - P_{0,0}^d)_{A_1 B_1} \otimes (\mathbb{1}_{m^2} - P_{0,0}^m)_{A_2 B_2} \right). \end{aligned} \quad (72)$$

Here, $P_{0,0}^d$ and $P_{0,0}^m$ are the projectors onto $|\Omega_{0,0}^d\rangle$ and $|\Omega_{0,0}^m\rangle$, respectively. The subscripts A_1 , A_2 , B_1 and B_2 indicate the corresponding Hilbert spaces on which the state acts. This state is indeed PPT in the cut $A_1 A_2 | B_1 B_2$. Furthermore, as the conversion of $|\Omega_{0,0}^m\rangle$ to $|\Omega_{0,0}^d\rangle$ is not possible for $d > m$ by LOCC alone, $\sigma_{m \rightarrow d}$ must be entangled. Hence, it is bound entangled.

The explicit protocol of probabilistically converting any entangled state $|\psi\rangle$ to any other state $|\phi\rangle$ by LOCC (and potentially using bound entangled states) is as follows.⁸⁹ Suppose the initial state $|\psi\rangle$ is entangled with Schmidt rank $r > 1$. If the target state $|\phi\rangle$ has Schmidt rank $r' \leq r$, then the usual conversion techniques^{87,88} are applicable. However, if $r' > r$, the initial state $|\psi\rangle$ is first converted to $|\Omega_{0,0}^r\rangle$ by LOCC.⁸⁸ Next, $|\Omega_{0,0}^r\rangle$ is converted into $|\Omega_{0,0}^{r'}\rangle$ by LOCC and the use of $\sigma_{r \rightarrow r'}$ as in (72). Finally, $|\Omega_{0,0}^{r'}\rangle$ can be converted to any desired state $|\phi\rangle$ with Schmidt rank r' . Importantly, the probability of success is nonzero in either case. Consequently, bound entangled states can be useful in pure state conversion tasks.

5.3.5. Bound entangled states and quantum channels

There are various ways of identifying quantum states with quantum channels. Two natural ones have been investigated for bipartite bound entangled states,^{68,90} and both relate to the concept of binding entanglement channels:

Definition 20 (Binding entanglement channel). A quantum channel $\mathcal{E}$ is called a binding entanglement channel if it can be used to create bipartite bound entangled states between two parties, but it cannot be used to share free entangled states.

By definition, binding entanglement channels are never entanglement breaking.⁹¹ The authors of Ref. 90 characterized binding entanglement channels by proving

Theorem 27. *A quantum channel $\mathcal{E}$ is a binding entanglement channel if and only if its Choi state (6) is a bound entangled state.*

Thus, utilizing the Choi–Jamiołkowski isomorphism,^{30,31} it is possible to unambiguously associate a binding entanglement channel $\mathcal{E}$ with a bound entangled state ρ whose reduced state is maximally mixed, i.e. $\rho_A = \text{Tr}_B(\rho) = (1/d)\mathbb{1}_d$. In particular, the identification is given by

$$\rho = (\mathcal{E} \otimes \mathbb{1})P_{0,0}, \quad (73)$$

where $P_{0,0}$ is the Bell projector defined in (16). The original bound entangled state is thus obtained by sending one half of a maximally entangled state $|\Omega_{0,0}\rangle$ as in (4) through the corresponding channel. As any Bell-diagonal state is locally maximally mixed (cf. Sec. 2.3), Eq. (73) fully characterizes any bound entangled state within this subset of states. For bound entangled states with reduced states not maximally mixed (but with full rank), it is still possible to find such an identification. However, an additional preliminary step of local filtering is necessary in this case.⁹⁰

An alternative route from bound entangled states to binding entanglement channels is presented in Ref. 68. There, the binding entanglement channel is given by the teleportation through the bound entangled state (cf. Sec. 5.3.1). No free entanglement can be shared by utilizing this channel. Otherwise, one could use it to construct a distillation scheme for the initial bound entangled state. Furthermore, sending one half of the maximally entangled state $|\Omega_{0,0}\rangle$ through the channel, the resulting state is bound entangled. Consequently, the channel satisfies Definition 20 and is thus a binding entanglement channel.

5.3.6. Quantum key distribution

The main objective of quantum key distribution (originally devised in Ref. 92) is to ensure secure classical communication between two parties, Alice and Bob. This is achieved by using “effectively distributed” entanglement⁹³ to establish a private key, i.e. a classical bit string known exclusively to them. In this creation process,

any potential interference of a third party, Eve, is detectable by Alice and Bob with nonzero probability. Alice and Bob can subsequently use the shared key to encode classical data before transmission and decode it upon receipt.

A common measure of a state's capability of allowing for a secure key generation protocol is given by:⁹⁴

Definition 21 (Private key distillation rate). The private key distillation rate (or simply distillable key) K_D is the maximally attainable asymptotic rate of converting n copies of some initial quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ via LOCC operations into a private qudit state of dimension d_n of the form

$$\gamma_{ABA'B'}^{(d_n)} = \frac{1}{d_n} \sum_{i,j=0}^{d_n-1} |ii\rangle\langle jj|_{AB} \otimes U_i \tilde{\rho}_{A'B'} U_j^\dagger, \quad (74)$$

where $\tilde{\rho}_{A'B'} \in \mathcal{D}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'})$ and the U_i are some unitaries acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$. In particular, the (device-dependent) distillable key is thus given by

$$K_D = \sup_{\mathcal{P}} \limsup_{n, d_n \rightarrow \infty} \frac{\log_2(d_n)}{n}, \quad (75)$$

where the supremum is taken over all possible LOCC protocols $\mathcal{P}$ converting $\rho_{AB}^{\otimes n}$ into $\gamma_{ABA'B'}^{(d_n)}$.

The name *private state* for states of the form (74) comes from Ref. 95 where it is shown that the most general state from which a private key can be obtained by LOCC is of this form. The private key distillation rate is very similar to the asymptotic entanglement distillation rate of Sec. 2.2.3, with the main difference being the different target state $\gamma_{ABA'B'}$.^{95,96} In general, distilling a private state is an easier task than distilling a maximally entangled state.²

For some quantum key distribution protocols, entangled states shared between the parties are a valuable resource. It is shown in Ref. 95 that an arbitrarily secure private key can also be obtained from bound entangled states. To this end, a specific bound entangled state with $K_D > 0$ is constructed. The downside of this constructive proof is that it requires many resource states on a high-dimensional Hilbert space. As demonstrated in Ref. 97, this high-dimensionality is not generally necessary. Therein, bound entangled states with dimensions $d_A = d_B = 4$ are constructed from which a private key can be obtained via the one-way distillation protocol from Ref. 98. This example is generalized in Ref. 99 to a wider class of states (requiring a two-way protocol for secret key distillation). Ref. 100 shows that the possibility to distill a secure key from certain PPT-entangled states relates to how much they violate a Bell inequality (c.f. Sec.5.6. A different construction of bound entangled states with a positive distillable key is presented in Ref. 101. This establishes a link between bound entangled states and classical probability theory. An important consequence of these examples is that a zero entanglement distillation rate, a feature of all bound entangled states, does not imply a zero private key distillation rate.

The authors of Ref. 97 argue that the volume of bound entangled states $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with positive distillable key, $K_D > 0$, is nonzero in the case of finite Hilbert space dimensions $d_A, d_B \geq 4$. This makes the experimental implementation of their theoretical considerations more feasible because a nonzero volume implies, in some sense, robustness against small perturbations (cf. Sec. 6). Furthermore, in Ref. 102, so-called “privacy witnesses” are introduced. They are an analog of entanglement witnesses (cf. Sec. 4.1) but for detecting a positive distillable key of quantum states. The advantage of these privacy witnesses is that they can facilitate the experimental verification of (bound entangled) private states compared to other methods.

The nonzero distillable key of bound entangled states might also be relevant for quantum repeaters.¹⁰³ Commonly, such repeaters utilize entanglement distillation and teleportation, making them inapplicable in scenarios with bound entanglement. An alternative realization is proposed in Ref. 104. This does not require distillability but only a nonzero distillable key and can, therefore, be used with bound entangled states.

Interestingly, the impossibility of using bound entangled states for distillation-based quantum repeaters can be turned into an advantage. Reference 105 proposes a hybrid quantum network architecture based on bound entangled states. While allowing the encryption of classical communication between two authorized parties, the impossibility of distilling such states protects against unauthorized secret-key generation via entanglement swapping. As distillable states used in such a network do not provide this security, bound entanglement can represent a unique resource for information processing tasks.

A further focus is on device-independent quantum key generation.¹⁰⁶ In this context, the authors of Ref. 107 found PPT-entangled states such that the device-independent distillable key K_{DI} is strictly less than the device-dependent one in (75), i.e. $K_{DI} < K_D$ for certain shared states. Regarding bound entanglement in device-independent protocols, a revised Peres conjectureⁱ is brought forward in Ref. 108:

Conjecture 1. *Bound entangled states are not useful for device-independent quantum key distribution.*

5.4. Bound entanglement and entanglement measures

A natural extension to whether a quantum state is entangled or not is to ask how much the state is entangled. *Entanglement measures* (cf. Ref. 109) aim to quantify the amount of entanglement. However, up to this date, no universal measure has been found. Instead, several entanglement measures have been proposed with different operational meanings. For mixed states, these measures are generally hard to

ⁱFor the original Peres conjecture, see the discussion in Sec. 5.6.

compute. However, for pure bipartite states the *entropy of entanglement* (cf. Ref. 15 and defined below), based on the *von Neumann entropy* provides a unique measure of entanglement (cf. Ref. 110). All proposed bipartite entanglement measures should reduce to it for pure states. Depending on the context, other properties like nonincreasing on average under LOCC, (sub- or super-) additivity or continuity can be required or desired.

In this section, we shortly introduce the most relevant entanglement measures for bipartite mixed states and elaborate on their relation to bound entanglement. Of special interest is the appearance of irreversibility of certain quantum information processing tasks. The difference between two important proposed entanglement measures, namely the so-called *entanglement cost* and the *distillable entanglement*, implies an irreversible loss of entanglement as a resource for some protocols. This phenomenon is closely related to bound entanglement. Another relation to bound entanglement is reflected in the additivity of the distillable entanglement and the existence of NPT bound entangled states.

The bipartite *entropy of entanglement* E of a pure state is defined as the *von Neumann entropy* $S(\rho) := -\text{Tr}(\rho \log_2(\rho))$ of the reduced state.

Definition 22 (Entropy of entanglement). Let $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The (von Neumann) entropy of entanglement of $|\psi\rangle$ is defined as

$$E(|\psi\rangle\langle\psi|) := S(\text{Tr}_A(|\psi\rangle\langle\psi|)) = S(\text{Tr}_B(|\psi\rangle\langle\psi|)). \quad (76)$$

A unique or “canonic” extension to mixed states is not known. One way to generalize it to mixed states is by so-called “convex hull constructions” based on all possible decompositions of ρ into pure states $|\psi_i\rangle$ with probabilities p_i . The *entanglement of formation* E_F is such an extension.

Definition 23 (Entanglement of formation). Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Given all decompositions of ρ into pure states $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, the entanglement of formation is defined as

$$E_F(\rho) := \inf_{\{p_i, \psi_i\}} \left\{ \sum_i p_i E(|\psi_i\rangle\langle\psi_i|) \right\}, \quad (77)$$

where the infimum is taken over all possible decompositions.

The entanglement of formation measures the minimum number of maximally entangled qubits required on average to produce the given state via LOCC.¹⁵ The authors of Ref. 2 show that this measure is nonadditive, i.e. in general, $E_F(\rho_1 \otimes \rho_2) \neq E_F(\rho_1) + E_F(\rho_2)$.

A closely related measure is the *entanglement cost* E_C . It is defined in the asymptotic limit of infinitely many copies and quantifies the maximal rate r at which maximally entangled states can be converted to copies of ρ using LOCC. The conversion is allowed to be approximate with vanishing error in the limit of a large number of copies. Let $P_{0,0}^{\otimes m}$ be m -copies of the bipartite maximally entangled

state of dimension d^2 (16). Denoting general LOCC operations by $\mathcal{E}_{\text{LOCC}}$ and the trace norm as $\text{Tr} |\rho| := \sqrt{\rho^\dagger \rho}$, we can define the entanglement cost as follows.

Definition 24 (Entanglement cost). Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The entanglement cost is defined as

$$E_C(\rho) := \inf \left\{ r : \lim_{n \rightarrow \infty} \left[\inf_{\mathcal{E}_{\text{LOCC}}} \text{Tr} |\rho^{\otimes n} - \mathcal{E}_{\text{LOCC}}(P_{0,0}^{\otimes rn})| \right] = 0 \right\}. \quad (78)$$

The close relation of the entanglement cost and the entanglement of formation is reflected in the equality of the *regularized* entanglement of formation and the entanglement cost (cf. Ref. 111),

$$E_C(\rho) = \lim_{n \rightarrow \infty} \frac{E_F(\rho^{\otimes n})}{n}. \quad (79)$$

The *distillable entanglement* E_D quantifies entanglement from the opposite viewpoint compared to the entanglement cost. While E_C represents the rate for transforming copies of a maximally entangled state to copies of the state ρ , E_D represents the rate for the inverse process of transforming multiple copies of ρ to maximally entangled states, i.e. entanglement distillation (cf. Sec. 2.2.3):

Definition 25 (Distillable entanglement). Let $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The distillable entanglement is defined as

$$E_D(\rho) := \sup \left\{ r : \lim_{n \rightarrow \infty} \left[\inf_{\mathcal{E}_{\text{LOCC}}} \text{Tr} |\mathcal{E}_{\text{LOCC}}(\rho^{\otimes n}) - P_{0,0}^{\otimes rn}| \right] = 0 \right\}. \quad (80)$$

Note that in the definitions of E_C and E_D , often $d = 2$ is assumed for the Bell state $P_{0,0}$. As discussed in Sec. 2.2.3, for $d > 2$ the asymptotic rates change according to $r \rightarrow r / \lceil \log_2(d) \rceil$.

Another important measure is the *relative entropy of entanglement* E_R . Using the *quantum relative entropy* $S(\rho || \sigma) := \text{Tr}(\rho \log_2(\rho) - \rho \log_2(\sigma))$, the relative entropy of entanglement can be defined as follows.

Definition 26 (Relative entropy of entanglement). Let $X \subset \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a set that is mapped onto itself under LOCC. The relative entropy of entanglement of ρ with respect to X is

$$E_{R,X}(\rho) := \inf_{\sigma \in X} S(\rho || \sigma). \quad (81)$$

The regularized relative entropy of entanglement is

$$E_{R,X}^\infty(\rho) := \lim_{n \rightarrow \infty} \frac{E_{R,X}(\rho^{\otimes n})}{n}. \quad (82)$$

If $X = \text{SEP}$, we simply write E_R . However, depending on which states are considered as “resource-neutral”, i.e. states for which zero measure is specified, also the sets PPT or the set of undistillable states can be considered as the set X .

Many other proposed entanglement measures exist (see Ref. 109 for an introduction to the topic), including the already introduced *distillable key* K_D (cf. Eq. (75)) or a measure related to the concurrence (cf. Ref. 112) and its quasi-pure approximation of Sec. 4.6.

5.4.1. Irreversibility and PPT bound entanglement

From the definitions of the entanglement cost and the distillable entanglement, it becomes clear that if they are equal, the processes of entanglement distillation using multiple copies of a state ρ is asymptotically reversible because ρ can be created from distilled maximally entangled states via LOCC at the same rate. However, in Refs. 68 and 113, it is shown that this does not generally hold and that the process of entanglement distillation is irreversible.

In particular, it is shown for a specific PPT bound entangled state ρ_b with $E_D(\rho_b) = 0$, that $E_C(\rho_b) > 0$. This is generalized in Ref. 114 to all bound entangled states. Consequently, interconverting the mixed-state entanglement of ρ and the entanglement of the pure Bell states via asymptotic LOCC operations is not reversible.

This also holds for catalytic LOCC operations of the form $\rho_b^{\otimes n} \otimes P_{0,0}^{\otimes m} \xrightarrow{\text{LOCC}} P_{0,0}^{\otimes(m+l)}$ (cf. Ref. 113), even if correlations of the catalyst with the systems are allowed at the output.¹¹⁵ This means that, even in the asymptotic limit, the entanglement needed to locally create a PPT bound entangled state ρ_b cannot be distilled using (catalytic) LOCC. Interpreting entanglement as a resource for information processing tasks, bound entangled states bind resources that cannot be accessed by distillation and thus motivate their name.

5.4.2. Additivity of distillable entanglement and NPT bound entanglement

The potential existence of NPT bound entanglement is linked to the additivity of the distillable entanglement. As mentioned in Sec. 5.3.3, with the help of certain bound entangled states, any NPT entangled state, including a potential NPT bound entangled states σ_b , can be activated and made distillable by local operations (cf. Refs. 86 and 84). Since any universal activator bound entangled state ρ_b and any NPT bound entangled state σ_b have by definition zero distillable entanglement, the distillability of the activated state would imply $E_D(\rho_b \otimes \sigma_b) > E_D(\rho_b) + E_D(\sigma_b) = 0$ and consequently the nonadditivity of the distillable entanglement.

5.4.3. Inequalities for entanglement measures

In Sec. 5.3.6, it is discussed that there exist bound entangled states ρ_b for which the distillable key (75) is nonzero, i.e. $K_D(\rho_b) > 0 = E_D(\rho_b)$. In Ref. 95, it is shown that the following hierarchy between the entanglement measures exists:

$$E_D \leq K_D \leq E_R^\infty \leq E_C. \quad (83)$$

While this inequality holds for all states, also the strict inequalities $E_D < K_D < E_C$ and $E_D < E_R^\infty$ are possible for certain states. Bound entangled states (with $E_D = 0$) can thus exhibit nonzero K_D , E_R^∞ and E_C .

5.5. Bound entanglement and quantum metrology

The field of quantum metrology studies precise measurement techniques with the aim of extracting physical parameters from quantum systems. One of the tasks is to estimate the dynamical parameter θ of a unitary evolution $U(\theta) = e^{-i\theta H}$ acting on a given quantum state ρ (see, e.g. Refs. 116 and 117). This evolution is induced by a Hamiltonian H . For any measurement, the estimation sensitivity is bounded by the *quantum Cramér-Rao bound*,

$$(\Delta\theta)^2 \geq \frac{1}{F_Q(\rho, H)}, \quad (84)$$

which involves the inverse of the *quantum Fisher information* (QFI). The latter is thus often the central object of study, and for the case under consideration, it is given by

$$F_Q(\rho, H) = 2 \sum_{i,j=0}^{d-1} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle \psi_i | H | \psi_j \rangle|^2. \quad (85)$$

Here, the $|\psi_i\rangle$ are eigenvectors of $\rho \in \mathcal{D}(\mathcal{H})$, $\dim(\mathcal{H}) = d$, with corresponding eigenvalue λ_i , and the sum is only over terms with $\lambda_i + \lambda_j \neq 0$. One of the results in Ref. 118 states that the QFI is upper bounded for any state ρ by

$$F_Q(\rho, H) \leq F_{Q,\max}(\rho, H) = 4(\Delta_\rho H)^2, \quad (86)$$

where $(\Delta_\rho H)^2 = \langle H^2 \rangle_\rho - \langle H \rangle_\rho^2$ is the variance of H . Furthermore, there exists an upper bound of the QFI for separable states, obtained by maximizing (85) over all states $\rho \in \text{SEP}$.¹¹⁹ This is generally lower than the maximal upper bound given by (86). A quantum state ρ is said to be *useful for quantum metrology* (for a given Hamiltonian H) if the QFI $F_Q(\rho, H)$ exceeds the separable bound.

Some investigations were conducted to find an answer to the question of whether bipartite bound entangled states can be useful for quantum metrology. The first examples are presented in Ref. 120. Semidefinite programming is used to find numerical instances of such states, including ones in Hilbert space dimensions 2×4 and $d \times d$ for $3 \leq d \leq 12$. In addition to the numerical findings, the authors also give the analytical form of an example in dimension 4×4 .

Subsequently, starting from these states, as well as from ones in Ref. 121, the authors of Ref. 122 constructed two distinct families of $(2d \times 2d)$ -dimensional bound entangled states for any $d \geq 2$ which are metrologically useful. On top of that, they even show that for their specific Hamiltonian, the general upper bound (86) is satisfied. Hence, these states are maximally useful for parameter estimation with a specific Hamiltonian.

In conclusion, even though bound entangled states can sometimes be useful for quantum metrology, generally, maximally entangled pure states offer a greater metrological advantage.

5.6. Bound entanglement and Bell inequalities

The origin of today's quantum information theory can be dated back to the publication of John Stewart Bell in 1964,¹²³ where he showed that the predictions of quantum theory of two joint quantum systems are in general incompatible with a local realistic hidden variable theory. Or in the words of Bell:¹²⁴

“If [a hidden-variable theory] is local, it will not agree with quantum mechanics, and if it agrees with quantum mechanics, it will not be local.”

In mathematical terms, any local realistic model assumes that the measurement statistics of a joint system are reproduced by

$$P(A, a; B, b) = \int d\lambda \rho(\lambda) p(A, a, \lambda) \cdot p(B, b, \lambda), \quad (87)$$

where A, B are observables chosen independently by Alice and Bob and a, b are their measurement outcomes. Here, $\rho(\lambda)$ is any normalized distribution of hidden parameters λ . The factorization of the joint probability distribution of A and B into the individual probabilities p is the assumption denoted as *local* by Bell.

Combining different probabilities of the form (87) allows us to find upper and lower bounds on those combinations, called Bell inequalities. Those discriminate between local realistic models and other models, particularly quantum theory. A plethora of experiments in many distinct quantum systems show a violation of such Bell inequalities and herewith, a contradiction to the assumptions behind Bell inequalities. A violation of a Bell inequality is only found in the case the joint system is entangled, but not all entangled systems violate a particular Bell inequality.

Here we present the result of Ref. 125, where the authors found a particular bound entangled state that does violate a particular Bell inequality. Consequently, the Peres conjecture¹²⁶ — that bound entangled states are equal to states that exhibit a local realistic explanation — is falsified. An alternative formulation of the conjecture is that any entangled state that leads to Bell inequality violation must be free entangled.

Let us consider the qutrit–qutrit state

$$\rho_{\text{bound}} = \begin{pmatrix} a_1 & 0 & 0 & 0 & a_2 & a_3 & a_4 & 0 & 0 \\ 0 & a_5 & a_3 & a_2 & 0 & 0 & 0 & -a_4 & -a_6 \\ 0 & a_3 & a_7 & a_3 & 0 & 0 & 0 & -a_6 & 0 \\ 0 & a_2 & a_3 & a_5 & 0 & 0 & 0 & -a_4 & a_8 \\ a_2 & 0 & 0 & 0 & a_1 & -a_3 & -a_4 & 0 & 0 \\ a_3 & 0 & 0 & 0 & -a_3 & a_7 & a_8 & 0 & 0 \\ a_4 & 0 & 0 & 0 & -a_4 & a_8 & a_9 & 0 & 0 \\ 0 & -a_4 & -a_6 & -a_4 & 0 & 0 & 0 & a_9 & 0 \\ 0 & -a_6 & 0 & a_8 & 0 & 0 & 0 & 0 & a_8 \end{pmatrix}, \quad (88)$$

where $a_1 = \frac{9789}{27536}$, $a_2 = \frac{3239}{27536}$, $a_3 = \frac{5\sqrt{\frac{131}{2}}}{13768}$, $a_4 = \frac{45\sqrt{\frac{131}{2}}}{6884}$, $a_5 = \frac{3311}{27536}$, $a_6 = \frac{9}{6884}$, $a_7 = \frac{1}{13768}$, $a_8 = \frac{9}{6884}$ and $a_9 = \frac{81}{3442}$. The state is bound entangled, which is detected by the realignment criterion (see Sec. 4.4). The following observable, in the form of an entanglement witness (see Sec. 4.1), can be defined:

$$W_{Bell} = \begin{pmatrix} \frac{1}{25} & 0 & 0 & 0 & -\frac{1}{25} & \frac{\sqrt{2}}{25} & -\frac{\sqrt{21}}{25} & 0 & 0 \\ 0 & \frac{3}{25} & \frac{\sqrt{2}}{25} & -\frac{1}{25} & 0 & 0 & 0 & \frac{\sqrt{21}}{25} & \frac{\sqrt{42}}{25} \\ 0 & \frac{\sqrt{2}}{25} & \frac{46}{25} & \frac{\sqrt{2}}{25} & 0 & 0 & 0 & \frac{\sqrt{42}}{25} & 0 \\ 0 & -\frac{1}{25} & \frac{\sqrt{2}}{25} & \frac{3}{25} & 0 & 0 & 0 & \frac{\sqrt{21}}{25} & -\frac{\sqrt{42}}{25} \\ -\frac{1}{25} & 0 & 0 & 0 & \frac{1}{25} & -\frac{\sqrt{2}}{25} & \frac{\sqrt{21}}{25} & 0 & 0 \\ \frac{\sqrt{2}}{25} & 0 & 0 & 0 & -\frac{\sqrt{2}}{25} & \frac{46}{25} & -\frac{\sqrt{42}}{25} & 0 & 0 \\ -\frac{\sqrt{21}}{25} & 0 & 0 & 0 & \frac{\sqrt{21}}{25} & -\frac{\sqrt{42}}{25} & \frac{21}{25} & 0 & 0 \\ 0 & \frac{\sqrt{21}}{25} & \frac{\sqrt{42}}{25} & \frac{\sqrt{21}}{25} & 0 & 0 & 0 & \frac{21}{25} & 0 \\ 0 & \frac{\sqrt{42}}{25} & 0 & -\frac{\sqrt{42}}{25} & 0 & 0 & 0 & 0 & \frac{8}{25} \end{pmatrix}. \quad (89)$$

This satisfies for all local realistic hidden-variable theories¹²⁵

$$0 \leq \text{Tr}(W_{Bell} \rho). \quad (90)$$

Computing $\text{Tr}(W_{Bell} \rho_{\text{bound}}) = -0.000263144$ shows a violation and thus the Peres conjecture is falsified. Undistillable states contradict the assumptions behind the derivation of Bell inequalities (the extension to the dimension $d = 4, \dots, 8$ is presented in Ref. 127). In Ref. 100, the question is asked how much one can violate a Bell inequality with PPT-entangled states and it is concluded that a violation cannot be large with respect to different measures, which affects quantum key generation (see Sec. 5.3.6).

Researchers have tried to find different families of bound entangled states that violate other Bell inequalities. For example, the systematic set of Bell inequalities for qudit systems,¹²⁸ so-called CGLMP-Bell inequalities, have been discussed for a subset of families within the Magic Simplex¹²⁹ and found a negative answer, i.e. no bound entangled state violated these Bell inequalities. It remains open whether a general set of Bell inequalities and bound entangled states can be constructed in a systematic way. With that, the relation of undistillability and assumptions behind Bell inequalities, such as nonlocality, can be further enlightened.

5.7. Bound entanglement and Lorentz invariance

One of the biggest open questions in physics is whether the two great theories, relativity and quantum theory, can be unified. Entanglement and quantum information theoretic tasks have therefore been intensively discussed in a relativistic setting (see, e.g. Ref. 130 and references therein). In Ref. 131, the authors put forward the question of whether bound entangled states differ from free entangled states by being invariant under a Lorentz boost. They found a negative answer, i.e. boosted bound entangled states can become separable, bound, or free entangled states depending on the applied relativistic boost. We sketch their derivation of this fact in the remainder of this section.

A strong guiding principle is that physics should not change for different observers. Let us choose as an observable an entanglement witness W (see Sec. 4.1), then the following relation has to hold

$$0 \leq \text{tr}(W\rho) = \text{tr}(W^{\text{boosted}}\rho^{\text{boosted}}), \quad (91)$$

where we denoted by “boosted” the observable that a Lorentz-boosted observer is assigning to the experimental setting.

Let us choose as a particular entanglement witness, the MUB-witness capable of detecting bound entangled states, defined in Eq. (43), $W = W_{\text{spin}}(\mathcal{M}_{d+1}, s \neq \frac{s}{2}) \otimes \mathbb{1}_{\text{momentum}}$ (see Sec. 4.2.1). Furthermore, let us choose $d = 3$ and $s = 1$ and a boost along the z -axis with a particular choice of energy (details can be found in Refs. 131 and 132). For any relativistic setting, one has to consider the spin and the momentum degrees of freedom of a quantum particle. Let us choose the Magic Simplex state $\rho_{\text{spin}}(x) = 2x P_{0,0} + x P_{1,1} + (\frac{1}{3} - x) \sum_{i=0}^2 P_{2,i}$ for two spin-1 particles with $x \in [0, \frac{1}{3}]$ for the spin part and a momentum state, e.g. $|\psi_{\text{momentum}}\rangle = \frac{1}{\sqrt{2}}\{|\vec{p}_+\vec{p}_-\rangle + |\vec{p}_-\vec{p}_+\rangle\}$ with $\vec{p}_{\pm}$ denoting opposite momenta. Then for $\rho = \rho_{\text{spin}}(x) \otimes \rho_{\text{momentum}}$ the inequality (91) turns to

$$0 \leq 1 - x, \quad (92)$$

which is not violated for any x . This is as expected since we have chosen the spin and the momentum part of our total state ρ to be separable.

The situation changes when we consider the spin or momentum subspaces, which we achieve by partial tracing over the respective subsystems. For the spin part, we find for the unboosted case

$$0 \leq \text{Tr}(\text{Tr}_{\text{momentum}}(W) \text{Tr}_{\text{momentum}}(\rho)) = -x. \quad (93)$$

Thus, a violation is found for any x that is not equal to zero. Indeed, the state is separable for $x = 0$, PPT entangled for $x \in \{0, \frac{2}{15}\}$ and else free entangled. Here, the MUB witness detects the full range of entanglement. On the other hand, for a particular relativistic boost, one finds that for no choice of x , a violation can be

found, i.e.

$$0 \leq \text{Tr}(\text{Tr}_{\text{momentum}}(W^{\text{boosted}}) \text{Tr}_{\text{momentum}}(\rho^{\text{boosted}})) = -x + 0.45. \quad (94)$$

Further detailed analyses show that the spin states $\text{Tr}_{\text{momentum}}(\rho^{\text{boosted}})$ are in this case separable.

In summary, a boosted state can lose its property of being bound or free entangled, i.e. the Lorentz invariance does not hold for either bound or free entangled states; it is not connected to the PPT property. From a practical point of view, different boosts allow to produce new sets of bound entangled states out of a particular bound entangled state.

6. Experiments Verifying Bound Entanglement

In recent years, impressive progress has been made in controlling the preparation, dynamics and measurement of high-dimensional quantum systems.^{133–137} Quantum computers, which work with qudits instead of qubits, are also no longer out of sight. With this progress, investigations into the preparation and observation of bound entangled states have also begun. In 2013, the first experiment based on MUB witnesses (cf. Sec. 4.2.1) was performed for two photons. We present the sketch of this experiment in Sec. 6.1, which confirmed for the first time the existence of this peculiar type of entangled states.

In general, verifying bound entanglement in the laboratory is an intrinsically hard problem since a large coherent volume of bound entangled states in Hilbert space must be found to enable verification despite experimental errors. In addition, a method must be found to distinguish between bound and free entanglement. With the discovery of the one-parameter family of states by the Horodecki family in 1998,⁴ it became clear that the search for experimental verification requires a larger family of bound entangled states to fulfill the experimental requirements. This led to investigating states within a subset of the entire Hilbert space, the Magic Simplex (see Sec. 2.3). These analyses show that bound entangled states are not rare in this subset, e.g. for $d = 3$ a relative volume of about 10% is bound entangled (see Sec. 5.2) and that bound entangled states can occupy large connected volumes in the Bell-diagonal state space that allow experimental investigations.

Once a family of states can be generated experimentally, the next task is to verify their entanglement. This can be done by any entanglement witness that can be realized on these physical systems but does not distinguish between bound and free entangled states. Alternatively, one can perform state tomography and then apply the tools presented in Secs. 3 and 4 to detect bound entanglement. This may not be the best strategy as it is time-consuming, often not realizable for experimental reasons, and usually requires additional assumptions. Another strategy is proposed in Ref. 138 via a hypothesis test tailored to the detection of bound entanglement and with a measure of statistical significance, but not yet applied experimentally. Ideas for preparing bound entangled states and proving their bound nature are urgently needed.

6.1. First experiment verifying the existence of bound entangled states for two qutrits

In 2013, the authors of Ref. 139 chose a particular state family of the Magic Simplex (cf. Definition 11) and applied a maximum complementarity protocol based on MUB observables to witness the entanglement of two photons entangled in their orbital angular momentum. Furthermore, they performed state tomography to verify that the generated states were PPT. Thus, the existence of bound entanglement in Nature was first experimentally verified 15 years after its discovery.⁴

Spatially entangled photon pairs can be generated by shining with a Gaussian laser beam ($\lambda = 413$ nm wavelength) onto a collinear Type-I spontaneous parametric downconversion (SPDC) crystal. The phase-matching in the crystal is tuned via temperature control to produce a similar amount of downconverted photons in the $l = 0$ and $l = \pm 1$ orbital angular momentum (OAM) mode. The conservation of orbital angular momentum forces opposite l , which allows only for the following state (which defines for this section the computational basis by setting $|l\rangle \mapsto |l+1\rangle$):

$$|\psi_{\text{SPDC}}\rangle = \frac{1}{\sqrt{3}}(|l=0\rangle \otimes |l=0\rangle + |l=+1\rangle \otimes |l=-1\rangle + |l=-1\rangle \otimes |l=+1\rangle). \quad (95)$$

This state is a Bell state that allows the construction of any other Bell state by applying a Weyl operator onto one of the subsystems (cf. Definition 10). The experiment achieves this via spatial light modulators (SLM), i.e. nonlinear materials. Projective measurements are done by appropriate mode transformation and subsequent imaging onto the core of a single mode fiber. The mixing was achieved by two different methods, one by considering only a subset of all possible measurements and the other by stochastic rotations applied to one of the photons. The photons are finally guided to an avalanche photodiode which functions as a single-photon counter.

In the experiment, four different correlation functions were measured, i.e. ($\mathcal{B}_0$ is the computational basis, and $|i_k\rangle$ are the basis states of the mutually unbiased bases $\mathcal{B}_k$; see Definition 18),

$$C_{\mathcal{B}_0, \mathcal{B}_0} = \sum_{i=0}^{d-1} \text{Tr}(|i\rangle\langle i| \otimes |i+2\rangle\langle i+2| \rho), \quad (96)$$

$$C_{\mathcal{B}_k, \mathcal{B}_k} = \sum_{i=0}^{d-1} \text{Tr}(|i_k\rangle\langle i_k| \otimes |i_k^*\rangle\langle i_k^*| \rho), \quad \text{with } 1 \leq k \leq 3. \quad (97)$$

The sum of these four correlation functions relates to the partial transpose of the MUB witness (43) for $m = 4$ and $s = 2$. If the sum is found to be greater than 2, entanglement is detected (cf. Eq. (46)). According to Theorem 11 this MUB witness forms a nondecomposable witness and can therefore detect bound entanglement. The experimental results were:¹³⁹

Correlation function	Theory	Experiment
$C_{\mathcal{B}_0, \mathcal{B}_0}$	0.675	0.667 ± 0.005
$C_{\mathcal{B}_1, \mathcal{B}_1}$	0.468	0.463 ± 0.005
$C_{\mathcal{B}_2, \mathcal{B}_2}$	0.468	0.463 ± 0.005
$C_{\mathcal{B}_3, \mathcal{B}_3}$	0.468	0.463 ± 0.005
$\sum_{i=0}^3 C_{\mathcal{B}_i, \mathcal{B}_i}$	2.079	2.066 ± 0.02

Here, the state was chosen from the Magic Simplex family B_1 with $\alpha = -0.07$, $\beta = -1.73$, defined and illustrated in Fig. 6.

The entanglement is detected since a value greater than 2 is only possible for entangled states. The experiment also showed excellent control since all entanglement class conserving symmetries (cf. Sec. 2.3) have been used to implement equivalent states.

The PPT property was verified via state tomography for different runs of the experiment and different choices of bases.¹⁴⁰ State tomography can always be performed by measuring all possible correlation functions based on MUBs, i.e. $\tilde{C}_k = \sum_{i,j=0}^{d-1} \text{Tr}(|i_k\rangle\langle i_k| \otimes |j_k\rangle\langle j_k| \rho)$.

7. Construction of Bound Entangled States

Constructing bound entangled states is generally difficult as no efficient solution exists to discriminate the sets SEP, BE and FE of general bipartite systems. There are various ways of approaching this problem, e.g. depending on the purpose of the construction or whether some bound entangled states are initially known.

An intensively studied strategy to construct bound entangled states relies on unextendible product bases (UPBs). We discuss this method in detail in Sec. 7.1. Alternatively, one can also analytically or numerically construct bound entangled states in different ways (e.g. Refs. 27 and 141–143). This can be accomplished by enforcing that the states satisfy the PPT criterion (Theorem 5) while violating at least one other entanglement criterion (cf. Sec. 4). In addition, depending on the problem investigated, further useful restrictions on the bound entangled states may be implemented this way.

The downside of the numerical approaches is that they often do not offer a deep insight into the analytic structure of bound entangled states. However, other properties, such as the frequency of bound entangled states, may become more feasible to investigate (cf. Sec. 5.2).

The situation is different if some bound entangled states are known initially (e.g. as the result of an analytic or numeric construction described above). Depending on the properties of these states, it is sometimes possible to construct further bound entangled states from them.^{69,70,144} Let us mention two strategies that build on concepts discussed in previous sections. First, one can easily utilize geometric sym-

metries to construct additional bound entangled states if the initial one is a Magic Simplex state, i.e. of standard Bell-diagonal form (cf. Sec. 2.3.1). Second, as discussed in Sec. 5.7, one may find new bound entangled states by Lorentz boosting others.

7.1. Construction from unextendible product bases

One way to analytically construct PPT bound entangled states utilizes the concept of unextendible product bases, defined by:

Definition 27 (Unextendible product basis). Given a multipartite Hilbert space $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$ of $n \geq 2$ parties and a set $S = \{|\phi_j\rangle = |\phi_j^1\rangle \otimes |\phi_j^2\rangle \otimes \cdots \otimes |\phi_j^n\rangle\}_{j=1}^m$ consisting of m mutually orthogonal product states in $\mathcal{H}$, i.e. $\langle \phi_i | \phi_j \rangle = \delta_{ij}$ for all $i, j = 1, \dots, m$. We denote by $\mathcal{H}_S$ the subspace of $\mathcal{H}$ spanned by the vectors in S , and by $\mathcal{H}_S^\perp$ the complementary subspace of $\mathcal{H}$ spanned by the set of vectors orthogonal to the ones in S . The set S is called unextendible product basis (UPB) if $\dim(\mathcal{H}_S^\perp) \neq 0$ and $\mathcal{H}_S^\perp$ contains no product vectors.

In Ref. 68 and 145, it is shown that such UPBs exist for specific dimensions. Given a UPB, one can construct PPT bound entangled states as follows.¹⁴⁵

Theorem 28. Let $S = \{|\phi_j\rangle\}_{j=1}^m$ be a UPB on a Hilbert space $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$ with $\dim(\mathcal{H}) = \prod_{i=1}^n d_i = D$ and $1 + \sum_{i=1}^n (d_i - 1) \leq m \leq D - 4$. Then the state

$$\rho_{\text{UPB}} = \frac{1}{D - m} \left(\mathbb{1}_D - \sum_{j=1}^m |\phi_j\rangle\langle\phi_j| \right) \quad (98)$$

is bound entangled.

To see that this is true, note that by construction, the state ρ_{UPB} is PPT because $\mathbb{1}_D$ is invariant under partial transposition and the states $|\phi_j\rangle$ are product states. In particular, $(|\phi_j\rangle\langle\phi_j|)^\Gamma = |\tilde{\phi}_j\rangle\langle\tilde{\phi}_j|$, where $|\tilde{\phi}_j\rangle$ are product states, and thus ρ_{UPB}^Γ is again a quantum state. Furthermore, Definition 27 entails that if S is a UPB, the complementary subspace $\mathcal{H}_S^\perp$ is nonempty and contains only entangled states. Thus, by the range criterion (Theorem 18), every state $\rho \in \mathcal{L}(\mathcal{H}_S^\perp)$ is entangled, and consequently ρ_{UPB} is bound entangled.

The state ρ_{UPB} in (98) is proportional to the projection operator onto the complementary subspace $\mathcal{H}_S^\perp$. As for the range of m in the theorem, it is proven in Ref. 145 that a lower bound on the number of elements in a UPB is given by $m_{\min} = 1 + \sum_{i=1}^n (d_i - 1)$, where $d_i = \dim(\mathcal{H}_i)$ (see also Ref. 146). Furthermore, it is clear that $\text{rank}(\rho_{\text{UPB}}) = D - m$. As no bound entangled states exist with rank smaller than 4 (cf. Theorem 21), a UPB cannot consist of more than $m_{\max} = D - 4$ product states. Conversely, such maximal UPBs can be used to construct bound entangled states with minimal rank.

Even though Theorem 28 establishes a way of constructing bound entangled states, it does not completely characterize them in general, i.e. not all bound entangled states can be constructed via a UPB. This is demonstrated in Ref. 145 by considering the bipartite bound entangled state from Ref. 32 with Hilbert space dimensions $d_1 = 2$ and $d_2 = 4$. As no UPBs exist in dimensions $2 \times n$ for any n , such a qubit-ququart state cannot be constructed in the above way. However, in special cases, UPBs enable a complete characterization of specific classes of bound entangled states. As shown in Ref. 147, for bipartite qutrit-qutrit systems, all rank-4 PPT bound entangled states can be obtained (up to local linear transformations) by maximal UPBs using Theorem 28.

Lastly, we mention that Theorem 28 does not present the only way to construct bound entangled states from UPBs. A different (but similar) approach is taken in Ref. 148. The authors consider the convex hull of the states in a UPB and obtain a more involved yet stronger result than the one from Ref. 145. In particular, they construct further bound entangled states in the vicinity of ρ_{UPB} from the same UPB.

8. Conclusion and Outlook

This review illuminates that many features of one of our most fundamental physical theories are still not fully understood. The tensor-product structure of the Hilbert space underlying bipartite finite-dimensional quantum systems is the origin of nonclassical correlations and thus, entanglement. In this context, the separability problem, i.e. distinguishing separable from nonseparable quantum states, is one of the main open questions. Solving this is shown to be NP-hard for general quantum systems of dimension $d > 6$. At its bottom lies the notion of local quantum operations and classical communications (LOCC) as entangled states are precisely those that cannot be created from separable states by LOCC alone. Even though we currently lack a good way to detect and quantify entanglement, it is clear that some states are more entangled than others in the sense that they exhibit stronger nonclassical correlations. This is especially true when considering probabilistic mixtures of pure quantum states, i.e. mixed states, which allows one to single out maximally entangled pure states from the rest. One question concerning the production of strongly entangled states via LOCC is: Given multiple copies of some weakly entangled states, is it possible to manipulate them by LOCC to obtain fewer but more strongly entangled states? Such a procedure is called entanglement distillation. It is proven that a necessary condition for the distillability of a quantum state is that it violates the PPT criterion. Entangled states that cannot be distilled to maximally entangled states are called bound entangled, and they naturally arise in the study of distillation. They seem to play a crucial role in our way to a full understanding of entanglement and its utilization as a resource.

Despite extensive research, the exact role of bound entanglement in connection to LOCC, entanglement distillation, and the PPT criterion still needs to be

determined. On the positive side, any new insight into bound entangled states will likely also lead to advances regarding the separability problem and vice versa. Many possible approaches to tackling this issue are summarized in this review. Thus, studying bound entangled states is crucial for fully understanding entanglement and the general nonclassical nature of quantum physics. Bound entangled states show limited use in currently known quantum technologies. Nonetheless, further investigation into the unique properties of bound entanglement may lead to the conception of specialized tasks only realizable with those states.

Setting aside these general motivations for the ongoing study of bound entanglement, the rest of this section will review the most relevant open questions from the previous sections. It thus presents a more detailed outlook on potential future research. Some questions that have found a definitive answer and are discussed in this review are collected in Table 4.

The current state of research concerning the detection of (bound) entanglement is summarized in Secs. 3 and 4. To date, the only sufficient criterion for the undistillability of entangled states is the PPT criterion (Theorem 5). This poses the question of whether NPT bound entangled states exist, a topic discussed in more detail at the end of this section. Section 4 emphasizes that almost all the presented

Table 4. This table summarizes some basic questions discussed in detail in this review. We have only selected dichotomous answers that always refer to all aspects mentioned in the respective sections.

Bound Entangled (BE) states		
Question	Answer	Section
Can BE states be distilled?	No	2 and 3
Does n -copy-undistillability for a state imply its undistillability?	No	2.2.3
Is BE invariant under LOCC?	No	2 and 3
Is there a single entanglement criterion that is strictly stronger than all others?	No	4
Can all entanglement witnesses detect BE states?	No	4.1.1
Is a complete set of MUBs needed for the detection of BE states?	No	4.2.1
Are BE states of measure zero in finite-dimensional Hilbert spaces?	No	5.2
Are BE states useful as resources in quantum teleportation?	No	5.3.1
Are BE states useful in device-dependent schemes as an additional resource?	Yes	5.3.3 and 5.3.4
Do all bound entangled states have zero entanglement of formation?	No	5.4
Can BE be useful for quantum metrology?	Yes	5.5
Can BE states violate a Bell inequality?	Yes	5.6
Are BE states Lorentz invariant?	No	5.7
Have BE states been verified experimentally?	Yes	6
Is there an analytic construction for all BE states?	No	7

entanglement criteria do not follow a strict hierarchy regarding the number of entangled states they detect. Hence, a universal criterion subsuming most of them has yet to be discovered. Finding such a universal criterion is generally difficult as the mathematical structure of the known criteria is often very different. One possible way forward is investigating the relation of bound entangled states (and their properties like entropic quantities, mixedness, etc.) to the criteria for detecting them. Concerning entanglement witnesses presented in Secs. 4.1 and 4.2, discovering an efficient way to find nondecomposable witnesses would lead to a novel method of constructing bound entangled states. One possible approach is via entanglement witnesses based on mutually unbiased bases (MUBs). The exact relation of MUBs and general nondecomposable witnesses is still not fully understood (cf. Sec. 4.2.2). Furthermore, nondecomposable witnesses can be constructed from nondecomposable PNCP maps via the Choi–Jamiołkowski isomorphism, possibly leading to the discovery of novel bound entangled states with interesting properties (cf. Ref. 149). Compared to other methods, entanglement witnesses constitute observables, allowing the detection of bound entangled states in experiments. As experimentally more accessible bound entangled states may be found using such a construction, verifying and subsequently utilizing bound entanglement in experiments could become easier.

Transitioning to the various characteristics of bound entangled states discussed in Sec. 5, one fundamental question concerns the frequency of these states. For a long time, it was argued that bound entanglement is a rare phenomenon. More recent studies show that the relative volume of bound entangled states in the state space is not zero (cf. Sec. 5.2), and hence, bound entanglement is not negligible in general. However, any volume calculation presumes the choice of a probability measure for the state space, yet no unique canonical measure has been found to date. Thus, a definitive relative volume determination is impossible. Investigating candidates for such a canonical measure might yield insight into bound entanglement and the geometry of the quantum state space. Additionally, numerical investigations suggest that PPT entanglement decreases approximately exponentially, independent of the chosen measure. Further insight into different state space measures could also clarify why this seems to be the case. As for the frequency of bound entanglement in the subset of Bell-diagonal states, it is remarkable that the largest share of bound entanglement can be found for the standard Bell basis (i.e. the Magic Simplex; cf. Sec. 2.3.1). The standard Bell basis possesses additional algebraic properties compared to nonstandard Bell bases. Consequently, one might conjecture that bound entanglement is linked to the underlying symmetries of the chosen basis. Investigating this specific family of states might help to understand the set of bound entangled states in the total state space.

In Sec. 5.3, it is elaborated that bound entangled states do not offer an advantage over separable states for most known quantum information processing tasks. The security in some networks utilizing secure key generation and the activation of bound entanglement are two exceptions. The latter usually refers to scenarios

where limited access to free entanglement with an additional supply of bound entangled states performs better in some quantum information protocols than the free or bound entangled states would allow on their own. Known examples include activating a quantum state's teleportation capability and its distillability. Activating the 1-distillability of any NPT state by PPT bound entangled states has very interesting implications if NPT bound entangled states would exist. In this case, one could obtain maximally entangled qudits from only bound entangled states and LOCC. Hence, the existence of NPT bound entanglement implies the nonadditivity of the distillable entanglement.

Including the distillable entanglement, some of the most relevant entanglement measures that attempt to quantify entanglement as a quantum resource are introduced in Sec. 5.4. As any such measure is generally very hard to evaluate for mixed states, precise results for bound entangled states are rare. Crucially, however, a fundamental irreversibility in (catalytic) LOCC operations is proven to exist via the difference between the distillable entanglement and the entanglement cost for bound entangled states. Bound entangled states require a nonzero amount of maximally entangled states for their formation, which, in turn, cannot be extracted by entanglement distillation. This irreversibility relates to, or may be, the main reason for the difficulties of defining a general entanglement measure quantifying entanglement as a resource. No such measure is known, and the use of a specific measure is usually motivated by a corresponding operational task. Finding and relating entanglement measures to operational tasks contributes to the identification of the most relevant properties they should satisfy. Investigating properties like (sub-/super-) additivity, asymptotic continuity, or lockability of proposed measures contributes to such operational interpretations. As bound entanglement is, by definition, related to entanglement distillation, the distillable entanglement is certainly a measure of interest. Evaluating and comparing different measures for certain NPT and PPT entangled states, such as Werner or Bell-diagonal states, may provide information about the entanglement structure and properties of bound entangled states. As these computations are generally not feasible, finding stricter bounds for entanglement measures, like the Rains bound,¹⁵⁰ is of crucial importance. While any other entanglement measure provides an upper bound for the distillable entanglement, lower bounds can, for instance, be found by analyzing Bell-diagonal states with respect to measures presented in this review or others like squashed entanglement.¹⁵¹ Even though irreversibility is present in the LOCC paradigm, other paradigms beyond LOCC have recently been proposed in which reversibility may be achieved.^{152,153} In contrast, however, the paradigms of nonentangling operations and PPT operations have been shown to exhibit the irreversibility related to bound entanglement.^{154,155} Notably, irreversibility under nonentangling operations distinguishes entanglement from other quantum resources (cf. Ref. 155). Since irreversibility lies at the heart of bound entanglement, investigations that explore this phenomenon under more general nonglobal paradigms than LOCC may contribute to our understanding of this type of entanglement.

One further question with direct physical motivation concerns bound entanglement in relativistic settings. This is particularly intriguing in the context of the nonlocal features of bound entangled states (cf. Sec. 5.6) and the local nature of special relativity. Examples show that Lorentz boosts can change the entanglement class of a state from free entangled to bound entangled or separable. However, a universal understanding is still missing. Such a comprehension could reveal crucial insights into the interplay of quantum physics and relativity and open new possibilities for constructing bound entangled states analytically.

Regarding general bound entanglement construction schemes described in Sec. 7, much research still needs to be done. Nowadays, numerically generating a plethora of bound entangled states with certain properties, e.g. inside the Magic Simplex, is possible. On the other hand, the main analytic method with some generality is via unextendible product bases (UPBs). As any construction necessarily exploits some property of bound entangled states, it will be crucial to investigate further constructions to gain a deeper understanding of the underlying physics and mathematics. So far, all schemes we know can only construct PPT bound entangled states.

This brings us to the central question of the existence of NPT bound entangled states, i.e. potential undistillable entangled states that violate the PPT criterion (Theorem 5). See Fig. 4 for a visualization of the problem. In this regard, the main open challenge is proving or disproving whether the PPT criterion is also necessary for the undistillability of bipartite quantum systems with general Hilbert space dimensions $d_A \times d_B$ (cf. Theorem 7). A consequence of Theorems 2 and 6 is that this is true for qubit-qubit (2×2) systems. In Ref. 156, it is furthermore shown that it extends to systems of Hilbert space dimension $2 \times d$ with $d > 2$. The difficulty of investigating undistillability lies in the definition of distillability itself (cf. Sec. 2.2.3). A state is said to be distillable if, for some $n > 0$, there is a LOCC protocol transforming n identical copies of the state to some maximally entangled state. An equivalent characterization of distillability uses any entangled qubit state as the target state of distillation (cf. Definition 8 and Theorem 2). However, for every $d > 2$ and $n > 0$, it is shown that there are states with Hilbert space dimensions $d^2 \times d^2$ that are not n -distillable, but are $(n + 1)$ -distillable (cf. Theorem 3). Thus, it is not enough to check that a state is n -undistillable for some finite $n > 0$ to establish that it is generally undistillable. Besides this complication, there is one major simplification to the problem of NPT bound entanglement. Reference 16 shows that to solve the general problem, it is sufficient to prove whether NPT bound entangled states exist among the one-parameter family of Werner states.¹⁵⁷ The reason for this is that every NPT state can be transformed by LOCC to an NPT Werner state. However, proving the undistillability of these highly symmetric Werner states has not been successful yet. There have been numerous attempts to characterize them. For every $n > 0$, Werner states are shown to be n -undistillable for a certain parameter range.^{156,158} Whether this parameter range vanishes as $n \rightarrow \infty$ is not known. Furthermore, it is conjectured that every 1-undistillable

Werner state is also 2-undistillable,¹⁵⁹ and numerical analyses show that such states are probably also 3-undistillable.¹⁵⁸ Further approaches and references can be found in Ref. 160 and 161. We mention in passing that, due to the Choi–Jamiołkowski isomorphism, the problem can also be restated in terms of positive maps.^{156,162} One major consequence of the existence of NPT bound entangled states is that it would imply the nonadditivity and nonconvexity of the distillable entanglement.¹⁶³ This is exemplified by the activation of 1-distillability discussed in Sec. 5.3.3 (cf. also Ref. 86). In particular, suppose ρ_{PPT} is a PPT bound entangled universal activator state and ρ_{NPT} is an NPT bound entangled state. Using Definition 25, one would obtain $E_D(\rho_{\text{NPT}} \otimes \rho_{\text{PPT}}) > 0 = E_D(\rho_{\text{NPT}}) + E_D(\rho_{\text{PPT}})$.

In summary, bound entanglement is a unique form of quantum entanglement that cannot be used for distillation. It remains a complex and not fully grasped phenomenon, indicating the limitations of our current understanding of these quantum correlations. The complexities associated with the detection, characterization, and quantification of bound entanglement underscore the incompleteness of our theoretical framework. The inability to fully quantify it as a resource in quantum information processing tasks further highlights the gaps in our knowledge, leaving the potential role of bound entanglement in practical applications of quantum information processing largely unexplored. Recent advancements in the detection and characterization of bound entanglement in specific systems offer more than just a glimmer of hope. Future research contributing to these developments could potentially lead us towards a comprehensive understanding of entanglement and its effective utilization in quantum technologies. The journey towards unraveling the mysteries of bound entanglement continues, paving the way for exciting discoveries in the realm of quantum physics and quantum information theory.

Acknowledgments

B.C.H. and C.P. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) project P36102-N (Grant DOI: 10.55776/P36102). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Appendix A. Minimal Fidelity of Teleportation

Here, we briefly recap how the measure-and-prepare teleportation protocol (cf. Ref. 77) works and calculate its fidelity of transmission. The general setup is identical to the one described at the beginning of Sec. 5.3.1. In the first step of the protocol, Alice measures her unknown quantum state $|\phi\rangle$ in some orthonormal basis $\{|v_i\rangle\}$. The probability of Alice obtaining outcome i is given by $p_i = |\langle v_i | \phi \rangle|^2$. Next, she classically transmits the information about the measurement basis and the outcome i to Bob, who prepares the respective quantum state $|v_i\rangle$. This will serve as his

guess as to what $|\phi\rangle$ was. Hence, taking into account all possible measurement results i , Bob prepares the mixed state $\sigma_\phi = \sum_{i=0}^{d-1} p_i |v_i\rangle\langle v_i| = \sum_{i=0}^{d-1} |\langle v_i|\phi\rangle|^2 |v_i\rangle\langle v_i|$. The corresponding fidelity of transmission (68) is thus given by

$$f = \int d\phi F(\sigma_\phi, |\phi\rangle) = \int d\phi \langle \phi | \sigma_\phi | \phi \rangle = \sum_{i=0}^{d-1} \int d\phi |\langle v_i | \phi \rangle|^4. \quad (\text{A.1})$$

To evaluate this integral, we write $|\phi\rangle = U_\phi|0\rangle$ and $|v_i\rangle = V_i|0\rangle$, where U_ϕ and V_i are unitary. This allows us to compute

$$\begin{aligned} f &= \sum_{i=0}^{d-1} \int d\phi |\langle v_i | \phi \rangle|^4 \\ &= \sum_{i=0}^{d-1} \int d\phi |\langle 0 | V_i^\dagger U_\phi | 0 \rangle|^4 \\ &= \sum_{i=0}^{d-1} \int d\phi |\langle 0 | \tilde{U}_i^\phi | 0 \rangle|^4 \\ &= \sum_{i=0}^{d-1} \frac{2}{d(d+1)} = \frac{2}{(d+1)}, \end{aligned} \quad (\text{A.2})$$

where we defined $\tilde{U}_i^\phi := V_i^\dagger U_\phi$ in the third equality. The fourth equality follows directly from the explicit calculation in Ref. 164 using the composite parametrization of unitary matrices.

As the measure-and-prepare teleportation protocol is applicable independent of the resource state shared by Alice and Bob, the minimal fidelity of transmission attainable in any scenario is lower bound by (A.2).

Appendix B. Protocol for Activation of Bound Entanglement

Here, we reproduce the example of a free entangled qutrit-qutrit state from Ref. 75, which is subsequently used in the activation protocol from Ref. 84.

Consider the free entangled qutrit-qutrit state

$$\rho^{\text{FE}} = \mathcal{F}_0 P_{0,0} + \frac{1 - \mathcal{F}_0}{3} \sum_{i=0}^2 P_{i,1}, \quad (\text{B.1})$$

with $0 < \mathcal{F}_0 < 1$ being its singlet fraction (69), and $P_{k,l}$ are the Bell state projectors (16). Note that this constitutes a mixture of the maximally entangled qutrit state (4) and a separable subgroup (line) state (19). Therefore, it is a Magic Simplex state (cf. Definition 11). That this state is free entangled can be shown by considering the local projection of ρ^{FE} onto the subspace of two degrees of freedom

$(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|)$, resulting in the effective qubit state

$$\rho_{2 \times 2}^{\text{FE}} = \frac{\mathcal{F}_0}{1 + \mathcal{F}_0} (|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|) + \frac{1 - \mathcal{F}_0}{1 + \mathcal{F}_0} |01\rangle\langle 01|. \quad (\text{B.2})$$

This state is NPT for $0 < \mathcal{F}_0 < 1$, and thus according to Theorems 2 and 6 it is distillable. Hence, also ρ^{FE} is distillable by the argument presented in Sec. 2.2.3. However, as proven in Ref. 75, the singlet fraction obtainable with nonzero probability by LOCC and a *single* copy of (B.1) is upper bounded by some constant $\mathcal{F}_{\text{max}} < 1$. This even holds when LOCC protocols with a success probability tending to zero are considered. By Theorem 24, faithful teleportation is thus impossible with only a single copy of ρ^{FE} because the maximally achievable fidelity of transmission is upper bounded by $f_{\text{max}} = (3\mathcal{F}_{\text{max}} + 1)/4 < 1$.

For the specific activation protocol from Ref. 84, the authors defined the activation state

$$\rho^{\text{BE}} = \frac{2}{7} P_{0,0} + \frac{\alpha}{21} \sum_{i=0}^2 P_{i,1} + \frac{5-\alpha}{21} \sum_{i=0}^2 P_{i,2}. \quad (\text{B.3})$$

Again, this is a mixture of the maximally entangled qutrit state (4) and two distinct subgroup (line) states (19). Thus, also (B.3) constitutes a Magic Simplex state. It is free entangled for $0 \leq \alpha < 1$ and $4 < \alpha \leq 5$, bound entangled for $1 \leq \alpha < 2$ and $3 < \alpha \leq 4$, and separable for $2 \leq \alpha \leq 3$. Separability follows from an explicit decomposition of (B.3) into a mixture of separable states, the free entanglement can be shown as for (B.1) above, and the bound entangled regions are detected by the realignment criterion (cf. Sec. 4.4).

Suppose now that Alice and Bob share one free entangled state (B.1) and additionally multiple copies of the bound entangled state (B.3) with $3 < \alpha \leq 4$. They are furthermore both able to apply a generalized d -dimensional XOR-gate locally. This quantum gate is represented by the unitary matrix

$$U_{\text{XOR}} = \sum_{i,j=0}^{d-1} |i\rangle\langle i| \otimes |j+i\rangle\langle j|. \quad (\text{B.4})$$

We call the unaltered first subsystem the control qudit and the altered second subsystem the target qudit. Alice and Bob can increase the singlet fraction of the free entangled state ρ^{FE} by LOCC as follows:

- (1) They each apply the XOR-gate to their parts of the free entangled state ρ^{FE} and the bound entangled state ρ^{BE} . The free (bound) entangled qudit is the control (target).
- (2) Alice and Bob each measure their respective target qudits in the computational basis and compare their results. If they differ, the procedure fails, and the free entangled resource is lost. However, if the results coincide, the protocol succeeds, and only the target qudits are discarded. The free entangled state now has a higher singlet fraction $\mathcal{F}_1 > \mathcal{F}_0$.

It is straightforward to compute the success probability of step two to be

$$P_{\mathcal{F}_0 \rightarrow \mathcal{F}_1} = \frac{2\mathcal{F}_0 + (1 - \mathcal{F}_0)(5 - \alpha)}{7}. \quad (\text{B.5})$$

After discarding the target qudits, the control pair is still of the form (B.1), but with an updated singlet fraction given by

$$\mathcal{F}_1 = \frac{2\mathcal{F}_0}{2\mathcal{F}_0 + (1 - \mathcal{F}_0)(5 - \alpha)}. \quad (\text{B.6})$$

A simple check confirms that for $3 < \alpha \leq 5$ one has $\mathcal{F}_1 > \mathcal{F}_0$ for any initial $0 < \mathcal{F}_0 < 1$. This includes the bound entangled range $3 < \alpha \leq 4$.

The above procedure can be repeated as often as Alice's and Bob's supply of bound entangled states allows. The probability of n consecutive successes P_n is given by the product of probabilities (B.5) with updated $\mathcal{F}_n$ for each round. This probability may be low, but it is nonzero. For each repetition, a bound entangled state is consumed in exchange for an increase in the singlet fraction of the free entangled state. This can be interpreted as transferring the entanglement from a bound entangled state, which cannot be accessed directly, to a free entangled state, where it can be used. By this procedure, the singlet fraction $\mathcal{F}_n$ can be boosted arbitrarily close to 1 for large n , provided Alice and Bob share enough bound entangled states. Consequently, $\mathcal{F}_n > \mathcal{F}_{\max}$ for some n , thus exceeding the maximal singlet fraction attainable by (B.1) alone. Interpreted as a distillation scheme, this suggests that the distillable entanglement (cf. Definition 25) may not be additive.²²

By virtue of Theorem 24 this can be viewed as follows: Generally, bound entangled states are not useful for teleportation as they do not outperform any classical teleportation protocol. This also holds true for the case of (B.3) with $3 < \alpha \leq 4$. Furthermore, the free entangled state (B.1) cannot achieve a fidelity of transmission (68) arbitrarily close to 1. However, if these two resources are combined, one can "activate" the bound entanglement contained in (B.3) by increasing the maximal fidelity of transmission of (B.1). In this way, bound entangled states can improve the performance in specific teleportation scenarios.

Lastly, for completeness, we add that the bound entanglement of (B.3) in the range $1 \leq \alpha < 2$ can also be activated, albeit with slight modifications. This requires the free entangled resource state

$$\tilde{\rho}^{\text{FE}} = \mathcal{F}_0 P_{0,0} + \frac{1 - \mathcal{F}_0}{3} \sum_{i=0}^2 P_{i,2}. \quad (\text{B.7})$$

The probability of equal outcomes in step two is

$$\tilde{P}_{\mathcal{F}_0 \rightarrow \mathcal{F}_1} = \frac{2\mathcal{F}_0 + (1 - \mathcal{F}_0)\alpha}{7}, \quad (\text{B.8})$$

after which the singlet fraction changes to

$$\tilde{\mathcal{F}}_1 = \frac{2\mathcal{F}_0}{2\mathcal{F}_0 + (1 - \mathcal{F}_0)\alpha}. \quad (\text{B.9})$$

It holds that $\tilde{\mathcal{F}}_1 > \mathcal{F}_0$ for $0 \leq \alpha < 2$ and therefore in particular in the bound entangled region $1 \leq \alpha < 2$.

ORCID

Beatrix C. Hiesmayr  <https://orcid.org/0000-0001-9062-6039>

Christopher Popp  <https://orcid.org/0000-0002-2352-3731>

Tobias C. Sutter  <https://orcid.org/0009-0008-3179-7979>

References

1. E. Schrödinger, *Math. Proc. Camb. Philos. Soc.* **31** (1935) 555.
2. R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, *Rev. Mod. Phys.* **81** (2009) 865.
3. L. Gurvits, Classical deterministic complexity of Edmonds' Problem and quantum entanglement, in *Proc. Thirty-Fifth Annual ACM Symp. Theory of Computing* (ACM, San Diego, CA, USA, 2003), pp. 10–19.
4. M. Horodecki, P. Horodecki and R. Horodecki, *Phys. Rev. Lett.* **80** (1998) 5239.
5. F. G. S. L. Brandão and M. B. Plenio, *Nat. Phys.* **4** (2008) 873.
6. M. Horodecki, J. Oppenheim and R. Horodecki, *Phys. Rev. Lett.* **89** (2002) 240403.
7. P. Horodecki, *Phys. Lett. A* **232** (1997) 333.
8. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 1 edn. (Cambridge University Press, 2012).
9. S. Arora and B. Barak, *Computational complexity: A Modern Approach*, 4th printing 2016 edn. (Cambridge University Press, New York, 2016).
10. J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, UK, 2018).
11. I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 1 edn. (Cambridge University Press, 2006).
12. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin and W. K. Wootters, *Phys. Rev. Lett.* **76** (1996) 722.
13. G. Alber, A. Delgado, N. Gisin and I. Jex, *J. Phys. A: Math. Gen.* **34** (2001) 8821.
14. D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu and A. Sanpera, *Phys. Rev. Lett.* **77** (1996) 2818.
15. C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, *Phys. Rev. A* **53** (1996) 2046.
16. M. Horodecki and P. Horodecki, *Phys. Rev. A* **59** (1999) 4206.
17. M. M. Wilde, *Quantum Information Theory*, 2nd edn. (Cambridge University Press, Cambridge, UK, 2017).
18. M. Horodecki, P. Horodecki and R. Horodecki, *Phys. Rev. Lett.* **78** (1997) 574.
19. J. Watrous, *Phys. Rev. Lett.* **93** (2004) 010502.
20. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, *Phys. Rev. Lett.* **70** (1993) 1895.
21. C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69** (1992) 2881.
22. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, *Phys. Rev. A* **54** (1996) 3824.
23. B. Baumgartner, B. Hiesmayr and H. Narnhofer, *J. Phys. A: Math. Theor.* **40** (2007) 7919.
24. W. K. Wootters and B. D. Fields, *Ann. Phys.* **191** (1989) 363.

25. B. Baumgartner, B. C. Hiesmayr and H. Narnhofer, *Phys. Lett. A* **372** (2008) 2190.
26. B. Baumgartner, B. C. Hiesmayr and H. Narnhofer, *Phys. Rev. A* **74** (2006) 032327.
27. C. Popp and B. C. Hiesmayr, *Sci. Rep.* **12** (2022) 12472.
28. C. Popp and B. C. Hiesmayr, *Sci. Rep.* **13** (2023) 2037.
29. C. Popp and B. C. Hiesmayr, *New J. Phys.* **26** (2024) 013039.
30. M.-D. Choi, *Linear Algebra Appl.* **10** (1975) 285.
31. A. Jamiołkowski, *Rep. Math. Phys.* **3** (1972) 275.
32. M. Horodecki, P. Horodecki and R. Horodecki, *Phys. Lett. A* **223** (1996) 1.
33. A. Peres, *Phys. Rev. Lett.* **77** (1996) 1413.
34. V. Vedral, M. B. Plenio, M. A. Rippin and P. L. Knight, *Phys. Rev. Lett.* **78** (1997) 2275.
35. D. Chruściński and G. Sarbicki, *J. Phys. A: Math. Theor.* **47** (2014) 483001.
36. J. Bae, D. Chruściński and B. C. Hiesmayr, *npj Quantum Inf.* **6** (2020) 1.
37. D. Chruściński, A. Bera, J. Bae and B. C. Hiesmayr, arXiv:2503.04158 [quant-ph].
38. A. Bera, J. Bae, B. C. Hiesmayr and D. Chruściński, *Sci. Rep.* **13** (2023) 10733.
39. H. F. Hofmann and S. Takeuchi, *Phys. Rev. A* **68** (2003) 032103.
40. O. Gühne and N. Lütkenhaus, *Phys. Rev. Lett.* **96** (2006) 170502.
41. J. Bae, B. C. Hiesmayr and D. McNulty, *New J. Phys.* **21** (2019) 013012.
42. I. D. Ivonovic, *J. Phys. A: Math. Gen.* **14** (1981) 3241.
43. T. Durt, B.-G. Englert, I. Bengtsson and K. Åzyczkowski, *Int. J. Quantum Inform.* **08** (2010) 535.
44. P. O. Boykin, M. Sitharam, M. Tarifi and P. Wocjan, arXiv:quant-ph/0502024.
45. C. Spengler, M. Huber, S. Brierley, T. Adaktylos and B. C. Hiesmayr, *Phys. Rev. A* **86** (2012) 022311.
46. B. C. Hiesmayr, D. McNulty, S. Baek, S. S. Roy, J. Bae and D. Chruściński, *New J. Phys.* **23** (2021) 093018.
47. J. Bae, A. Bera, D. Chruściński, B. C. Hiesmayr and D. McNulty, *J. Phys. A: Math. Theor.* **55** (2022) 505303.
48. A. J. Scott, arXiv:1703.03993 [quant-ph].
49. M. Grassl, D. McNulty, L. Miřta and T. Paterek, *Phys. Rev. A* **95** (2017) 012118.
50. B. C. Stacey, *Found. Phys.* **47** (2017) 1060.
51. G. Sarbicki, G. Scala and D. Chruściński, *Phys. Rev. A* **101** (2020) 012341.
52. J. De Vicente, *Quantum Inf. Comput.* **7** (2007) 624.
53. K. Chen and L.-A. Wu, *Quantum Inf. Comput.* **3** (2003) 193.
54. O. Rudolph, *Quantum Inf. Process.* **4** (2005) 219.
55. J. Shang, A. Asadian, H. Zhu and O. Gühne, *Phys. Rev. A* **98** (2018) 022309.
56. L.-M. Lai, T. Li, S.-M. Fei and Z.-X. Wang, *Quantum Inf. Process.* **17** (2018) 314.
57. J. Bae, M. Tiersch, S. Sauer, F. De Melo, F. Mintert, B. Hiesmayr and A. Buchleitner, *Phys. Rev. A* **80** (2009) 022317.
58. W. K. Wootters, *Phys. Rev. Lett.* **80** (1998) 2245.
59. B. C. Hiesmayr and M. Huber, *Phys. Rev. A* **78** (2008) 012342.
60. F. Mintert, M. Kuś and A. Buchleitner, *Phys. Rev. Lett.* **95** (2005) 260502.
61. F. Mintert, A. R. R. Carvalho, M. Kuś and A. Buchleitner, *Phys. Rep.* **415** (2005) 207.
62. B. Bhattacharya, S. Goswami, R. Mundra, N. Ganguly, I. Chakrabarty, S. Bhattacharya and A. S. Majumdar, *J. Phys. Commun.* **5** (2021) 065008.
63. P. Horodecki, J. A. Smolin, B. M. Terhal and A. V. Thapliyal, *Theor. Comput. Sci.* **292** (2003) 589.
64. L. Chen and D. Z. Đoković, *J. Phys. A: Math. Theor.* **44** (2011) 285303.
65. P. Horodecki, M. Lewenstein, G. Vidal and I. Cirac, *Phys. Rev. A* **62** (2000) 032310.

66. L. Chen and Y.-X. Chen, *Phys. Rev. A* **78** (2008) 022318.
67. D. Bruß and A. Peres, *Phys. Rev. A* **61** (2000) 030301.
68. D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, *Commun. Math. Phys.* **238** (2003) 379.
69. P. Bej and S. Halder, *Phys. Lett. A* **386** (2021) 126992.
70. S. Bandyopadhyay, S. Ghosh and V. Roychowdhury, *Phys. Rev. A* **71** (2005) 012316.
71. P. Horodecki, J. I. Cirac and M. Lewenstein, Bound entanglement for continuous variables is a rare phenomenon, in *Quantum Information with Continuous Variables*, eds. S. L. Braunstein and A. K. Pati (Springer Netherlands, Dordrecht, 2003), pp. 211–228.
72. K. Życzkowski, P. Horodecki, A. Sanpera and M. Lewenstein, *Phys. Rev. A* **58** (1998) 883.
73. P. B. Slater, *J. Phys. A: Math. Gen.* **32** (1999) 5261.
74. K. Życzkowski, *Phys. Rev. A* **60** (1999) 3496.
75. M. Horodecki, P. Horodecki and R. Horodecki, *Phys. Rev. A* **60** (1999) 1888.
76. R. Jozsa, *J. Mod. Opt.* **41** (1994) 2315.
77. N. Linden and S. Popescu, *Phys. Rev. A* **59** (1999) 137.
78. E. Chitambar and F. Leditzky, *IEEE Trans. Inf. Theory* **70** (2024) 3529.
79. C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal, *Phys. Rev. Lett.* **83** (1999) 3081.
80. C. Bennett, P. Shor, J. Smolin and A. Thapliyal, *IEEE Trans. Inf. Theory* **48** (2002) 2637.
81. M. Horodecki, P. Horodecki, R. L. Horodecki, D. Leung and B. Terha, *Quantum Inf. Comput.* **1** (2001) 70.
82. N. Linden, S. Massar and S. Popescu, *Phys. Rev. Lett.* **81** (1998) 3279.
83. A. Kent, *Phys. Rev. Lett.* **81** (1998) 2839.
84. P. Horodecki, M. Horodecki and R. Horodecki, *Phys. Rev. Lett.* **82** (1999) 1056.
85. F. Ozaydin, C. Bayindir, A. A. Altintas and C. Yesilyurt, *Phys. Rev. A* **105** (2022) 022439.
86. K. G. H. Vollbrecht and M. M. Wolf, *Phys. Rev. Lett.* **88** (2002) 247901.
87. H.-K. Lo and S. Popescu, *Phys. Rev. A* **63** (2001) 022301.
88. G. Vidal, *Phys. Rev. Lett.* **83** (1999) 1046.
89. S. Ishizaka, *Phys. Rev. Lett.* **93** (2004) 190501.
90. P. Horodecki, M. Horodecki and R. Horodecki, *J. Mod. Optics* **47** (2000) 347.
91. M. Horodecki, P. W. Shor and M. B. Ruskai, *Rev. Math. Phys.* **15** (2003) 629.
92. C. H. Bennett and G. Brassard, *Theor. Comput. Sci.* **560** (2014) 7.
93. M. Curty, M. Lewenstein and N. L. Aijtkenhaus, *Phys. Rev. Lett.* **92** (2004) 217903.
94. P. Horodecki and R. Augusiak, arXiv:0712.3999 [quant-ph].
95. K. Horodecki, M. Horodecki, P. Horodecki and J. Oppenheim, *Phys. Rev. Lett.* **94** (2005) 160502.
96. K. Horodecki, M. Horodecki, P. Horodecki and J. Oppenheim, *IEEE Trans. Inf. Theory* **55** (2009) 1898.
97. K. Horodecki, L. Pankowski, M. Horodecki and P. Horodecki, *IEEE Trans. Inf. Theory* **54** (2008) 2621.
98. I. Devetak and A. Winter, *Proc. R. Soc. A* **461** (2005) 207.
99. L. Pankowski and M. Horodecki, *J. Phys. A: Math. Theor.* **44** (2011) 035301.
100. K. Horodecki and G. Murta, *Phys. Rev. A* **92** (2015) 010301.
101. M. Ozols, G. Smith and J. A. Smolin, *Phys. Rev. Lett.* **112** (2014) 110502.
102. K. Banaszek, K. Horodecki and P. Horodecki, *Phys. Rev. A* **85** (2012) 012330.
103. H.-J. Briegel, W. Dür, J. I. Cirac and P. Zoller, *Phys. Rev. Lett.* **81** (1998) 5932.

104. S. Bäuml, M. Christandl, K. Horodecki and A. Winter, *Nat. Commun.* **6** (2015) 6908.
105. O. Sakarya, M. Winczewski, A. Rutkowski and K. Horodecki, *Phys. Rev. Res.* **2** (2020) 043022.
106. W. Zhang, T. Van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C. C.-W. Lim and H. Weinfurter, *Nature* **607** (2022) 687.
107. M. Christandl, R. Ferrara and K. Horodecki, *Phys. Rev. Lett.* **126** (2021) 160501.
108. R. Arnon-Friedman and F. Leditzky, *IEEE Trans. Inf. Theory* **67** (2021) 6606.
109. M. Plenio and S. Virmani, *Quantum Inf. Comput.* **7** (2007) 1.
110. M. J. Donald, M. Horodecki and O. Rudolph, *J. Math. Phys.* **43** (2002) 4252.
111. P. M. Hayden, M. Horodecki and B. M. Terhal, *J. Phys. A: Math. Gen.* **34** (2001) 6891.
112. B. C. Hiesmayr, M. Huber and P. Krammer, *Phys. Rev. A* **79** (2009) 062308.
113. G. Vidal and J. I. Cirac, *Phys. Rev. Lett.* **86** (2001) 5803.
114. D. Yang, M. Horodecki, R. Horodecki and B. Synak-Radtke, *Phys. Rev. Lett.* **95** (2005) 190501.
115. L. Lami, B. Regula and A. Streltsov, *Phys. Rev. A* **109** (May 2024) L050401.
116. G. Tóth and I. Apellaniz, *J. Phys. A: Math. Theor.* **47** (2014) 424006.
117. J. Liu, H. Yuan, X.-M. Lu and X. Wang, *J. Phys. A: Math. Theor.* **53** (2019) 023001.
118. S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72** (May 1994) 3439.
119. L. Pezzé and A. Smerzi, *Phys. Rev. Lett.* **102** (2009) 100401.
120. G. Tóth and T. Vértesi, *Phys. Rev. Lett.* **120** (January 2018) 020506.
121. P. Badziąg, K. Horodecki, M. Horodecki, J. Jenkinson and S. J. Szarek, *Phys. Rev. A* **90** (2014) 012301.
122. K. F. Pál, G. Tóth, E. Bene and T. T. Vértesi, *Phys. Rev. Res.* **3** (2021) 023101.
123. J. S. Bell, *Physics* **1** (1964) 195.
124. J. S. Bell, *Speakable and Unsayable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, 2 edn. (Cambridge University Press, Cambridge, 2004).
125. T. T. Vértesi and N. Brunner, *Nat. Commun.* **5** (2014) 5297.
126. A. Peres, *Found. Phys.* **29** (1999) 589.
127. K. F. Pál and T. T. Vértesi *Phys. Rev. A* **96** (2017) 022123.
128. D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu, *Phys. Rev. Lett.* **88** (2002) 040404.
129. C. Spengler, M. Huber and B. C. Hiesmayr, *J. Phys. A: Math. Theor.* **44** (2011) 065304.
130. N. Friis, R. A. Bertlmann, M. Huber and B. C. Hiesmayr, *Phys. Rev. A* **81** (2010) 042114.
131. P. Caban and B. C. Hiesmayr, *Sci. Rep.* **13** (2023) 11189.
132. P. Caban and J. Rembieliński, *Phys. Rev. A* **72** (2005) 012103.
133. Y. Wang, Z. Hu, B. C. Sanders and S. Kais, *Front. Phys.* **8** (2020) 589504.
134. Y. Chi, J. Huang, Z. Zhang, J. Mao, Z. Zhou, X. Chen, C. Zhai, J. Bao, T. Dai, H. Yuan, M. Zhang, D. Dai, B. Tang, Y. Yang, Z. Li, Y. Ding, L. K. Oxenlind, M. G. Thompson, J. L. O'Brien, Y. Li, Q. Gong and J. Wang, *Nat. Commun.* **13** (2022) 1166.
135. M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler and T. Monz, *Nat. Phys.* **18** (2022) 1053.
136. G. Carvacho, F. Graffitti, V. D'Ambrosio, B. C. Hiesmayr and F. Sciarrino, *Sci. Rep.* **7** (2017) 13265.
137. B. Hiesmayr, M. de Dood and W. Löffler, *Phys. Rev. Lett.* **116** (2016) 073601.

138. G. Sentís, J. N. Greiner, J. Shang, J. Siewert and M. Kleinmann, *Quantum* **2** (2018) 113.
139. B. C. Hiesmayr and W. Löffler, *New J. Phys.* **15** (2013) 083036.
140. W. Löffler and B. C. Hiesmayr, Detection of bound entanglement of OAM photons using complementarity, in *Frontiers in Optics 2013* (Optica Publishing Group, 2013), paper LW2G.3.
141. W. Li, R. Han, J. Shang, H. K. Ng and B.-G. Englert, arXiv:2109.14215 [quant-ph].
142. E. Sindici and M. Piani, *Phys. Rev. A* **97** (2018) 032319.
143. J. Moerland, N. Wyderka, H. Kampermann and D. Bruß, *Phys. Rev. A* **109** (2024) 012217.
144. H. Zhao, S. Guo, N. Jing and S. Fei, *Quantum Inf. Process.* **15** (2016) 1529.
145. C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, *Phys. Rev. Lett.* **82** (1999) 5385.
146. N. Alon and L. Lovász, *J. Comb. Theory, Ser. A* **95** (2001) 169.
147. L. Skowronek, *J. Math. Phys.* **52** (2011) 122202.
148. A. O. Pittenger, *Linear Algebra Appl.* **359** (2003) 235.
149. J. K. Korbicz, M. L. Almeida, J. Bae, M. Lewenstein and A. Acín, *Phys. Rev. A* **78** (2008) 062105.
150. E. Rains, *IEEE Trans. Inf. Theory* **47** (2001) 2921.
151. M. Christandl and A. Winter, *J. Math. Phys.* **45** (2004) 829.
152. B. Regula and L. Lami, *Nat. Commun.* **15** (2024) 3096.
153. L. Lami and B. Regula, arXiv:2307.11008 [math-ph, physics:quant-ph].
154. X. Wang and R. Duan, *Phys. Rev. Lett.* **119** (2017) 180506.
155. L. Lami and B. Regula, *Nat. Phys.* **19** (2023) 184.
156. D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal and A. V. Thapliyal, *Phys. Rev. A* **61** (2000) 062312.
157. R. F. Werner, *Phys. Rev. A* **40** (1989) 4277.
158. W. Dür, J. I. Cirac, M. Lewenstein and D. Bruß, *Phys. Rev. A* **61** (2000) 062313.
159. D. Đoković, *Entropy* **18** (2016) 216.
160. L. Pankowski, M. Piani, M. Horodecki and P. Horodecki, *IEEE Trans. Inf. Theory* **56** (2010) 4085.
161. L. Clarisse, *Quantum Inf. Comput.* **6** (2006) 539.
162. A. Müller-Hermes, D. Reeb and M. M. Wolf, *J. Math. Phys.* **57** (2016) 015202.
163. P. W. Shor, J. A. Smolin and B. M. Terhal, *Phys. Rev. Lett.* **86** (2001) 2681.
164. C. Spengler, M. Huber and B. C. Hiesmayr, *J. Math. Phys.* **53** (2012) 013501.

B.11. The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks

The Impact of Architecture and Cost Function on Dissipative Quantum Neural Networks

Tobias C. Sutter, Christopher Popp, and Beatrix C. Hiesmayr

University of Vienna, Faculty of Physics, Währingerstrasse 17, 1090 Vienna

Combining machine learning and quantum computation is a potential path towards powerful applications on quantum devices. Regarding this, quantum neural networks are a prominent approach. In this work, we present a novel architecture for dissipative quantum neural networks (DQNNs) in which each building block can implement any quantum channel, thus introducing a clear notion of universality suitable for the quantum framework. To this end, we reformulate DQNNs using isometries instead of conventionally used unitaries, thereby reducing the number of parameters in these models. We furthermore derive a versatile one-to-one parametrization of isometries, allowing for an efficient implementation of the proposed structure. Focusing on the impact of different cost functions on the optimization process, we numerically investigate the trainability of extended DQNNs. This unveils significant training differences among the cost functions considered. Our findings facilitate both the theoretical understanding and the experimental implementability of quantum neural networks.

1 Introduction

Classical machine learning (CML) and quantum computing are established computational paradigms. While the former has already proven valuable in widely used applications like large language models, the theoretically promised advantages of the latter [1] are yet to be confirmed experimentally. This is due to the experimental challenges accompanying the realization of

quantum computers [2]. Nonetheless, machine learning on quantum hardware, i.e., quantum machine learning (QML), promises several benefits, like reducing the complexity of specific machine learning algorithms [3]. One explicit model, often called dissipative quantum neural network (DQNN), has been proposed in Ref. [4]. It can be understood as a straightforward quantization of classical feedforward artificial neural networks. However, it manifestly does not contain nonlinearities, which are crucial for the universality (i.e., the ability to approximate any continuous function on a compact domain arbitrarily well) of its classical counterpart [5]. In this regard, it is essential to emphasize that the notions of universality for CML and QML may differ, and obtaining a quantum advantage (e.g., speed-up over any classical algorithm) may be only one of many reasonable goals of QML [6]. Furthermore, (linear) DQNNs are intriguing from a quantum information theoretic viewpoint as fundamental concepts like the Heisenberg uncertainty relation appear in their optimization process [7]. Despite the potential of QML, several factors affect the expressivity and trainability of these models: Besides quantum hardware [2] and data-related factors [8], the cost function and network architecture [9, 10], and entanglement within the network [11, 12] crucially influence a model's performance.

The aim of this contribution is twofold: We first extend the conventional DQNN architecture so that each building block satisfies a specific notion of universality and subsequently focus on the impact of the cost function on the training process. To avoid problems arising from the exponentially growing Hilbert space dimension, we concentrate on shallow DQNNs with a small output Hilbert space. Thus, we provide small-scale results for the proposed extended architecture. As an interesting use case, we mention that DQNNs with a single output qubit already allow

Tobias C. Sutter: tobias.christoph.sutter@univie.ac.at
 Christopher Popp: christopher.popp@univie.ac.at
 Beatrix C. Hiesmayr: beatrix.hiesmayr@univie.ac.at

us to infer three properties of the input state (one for each degree of freedom of the output state). This is sufficient for specific quantum information processing tasks like determining the input state's purity or the concurrence [13].

The work is structured as follows. We formally introduce DQNNs in Sec. 2.1 before reformulating them using isometries instead of unitaries in Sec. 2.1.2. We derive a composite parametrization of isometries to leverage the resulting reduction of the variational parameters. Sec. 2.1.3 discusses two distinct training approaches: One based on random state sampling and the other on the Choi state of a quantum channel. In Sec. 2.2, we propose an extension of the conventional DQNN architecture based on considerations about the universality of DQNNs. This ensures that each quantum perceptron has the power to implement a general quantum channel. Sec. 3 introduces various cost functions for mixed output and target states before we report our numerical results in Sec. 4. Here, we conduct numerical simulations to assess the impact of different cost functions on the optimization process of a minimal extended DQNN. Sec. 4.1 is concerned with learning randomly sampled quantum channels, while Sec. 4.2 investigates the trainability of the Werner channel. We conclude by discussing our results in Sec. 5.

2 Dissipative Quantum Neural Networks

Dissipative quantum neural networks (DQNNs) are a straightforward quantization of classical feedforward artificial neural networks, where the artificial neurons are replaced by quantum systems. Usually, these models' trainable weight and bias matrices are represented by variational unitary gates that are applied to the layers consecutively. During the training phase, the unitary parameters are adjusted to optimize a given cost function that compares the network's output to a desired target output (supervised learning). Due to the freedom of initializing the hidden and output layer neurons in fiducial quantum states, these unitaries can be considered as isometries. This reduces the degrees of freedom and, thus, the computation effort required for the optimization procedure of DQNNs. We derive a versatile one-to-one parametrization of isometries from the

composite parametrization of the unitary group $\mathcal{U}(d)$ [14, 15]. The details can be found in App. A.

Moreover, we define a DQNN as *quantum channel universal* if it can implement any completely positive and trace-preserving (CPTP) map from the input to the output state. This allows for a standardization of DQNNs and a meaningful performance comparison. These considerations lead to an extended DQNN architecture where each building block naturally implements a general CPTP map. In contrast to conventional DQNNs, a minimal version of our modified architecture (comprised of three neurons) is quantum channel universal. The isometry viewpoint also gives a straightforward interpretation of the training process: The network aims to learn the Stinespring representation of a target quantum channel by adjusting its isometry degrees of freedom.

2.1 Conventional DQNNs

As introduced in Ref. [4], the conventional architecture for DQNNs aims to mimic classical feedforward artificial neural networks: The artificial neurons are represented by d -dimensional quantum systems called qudits, and unitary interactions represent the weight and bias matrices. Choosing an architecture requires arranging these N quantum neurons into L layers, as visualized in Fig. 1. Layers 1 and L constitute the input and output layers, respectively, and layers 2 to $L - 1$ represent the hidden layers. Each layer $\ell \in \{1, \dots, L\}$ consists of n_ℓ neurons. Furthermore, we can formally assign a Hilbert space $\mathcal{H}_\ell = \bigotimes_{i=1}^{n_\ell} \mathcal{H}_\ell^{(i)}$ to each layer $\ell \in \{1, \dots, L\}$ of the network, where $\mathcal{H}_\ell^{(i)}$ is the Hilbert space of the i th neuron in layer ℓ .

2.1.1 Unitary Formulation

Initially, layers $\ell \geq 2$ are prepared in a fiducial state, e.g., the computational basis state $|0\rangle^{\otimes n_\ell} \in \mathcal{H}_\ell$, and layer 1 holds a generally mixed input quantum state $\rho_{\text{in}} \in \mathcal{D}(\mathcal{H}_1)$. The set $\mathcal{D}(\mathcal{H}_1)$ denotes the set of positive semi-definite linear operators mapping $\mathcal{H}_1$ into itself and satisfying $\text{Tr}(\rho_{\text{in}}) = 1$. Analogous to the weight and bias matrices in classical networks, neurons in adjacent layers (ℓ and $\ell + 1$) of DQNNs are connected by variational unitary transformations $U_k^{(\ell, \ell+1)}$, called (quantum) perceptrons. Here,

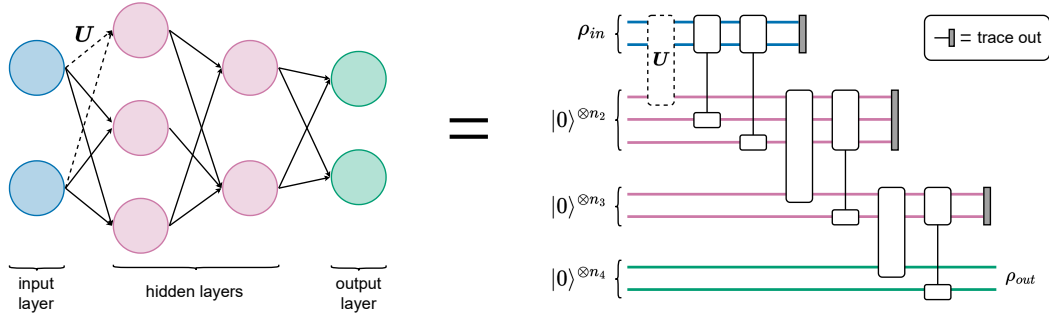


Figure 1: Conventional dissipative quantum neural network (DQNN) with $N = 9$, $n_1 = n_3 = n_4 = 2$, and $n_2 = 3$. It consists of input (blue), hidden (violet), and output layers (green). In the unitary formulation, the hidden and output layers are initialized in fiducial states. The black arrows in the left diagrams represent the unitary perceptrons and, thus, the information flow. Dashed lines indicate the first unitary $U = U_1^{(1,2)}$, while the full unitary is given by $U_2^{(3,4)} U_1^{(3,4)} U_2^{(2,3)} U_1^{(2,3)} U_3^{(1,2)} U_2^{(1,2)} U_1^{(1,2)}$. Hence, the parameter $k \in \{1, \dots, K_{\ell+1}\}$ that counts the neurons of each layer in (3) increases from the top to the bottom neurons. After the unitaries are applied, the input and hidden layers are traced out, leaving the network in the output state (1). In the quantum circuit diagram on the right, each horizontal line corresponds to one neuron (qudit), and time increases from the left to the right. The boxes represent the unitary perceptrons, successively applied in a specific order to adjacent layers.

$\ell \in \{1, \dots, L-1\}$ and $k \in \{1, \dots, K_{\ell+1}\}$, where $K_{\ell+1}$ is the total number of unitaries connecting layers ℓ and $\ell+1$. For simplicity, we assume that each unitary $U_k^{(\ell, \ell+1)}$ acts on all neurons of layer ℓ (global perceptron) and one neuron of layer $\ell+1$. Hence, $K_{\ell+1} = n_{\ell+1}$ and $U_k^{(\ell, \ell+1)} \in \mathcal{U}(\mathcal{H}_{\ell+1}^{(k)} \otimes \mathcal{H}_\ell)$, where $\mathcal{U}(\mathcal{H})$ denotes the set of unitary matrices acting on $\mathcal{H}$. After all unitaries are applied, layers 1 to $L-1$ are traced out, yielding output state $\rho_{\text{out}} \in \mathcal{D}(\mathcal{H}_L)$, given by

$$\rho_{\text{out}} = \text{Tr}_{1, \dots, L-1} \left[U (|0\rangle\langle 0|_{L, \dots, 2} \otimes \rho_{\text{in}}) U^\dagger \right], \quad (1)$$

where

$$|0\rangle_{L, \dots, 2} = |0\rangle_L \otimes |0\rangle_{L-1} \otimes \dots \otimes |0\rangle_2, \quad (2)$$

$$U = \prod_{\ell=1}^{L-1} \left(\prod_{k=1}^{K_{\ell+1}} U_{1-k+K_{\ell+1}}^{(\ell', \ell+1)} \right), \quad (3)$$

where $\ell' = L - \ell$, and $|0\rangle_\ell = \bigotimes_{i=1}^{n_\ell} |0\rangle_\ell^{(i)}$ with $|0\rangle_\ell^{(i)} \in \mathcal{H}_\ell^{(i)}$. The first product in (3) concerns the layers, while the second regards the unitaries within one layer. Two remarks are in order. First, our choice for arranging the Hilbert space throughout this work is reversed in the sense that we consider the total Hilbert space of the DQNN as $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_{L-1} \otimes \dots \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$. This is in preparation for using the composite parametriza-

tion of isometries for which this ordering is essential (cf. App. A). Second, we define the product in (3) as $\prod_{i=1}^n A_i = A_1 \cdot A_2 \cdots A_n$. Thereby, we ensure that the DQNN applies the unitaries layer-wise and ordered according to the label k (cf. Fig. 1). This is important because they generally do not commute within each layer.

We note in passing that this quantum machine learning ansatz crucially differs from its classical counterpart in that it does not involve any nonlinearities, which are essential for the universality property of classical feedforward artificial neural networks. However, linear transformations are sufficient for the notion of universality we consider in Sec. 2.2.

2.1.2 From Unitaries to Isometries

A different and computationally advantageous perspective on DQNNs can be adopted by considering the perceptrons not as unitary transformations but as isometries. To do so, the neurons in the layers $\ell \in \{2, \dots, L\}$ are not initialized in a fiducial state. Instead, the network's initial state is simply $\rho_{\text{in}} \in \mathcal{D}(\mathcal{H}_1)$. Subsequently, the perceptron isometries $V_k^{(\ell, \ell+1)} := U_k^{(\ell, \ell+1)} |0\rangle_{\ell+1}^{(k)} \in \text{Iso}(\mathcal{H}_\ell, \mathcal{H}_{\ell+1}^{(k)} \otimes \mathcal{H}_\ell)$ get applied sequentially, thus each enlarging the network's Hilbert space $\mathcal{H}$ by one neuron. The output state is

$$\rho_{\text{out}} = \text{Tr}_{1, \dots, L-1} \left[V \rho_{\text{in}} V^\dagger \right], \quad (4)$$

where

$$V = \prod_{\ell=1}^{L-1} \left(\prod_{k=1}^{K_{\ell'+1}} V_{1-k+K_{\ell'+1}}^{(\ell', \ell'+1)} \right), \quad (5)$$

with $\ell' = L - \ell$. One advantage of this formulation is that it illuminates the DQNN's implementation of a completely positive and trace-preserving (CPTP) map $\mathcal{E}_{\text{net}} : \mathcal{D}(\mathcal{H}_1) \rightarrow \mathcal{D}(\mathcal{H}_L)$, where V can be viewed as the network's Stinespring isometry [16]. We can therefore write $\rho_{\text{out}} = \mathcal{E}_{\text{net}}(\rho_{\text{in}})$, describing every transformation that is theoretically possible for DQNNs. Another advantage is that isometries have fewer degrees of freedom than unitaries. Hence, this reformulation also reduces the number of parameters to optimize during the network's training phase (described below). In particular, a unitary perceptron acting on d -dimensional input and hidden layer qudits has d^4 free parameters, while an isometry perceptron acting on the same qudits only has $d^2(2d-1)$. Already for $d = 4$, this more than halves the number of parameters to optimize.

However, a suitable variational parametrization of isometries is required to exploit this. Based on the composite parametrization (CP) of the unitary group [14, 15], we derive a corresponding one-to-one parametrization of isometries in App. A. As a result, any isometry $V_{\text{CP}} \in \text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$ can be written as

$$V_{\text{CP}} = \left[\prod_{m=0}^{d_1-1} \prod_{n=m+1}^{d_2-1} \Lambda_{m,n} \right] \left[\prod_{l=0}^{d_1-1} e^{iP_l \lambda_{ll}} \right] \mathbb{1}_{d_2 \times d_1}, \quad (6)$$

where $d_\ell = \dim(\mathcal{H}_\ell)$, $\{|i\rangle_\ell\}_{i=0}^{d_\ell-1}$ is a basis of $\mathcal{H}_\ell$, and

$$\mathbb{1}_{d_2 \times d_1} = \sum_{i=0}^{d_1-1} |i\rangle_2 \langle i|_1, \quad (7)$$

$$P_n = |n\rangle_2 \langle n|_2, \quad (8)$$

$$Y_{m,n} = -i|m\rangle_2 \langle n|_2 + i|n\rangle_2 \langle m|_2, \quad (9)$$

$$\Lambda_{m,n} = e^{iP_n \lambda_{n,m}} e^{iY_{m,n} \lambda_{m,n}}. \quad (10)$$

The set $\{\lambda_{m,n} | 0 \leq m, n < d_2, m < d_1 \vee n < d_1\}$ contains the $2d_1d_2 - d_1^2$ parameters of V_{CP} .

2.1.3 Gradient Optimization

The standard procedure for training the network toward implementing a desired target transfor-

mation $\mathcal{E}_{\text{tar}} : \mathcal{D}(\mathcal{H}_1) \rightarrow \mathcal{D}(\mathcal{H}_L)$ involves sampling a set of input states $\{\rho_{\text{in}}^{(i)}\}_{i=1}^{N_t}$. This random element can speed up the optimization process, similar to stochastic gradient descent in classical machine learning. However, the geometry of quantum state space is non-unique [17], so this scheme can suffer from choosing the “wrong” sampling method. For each element of the input state set, the corresponding network output state $\rho_{\text{out}}^{(i)} = \mathcal{E}_{\text{net}}(\rho_{\text{in}}^{(i)})$ and target output state $\rho_{\text{tar}}^{(i)} = \mathcal{E}_{\text{tar}}(\rho_{\text{in}}^{(i)})$ are computed.

It is imperative that $\mathcal{E}_{\text{tar}}$ is (close to) a CPTP map. Otherwise, the DQNN will inevitably fail in the training process as $\mathcal{E}_{\text{net}}$ is necessarily a quantum channel, and the perfect network satisfies $\mathcal{E}_{\text{net}}(\sigma) = \mathcal{E}_{\text{tar}}(\sigma)$ for all $\sigma \in \mathcal{D}(\mathcal{H}_1)$. Hence, a good strategy to avoid trainability issues is to ensure that the target transformation $\mathcal{E}_{\text{tar}}$ represents a quantum channel, i.e., is linear and CPTP.

To evaluate how well the network reproduces $\mathcal{E}_{\text{tar}}$, a cost/loss function $C : \mathcal{D}(\mathcal{H}_L) \times \mathcal{D}(\mathcal{H}_L) \rightarrow \mathbb{R}$ is applied to each element of $\{(\rho_{\text{out}}^{(i)}, \rho_{\text{tar}}^{(i)})\}_{i=1}^{N_t}$. The total cost of the network is the average cost over all training states,

$$C_{\text{tot}} = \frac{1}{N_t} \sum_{i=1}^{N_t} C(\rho_{\text{tar}}^{(i)}, \rho_{\text{out}}^{(i)}). \quad (11)$$

The function C is usually a similarity or distinguishability measure on the output state space. We discuss potential candidates in Sec. 3 and their impact on the training in Sec. 4.

Once a cost function is chosen, the network trains by updating the variational isometry parameters $\{\lambda_\mu\}_\mu$ according to gradient descent (if C is a distinguishability measure) or gradient ascent (if C is a similarity measure). This requires taking the derivative of (11) with respect to every λ_μ ,

$$\frac{\partial C_{\text{tot}}}{\partial \lambda_\mu} = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\partial}{\partial \lambda_\mu} C(\rho_{\text{tar}}^{(i)}, \rho_{\text{out}}^{(i)}), \quad (12)$$

and adjust the network's parameters according to, e.g., the ADAM optimizer [18]. Repeating this feedback loop of computing the network's output state for each training input state and updating the isometry parameters leads to a (local) optimum in the cost function landscape. We call this scheme random state training.

Due to the fact that DQNNs can only realize quantum channels, a different optimization

method can be considered. It does not rely on random state sampling but the Choi representation of quantum channels [19, 20]. The maps $\mathcal{E}_{\text{net}}$ and $\mathcal{E}_{\text{tar}}$ are identified with their respective Choi state [16]

$$J(\mathcal{E}_{\text{net}/\text{tar}}) = \frac{1}{d_1} \sum_{i,j=0}^{d_1-1} \mathcal{E}_{\text{net}/\text{tar}}(|i\rangle_1 \langle j|_1) \otimes |i\rangle_1 \langle j|_1. \quad (13)$$

Operationally, the state $J(\mathcal{E}_{\text{net}})$ can be created by sending one half of the maximally entangled state $|\Omega\rangle = 1/\sqrt{d_1} \sum_{i=0}^{d_1-1} |i\rangle \otimes |i\rangle$ through the network. Because the Choi representation is unique, the DQNN perfectly represents the target transformation if and only if $J(\mathcal{E}_{\text{net}}) = J(\mathcal{E}_{\text{tar}})$. Thus, defining a cost function $C : \mathcal{D}(\mathcal{H}_L \otimes \mathcal{H}_1) \times \mathcal{D}(\mathcal{H}_L \otimes \mathcal{H}_1) \rightarrow \mathbb{R}$, we can optimize the DQNN by computing

$$\frac{\partial C(J(\mathcal{E}_{\text{tar}}), J(\mathcal{E}_{\text{net}}))}{\partial \lambda_\mu}, \quad (14)$$

and using gradient optimization as before. We refer to this method as Choi training. The drawbacks are that the target channel must be entirely known, and the cost function acts on a larger Hilbert space. However, it does not suffer from a potentially unsuitable sampling of (finitely many) input quantum states and thus allows more objective trainability statements. Therefore, we use it to benchmark the performance of different cost functions in Sec. 4.

2.2 Extended DQNNs

It is clear from (4) that DQNNs implement a CPTP map from $\mathcal{D}(\mathcal{H}_1)$ to $\mathcal{D}(\mathcal{H}_L)$. Consequently, the most general learnable transformation $\mathcal{E}_{\text{tar}}$ is also of this kind. This leads us to regard a DQNN as *quantum channel universal* if it can realize any CPTP map $\mathcal{E} : \mathcal{D}(\mathcal{H}_1) \rightarrow \mathcal{D}(\mathcal{H}_L)$. If the DQNN can realize such a map only approximately, we say it is a universal quantum channel approximator. Note that this definition only covers linear maps from the input to the output state; classical post-processing is needed to obtain nonlinear functions of the input state. Furthermore, it contrasts universal quantum computation, which only regards approximating unitary transformations.

Conventional DQNNs (Sec. 2.1) do not necessarily have a structure that enables quantum

channel universality. Take, e.g., a network consisting of one input, one hidden, and one output layer neuron. There are two perceptrons in such a network, and according to (4) and (5), the output state is

$$\rho_{\text{out}} = \text{Tr}_{1,2} [V \rho_{\text{in}} V^\dagger], \quad (15)$$

where $V = V_1^{(2,3)} V_1^{(1,2)}$. However, a Stinespring isometry $V_{\mathcal{E}}$ of a quantum channel $\mathcal{E} : \mathcal{D}(\mathcal{H}_1) \rightarrow \mathcal{D}(\mathcal{H}_3)$ can generally not be written as the product of two isometries, i.e., $V_{\mathcal{E}} \neq V_1^{(2,3)} V_1^{(1,2)}$. Hence, this network is not quantum channel universal.

For this reason, we extend the input-hidden-output layer structure of DQNNs by adding ancilla layers. Every perceptron adds to the network not only one hidden or output neuron but also an ancilla neuron, which is subsequently traced out (see Fig. 2). Hence, the isometries are given by $V_k^{(\ell, \ell+1, \ell+2)} \in \text{Iso}(\mathcal{H}_\ell, \mathcal{H}_{\ell+2}^{(k)} \otimes \mathcal{H}_{\ell+1}^{(k)} \otimes \mathcal{H}_\ell)$ and the output state ρ_{out} results from (4) together with

$$V = \prod_{\ell=1}^{(L-1)/2} \left(\prod_{k=1}^{K_{\ell'+2}} V_{1-k+K_{\ell'+2}}^{(\ell', \ell'+1, \ell'+2)} \right), \quad (16)$$

where $\ell' = L - 2\ell$. The additional degree of freedom ensures that a minimal network consisting of an input, an ancilla, and an output layer connected by a single perceptron is quantum channel universal, provided that $\dim(\mathcal{H}_2) = \dim(\mathcal{H}_3)$. The isometry viewpoint allows a straightforward interpretation of the training process: Given a target quantum channel, the network aims to learn its Stinespring representation.

This minimal extended DQNN can be considered the blueprint for the perceptrons of larger networks comprising multiple layers with more than one neuron each (see Fig. 3). Consequently, each perceptron of an extended DQNN is quantum channel universal. Note, however, that this does not ensure that the whole network also has this property.

3 Choices of Cost Functions

The trainability of DQNNs depends on the cost function used in the optimization process [9, 10]. Thus, a suitable cost function is essential for designing a useful DQNN. In principle, one may choose any reasonable function $C : \mathcal{D}(\mathcal{H}_L) \times$

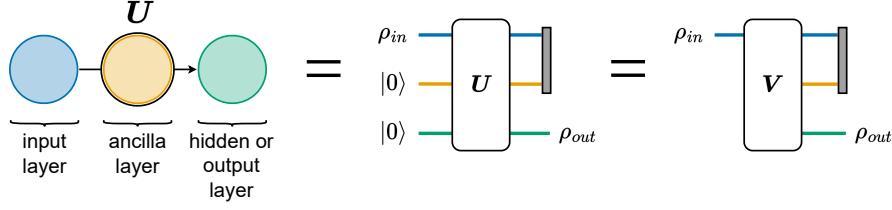


Figure 2: The minimal version of an extended DQNN consists of three neurons and can be viewed as the prototype for perceptrons of larger extended DQNNs. The black ring and arrow on the left denote the perceptron. It can learn any quantum channel from the input to the hidden/output neurons if the dimensions of the ancilla and the hidden/output neurons coincide. The middle and right figures show the quantum circuit of the perceptron in the unitary and isometry formulation, respectively, for which we have $V = U(|0\rangle_3 \otimes |0\rangle_2 \otimes \mathbb{1}_1)$. For general CPTP maps, V does not factorize, i.e., $V \neq V_1^{(2,3)} V_1^{(1,2)}$.

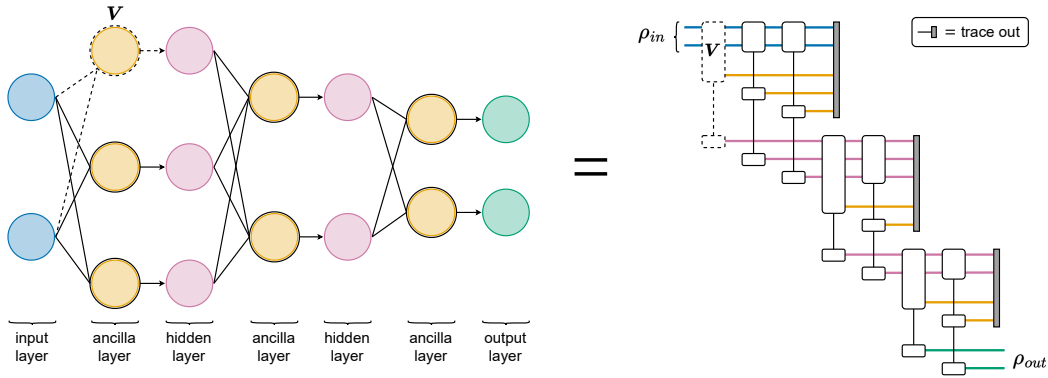


Figure 3: An extended version of the network in Fig. 3, consisting of 16 instead of 9 neurons and thus almost doubling the size. It comprises input (layer 1; blue), ancilla (layers 2, 4, and 6; gold), hidden (layers 3 and 5; violet), and output layers (layer 7; green). This comes with the benefit that every perceptron can implement a general CPTP map. The dashed lines indicate the first isometry perceptron $V = V_1^{(1,2,3)}$, while the full isometry (16) is given by $V_2^{(5,6,7)} V_1^{(5,6,7)} V_2^{(3,4,5)} V_1^{(3,4,5)} V_3^{(1,2,3)} V_2^{(1,2,3)} V_1^{(1,2,3)}$. The quantum circuit diagram on the right utilizes the isometric formulation of extended DQNNs.

$\mathcal{D}(\mathcal{H}_L) \rightarrow \mathbb{R}$. However, we focus on distance and similarity measures on $\mathcal{D}(\mathcal{H}_L)$ as they align with the usual “cost” or “reward” imposed for assessing the network’s output. In this section, we present several candidates for C that are applicable to mixed output and target states of a DQNN. Special attention is paid to experimental measurability and information-theoretic interpretation of the presented quantities. Furthermore, we need the gradient of C to optimize a DQNN using gradient descent/ascent. This involves taking derivatives of C with respect to the variational parameters of the network. We present analytical expressions for this in App. B whenever possible.

Typically, *distance measures* $D : \mathcal{D}(\mathcal{H}) \times \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$ on the space of density matrices are defined by the following properties: D must be

nonnegative ($D(\rho, \sigma) \geq 0$), symmetric ($D(\rho, \sigma) = D(\sigma, \rho)$), zero if and only if the states are equal ($D(\rho, \sigma) = 0 \Leftrightarrow \rho = \sigma$), and satisfy the triangle inequality ($D(\rho, \sigma) \leq D(\rho, \chi) + D(\chi, \sigma)$). Additionally, Ref. [21] proposes that a quantum distance measure should satisfy the so-called *data-processing inequality*

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma), \quad (17)$$

where $\mathcal{E}$ is any CPTP map. This allows using D to quantify entanglement in a meaningful way.

Nevertheless, dropping some of these properties in favor of a clear operational interpretation can help solve specific problems. In this case, one considers *divergences*, which are not required to be symmetric or satisfy the triangle inequality. The essential property is that they satisfy the

data-processing inequality. They find meaning, e.g., in asymmetric hypothesis testing scenarios, by quantifying how distinguishable one state is from another.

Similarly, *fidelities* are a pivotal similarity measure between two quantum states. By definition, every fidelity function F must satisfy a set of axioms [22]. One of them demands that if $F(\rho, \sigma)$ is a fidelity for $\rho, \sigma \in \mathcal{D}(\mathcal{H}_L)$, it reduces to $F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\rho|\psi\rangle$ if $\sigma = |\psi\rangle\langle\psi|$ is a pure state. Hence, F generalizes the notion of the transition probability of two pure states to the mixed case. Despite this, the axioms do not single out a unique quantum fidelity.

The remainder of this section introduces several well-known distances, fidelities, and one divergence, representing potential cost functions.

Hilbert-Schmidt Distance. The Hilbert-Schmidt inner product $\langle\rho, \sigma\rangle_{\text{HS}} = \text{Tr}(\rho^\dagger\sigma)$ induces a norm on $\mathcal{D}(\mathcal{H})$. This can be used to define the *Hilbert-Schmidt distance*,

$$D_{\text{HS}}(\rho, \sigma) = \sqrt{\text{Tr}((\rho - \sigma)^2)}, \quad (18)$$

where we used that $(\rho - \sigma)^\dagger = \rho - \sigma$ for $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. It has a clear operational meaning as an information distance between two quantum states [23]. One advantage of this cost function choice is that it is readily measurable on a quantum computer using the SWAP test [24, 25]. However, it violates the data-processing inequality [26].

Trace Distance. The trace distance is given by

$$D_{\text{Tr}}(\rho, \sigma) = \frac{1}{2} \text{Tr}(|\rho - \sigma|), \quad (19)$$

where $|A| = \sqrt{A^\dagger A}$. It can be interpreted as follows: Given two quantum states ρ and σ , each with probability 1/2, the trace distance quantifies the lowest error probability for distinguishing them upon performing any POVM [27]. Furthermore, the trace distance satisfies the data-processing inequality [28].

Generalized p -Fidelities. The p -fidelity [29] is a general approach that covers multiple interesting similarity and distance measures. It is defined as

$$F_p(\rho, \sigma) = \frac{\|\sqrt{\sigma}\sqrt{\rho}\|_p^2}{\max(\|\sigma\|_p^2, \|\rho\|_p^2)}, \quad (20)$$

where the p -norm is $\|A\|_p := \text{tr}((A^\dagger A)^{p/2})^{1/p}$. This satisfies all fidelity axioms for $p \geq 1$.

We consider two special cases. For $p = 1$, we obtain the *Uhlmann-Jozsa fidelity* [22, 30]

$$F_1(\rho, \sigma) = \text{Tr}\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2, \quad (21)$$

which satisfies the data-processing inequality [31]. Furthermore, Uhlmann's theorem [30] allows to connect F_1 to the Bures metric, a natural Riemannian metric on the space of mixed quantum states [17]. The *Bures distance* is given by

$$D_1(\rho, \sigma) = \sqrt{2\left(1 - \sqrt{F_1(\rho, \sigma)}\right)}. \quad (22)$$

Despite having a solid theoretic foundation, F_1 and D_1 are challenging to measure experimentally.

The case $p = 2$ leads to the *Hilbert-Schmidt fidelity*, given by

$$F_2(\rho, \sigma) = \frac{\text{Tr}(\rho\sigma)}{\max\{\text{Tr}(\rho^2), \text{Tr}(\sigma^2)\}}. \quad (23)$$

Contrary to the Hilbert-Schmidt inner product, it satisfies the fidelity axioms. The main advantage of F_2 is that it is easily calculable and experimentally measurable, as demonstrated in Ref. [32]. However, one downside is that it violates the data-processing inequality [29]. Lastly, as shown in Ref. [29], one can define a distance based on (23) by

$$D_2(\rho, \sigma) = \sqrt{2(1 - F_2(\rho, \sigma))}. \quad (24)$$

To avoid confusion with the Hilbert-Schmidt distance (18), we refer to (24) as the D_2 distance.

Quantities from Hypothesis Testing.

Quantum hypothesis testing is a fundamental quantum processing task where an observer receives a quantum system known to be in one of two possible states, and the goal is to correctly guess which state it is after performing a POVM measurement [33].

This setting gives rise to two fundamental asymptotic quantities. The first one is the *Quantum Chernoff Bound* [34, 35]

$$F_{\text{QCB}}(\rho, \sigma) = \min_{0 \leq s \leq 1} \text{tr}(\rho^s \sigma^{1-s}). \quad (25)$$

Its interpretation is the following: The quantity $-\log_2(F_{\text{QCB}}(\rho, \sigma))$ is the optimal asymptotic error exponent for symmetric hypothesis testing, i.e., quantum state discrimination. Additionally, it satisfies the data-processing inequality [34].

The second is the *quantum relative entropy*. It is defined as

$$D_{\text{QRE}}(\rho||\sigma) = \text{Tr}(\rho \log(\rho) - \rho \log(\sigma)), \quad (26)$$

where $\log$ is the matrix logarithm. It is a divergence, not a distance measure, as it is not symmetric under exchanging ρ and σ . Nonetheless, it can be used for state discrimination because it is non-negative, and zero if and only if $\rho = \sigma$ due to Klein's inequality [36]. It also satisfies the data-processing inequality [37]. The quantum relative entropy gains operational meaning from the quantum Stein's lemma as the optimal rate in asymmetric quantum hypothesis testing [33].

4 Numerical Trainability Results

To demonstrate the trainability of the extended DQNN architecture and to quantify the effect of different cost functions on the learning rate of quantum neural networks, we conduct numerical simulations of a minimal network consisting of three qubits (see Fig. 2). The network implements the quantum channel $\mathcal{E}_{\text{net}}$, and its isometry parameters are initialized randomly but close to zero. This corresponds to canonically embedding the input state in the larger Hilbert space of the whole network, with an additional small numerical perturbation. We found that without this minor disturbance of the initial parameters, the convergence to a cost function optimum is slower. A similar initialization strategy mitigates barren plateaus in variational quantum circuits [38]. The training objective is to learn a target quantum channel $\mathcal{E}_{\text{tar}}$. Due to the network's extended structure, the DQNN we consider is quantum channel universal for qubit-qubit channels, i.e., it can represent any such channel exactly. This avoids the problem of $\mathcal{E}_{\text{tar}}$ being impossible to learn. The optimization is done with Choi and random state training separately (discussed in detail in Sec. 2.1.3). In both cases, we use the ADAM algorithm for gradient optimization [18] for 1000 training iterations.

Objective assessment of the cost functions' performance requires a suitable and independent distinguishability measure for $\mathcal{E}_{\text{net}}$ and $\mathcal{E}_{\text{tar}}$. A useful quantity is the *diamond distance* $\|\mathcal{E}_{\text{net}} - \mathcal{E}_{\text{tar}}\|_{\diamond}$ [39]. It is induced by the *diamond norm* [40]

$$\|\mathcal{E}\|_{\diamond} = \max_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \|(\mathbb{1}_d \otimes \mathcal{E})(\rho)\|_1, \quad (27)$$

where $\mathcal{E}$ is a CPTP map acting on $\mathcal{H}$, $d = \dim(\mathcal{H})$, and $\|A\|_1 = \text{Tr}(|A|)$ denotes the trace norm. It can be interpreted as the best-case distinguishability of the output of the two channels when applied to part of a quantum state. Further note that $D_{\text{Tr}}(J(\mathcal{E}_{\text{tar}}), J(\mathcal{E}_{\text{net}})) \leq \frac{1}{2} \|\mathcal{E}_{\text{net}} - \mathcal{E}_{\text{tar}}\|_{\diamond}$. We employ a numerical implementation of the diamond distance using a Monte Carlo algorithm described in Ref. [41].

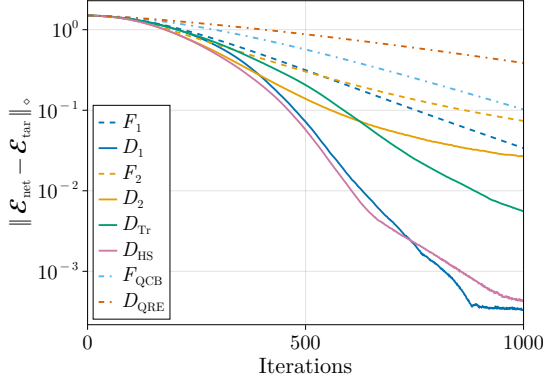
In Sec. 4.1, we optimize the DQNN using randomly sampled target quantum channels, while in Sec. 4.2 we consider the highly symmetric Werner channel as the target objective.

4.1 Learning Random Channels

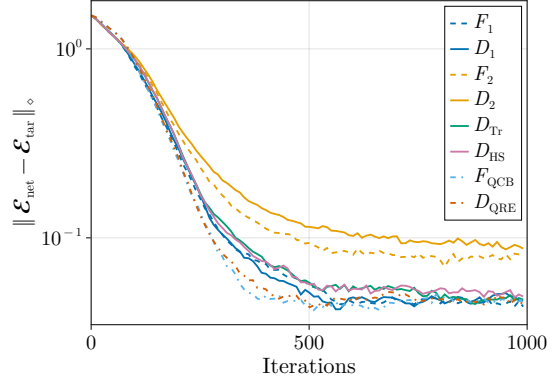
To determine the performance of the different cost functions, we begin the numerical analysis by training the DQNN using 100 random qubit-qubit target channels $\mathcal{E}_{\text{tar}}$. The channel sampling is implemented using [42, 43].

Our first benchmark comes from Choi training. Fig. 4a shows the average diamond distance $\|\mathcal{E}_{\text{net}} - \mathcal{E}_{\text{tar}}\|_{\diamond}$ for 1000 optimization iterations. After the training, the distances D_1 , D_2 , D_{Tr} , and D_{HS} perform better than any fidelity. However, the learning rate of F_1 suggests that it may surpass D_2 with additional training rounds. Nonetheless, D_1 and D_{HS} achieve the best result with a mean diamond distance of 3.43×10^{-4} and 4.55×10^{-4} , respectively. Interestingly, F_{QCB} and D_{QRE} , both related to asymptotic hypothesis testing, lead to the least optimized networks after the training. In these cases, the final mean diamond distance is 0.102 and 0.386, respectively.

The second benchmark is obtained using random state training. The input training states are sampled using the Hilbert-Schmidt distribution on the set of quantum states (implemented using [42, 44]). For the training, we use eight batches containing four states each. Once a cost optimum is reached for a batch, 32 new training states are generated. Fig. 4b shows the mean diamond distance between $\mathcal{E}_{\text{tar}}$ and $\mathcal{E}_{\text{net}}$. The convergence



(a) Mean diamond distance for Choi training.



(b) Mean diamond distance for random state training.

Figure 4: The plots show the mean diamond distance $\|\mathcal{E}_{\text{net}} - \mathcal{E}_{\text{tar}}\|_{\diamond}$ for 1000 training iterations, averaged over 100 random target channels $\mathcal{E}_{\text{tar}}$. (a): Using Choi training, the best-performing cost functions are D_1 followed by D_{HS} , reaching a mean diamond distance of less than 10^{-3} . (b): Training the network with randomly sampled input states leads to faster convergence to a cost optimum. However, this optimum is worse, reaching only a mean diamond distance of around 5×10^{-2} for all cost functions except F_2 and D_2 .

to the cost function optimum is faster but does not reach the same values as the Choi training. Specifically, it converges to about 5×10^{-2} for almost all examined cost functions, the exceptions being F_2 and D_2 , which perform significantly worse than the others.

4.2 Learning the Werner Channel

Lastly, we investigate the trainability of the Werner channel, given by

$$\mathcal{E}_{\text{W},\alpha}(\rho) = \frac{1}{\alpha + d} \left(\text{Tr}(\rho) \mathbb{1}_d + \alpha \rho^T \right), \quad (28)$$

where $\alpha \in [-1, 1]$, $\rho \in \mathcal{D}(\mathcal{H})$, $d = \dim(\mathcal{H})$, and ρ^T denotes the transpose of ρ . The name stems from the fact that the Choi state $J(\mathcal{E}_{\text{W},\alpha})$ is the Werner state [45], an exceptionally symmetric bipartite quantum state with a deep connection to the foundations of quantum theory. The Werner channel inherits many interesting features from its Choi state. For example, it has full Kraus rank for $\alpha \in (-1, 1)$, is unital, mixed-unitary for $d = 2$ [16], and the output state is generally highly mixed [46]. Furthermore, the Werner channel is entanglement breaking for $\alpha \in [-\frac{1}{d}, 1]$ as the Werner state is separable for this parameter region. The case $\alpha = 0$ corresponds to the completely depolarizing channel, outputting the maximally mixed state $\frac{1}{d} \mathbb{1}_d$ for any input state.

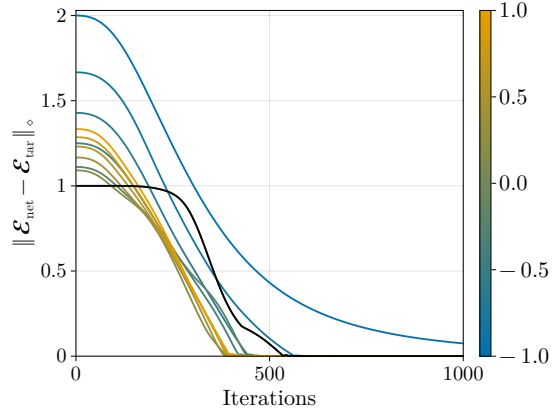


Figure 5: Learning the Werner channel $\mathcal{E}_{\text{W},\alpha}$ using Choi training with the Hilbert-Schmidt distance (18) as the cost function. The colors indicate the value of $\alpha \in [-1, 1]$. The completely depolarizing channel $\mathcal{E}_{\text{W},0}$ is highlighted in black.

Fig. 5 depicts the optimization of the minimal extended DQNN using Choi training and the Hilbert-Schmidt distance cost function (19) (other cost functions show similar behavior). We find that the convergence properties correlate with α : The higher this value, the faster the convergence to a small diamond distance. While the diamond distance for every $\mathcal{E}_{\text{W},\alpha}$ with $-0.7 \leq \alpha \leq 1$ (except $\alpha = 0$) is at most 6.7×10^{-3} after 500 training iterations, $\mathcal{E}_{\text{W},-1}$ only achieves a value of 0.075 after 1000 rounds. Interest-

ingly, the network seems to have problems finding an optimum in the cost landscape for $\alpha = 0$: The diamond distance decreases rapidly only after around 350 optimization steps.

5 Discussion and Conclusion

In this contribution, we developed an extension of the conventional dissipative quantum neural network architecture so that the perceptrons realize general quantum channels and investigated the impact of the cost function on the optimization process. In particular, we found that using isometries instead of unitaries in formulating DQNNs considerably reduces the number of parameters to optimize during training. To leverage this, we derived a versatile one-to-one composite parametrization of isometries. Besides the established way of using randomly sampled states for training, we presented a different training method based on the network’s Choi state. The main advantage is that the optimization does not rely on the sampling method (which requires choosing a non-unique geometry of quantum states), thus allowing more objective trainability statements. However, the target channel must be known entirely, and the cost function is applied to states with a larger Hilbert space dimension. This Choi approach distinguishes the quantum from the classical version of feed-forward neural networks, for which random inputs are required. We then defined a DQNN as *quantum channel universal* if it can learn arbitrary quantum channels from the input to the output state. Based on this, we argued for extending the conventional architecture by adding ancilla neurons to increase its expressivity. This way, the individual building blocks of (large) networks are quantum channel universal at the prize of increasing their size.

We simulated a minimal extended network consisting of three qubits and one perceptron to evaluate the influence of different cost functions on gradient optimization. The first objective was to learn random quantum channels to obtain insight into the general convergence behavior. Using Choi training, we found that the Hilbert-Schmidt and Bures distance performed best. Due to the fact that the former is easily calculable and readily measurable on quantum hardware, we suggest this to be the preferred cost function for Choi training. Furthermore, as its compu-

tation only involves functions that are at most quadratic in the quantum states, shadow tomography via randomized measurements [47, 48] is an alternative to full quantum state tomography [49, 50]. For random state training, almost all cost functions performed equally well. Nonetheless, for the Hilbert-Schmidt and Bures distance, the final distinguishability between the network and target channel was about two orders of magnitude greater than for Choi training. However, the convergence to an optimum is faster than for Choi training. This can be interpreted as the DQNN showing signs of barren plateaus for Choi training (i.e., gradients that vanish exponentially with the Hilbert space dimension) due to the cost function acting on a larger Hilbert space [9, 10]. This well-known trainability issue is not exclusive to gradient-based optimization, which we used in this work, but also appears in gradient-free schemes [51]. The presented results suggest that this phenomenon does not affect all cost functions equally (compare, e.g., the differences between Choi and random state training for the Hilbert-Schmidt distance and the quantum relative entropy, respectively). This raises the question of what properties a cost function must have to be less prone to barren plateaus. Our findings indicate that satisfying the data-processing inequality is not the decisive factor.

Lastly, we studied the trainability of the Werner channel. The results indicate a correlation between the learning rate and the Werner channel’s parameter. This suggests a connection between the trainability and the target channel’s properties. For example, the optimization takes less iterations if the channel is entanglement breaking. The exception is the completely depolarizing channel, for which the network has initial difficulties finding an optimum in the cost function landscape.

In conclusion, our results shed new light on two crucial aspects of quantum neural network design: architecture and cost function. We believe that the isometry formulation of extended DQNNs will aid in the theoretical development of this growing field. Furthermore, having found a suitable and readily measurable cost function will influence the experimental realization of quantum machine learning models.

Acknowledgments

T.C.S. wants to thank Felix Hitzelhammer for valuable discussions and comments. B.C.H. and C.P. acknowledge gratefully that this research was funded in whole, or in part, by the Austrian Science Fund (FWF) project P36102-N (Grant DOI: 10.55776/P36102). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

References

- [1] John Preskill. “Quantum computing and the entanglement frontier” (2012). arXiv:1203.5813 [quant-ph].
- [2] John Preskill. “Quantum Computing in the NISQ era and beyond”. *Quantum* **2**, 79 (2018).
- [3] M. Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J. Coles. “Challenges and opportunities in quantum machine learning”. *Nature Computational Science* **2**, 567–576 (2022).
- [4] Kerstin Beer, Dmytro Bondarenko, Terry Farrelly, Tobias J. Osborne, Robert Salzmann, Daniel Scheiermann, and Ramona Wolf. “Training deep quantum neural networks”. *Nature Communications* **11**, 808 (2020).
- [5] Kurt Hornik. “Approximation capabilities of multilayer feedforward networks”. *Neural Networks* **4**, 251–257 (1991).
- [6] Maria Schuld and Nathan Killoran. “Is Quantum Advantage the Right Goal for Quantum Machine Learning?”. *PRX Quantum* **3**, 030101 (2022).
- [7] Beatrix C. Hiesmayr. “A quantum information theoretic view on a deep quantum neural network”. *AIP Conference Proceedings* **3061**, 020001 (2024).
- [8] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. “Effect of data encoding on the expressive power of variational quantum-machine-learning models”. *Physical Review A* **103**, 032430 (2021).
- [9] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. “Cost function dependent barren plateaus in shallow parametrized quantum circuits”. *Nature Communications* **12**, 1791 (2021).
- [10] Kunal Sharma, M. Cerezo, Lukasz Cincio, and Patrick J. Coles. “Trainability of Dissipative Perceptron-Based Quantum Neural Networks”. *Physical Review Letters* **128**, 180505 (2022).
- [11] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe. “Entanglement-Induced Barren Plateaus”. *PRX Quantum* **2**, 040316 (2021).
- [12] Taylor L. Patti, Khadijeh Najafi, Xun Gao, and Susanne F. Yelin. “Entanglement devised barren plateau mitigation”. *Physical Review Research* **3**, 033090 (2021).
- [13] William K. Wootters. “Entanglement of Formation of an Arbitrary State of Two Qubits”. *Physical Review Letters* **80**, 2245–2248 (1998).
- [14] Christoph Spengler, Marcus Huber, and Beatrix C Hiesmayr. “A composite parameterization of unitary groups, density matrices and subspaces”. *Journal of Physics A: Mathematical and Theoretical* **43**, 385306 (2010).
- [15] Christoph Spengler, Marcus Huber, and Beatrix C. Hiesmayr. “Composite parameterization and Haar measure for all unitary and special unitary groups”. *Journal of Mathematical Physics* **53**, 013501 (2012).
- [16] John Watrous. “The Theory of Quantum Information”. *Cambridge University Press*. (2018). 1 edition.
- [17] Ingemar Bengtsson and Karol Zyczkowski. “Geometry of Quantum States: An Introduction to Quantum Entanglement”. *Cambridge University Press*. Cambridge (2006).
- [18] Diederik P. Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization” (2017). arXiv:1412.6980 [cs].
- [19] Man-Duen Choi. “Completely positive linear maps on complex matrices”. *Linear Algebra and its Applications* **10**, 285–290 (1975).
- [20] A. Jamiołkowski. “Linear transformations which preserve trace and positive semidefiniteness of operators”. *Reports on Mathematical Physics* **3**, 275–278 (1972).
- [21] V. Vedral, M. B. Plenio, M. A. Rippin,

- and P. L. Knight. “Quantifying Entanglement”. *Physical Review Letters* **78**, 2275–2279 (1997).
- [22] Richard Jozsa. “Fidelity for Mixed Quantum States”. *Journal of Modern Optics* **41**, 2315–2323 (1994).
- [23] Jinhyoung Lee, M. S. Kim, and Časlav Brukner. “Operationally Invariant Measure of the Distance between Quantum States by Complementary Measurements”. *Physical Review Letters* **91**, 087902 (2003).
- [24] Adriano Barenco, André Berthiaume, David Deutsch, Artur Ekert, Richard Jozsa, and Chiara Macchiavello. “Stabilization of Quantum Computations by Symmetrization”. *SIAM Journal on Computing* **26**, 1541–1557 (1997).
- [25] Harry Buhrman, Richard Cleve, John Watrous, and Ronald De Wolf. “Quantum Fingerprinting”. *Physical Review Letters* **87**, 167902 (2001).
- [26] Masanao Ozawa. “Entanglement measures and the Hilbert–Schmidt distance”. *Physics Letters A* **268**, 158–160 (2000).
- [27] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. “Distance measures to compare real and ideal quantum processes”. *Physical Review A* **71**, 062310 (2005).
- [28] Mary Beth Ruskai. “Beyond strong subadditivity? improved bounds on the contraction of generalized relative entropy”. *Reviews in Mathematical Physics* **06**, 1147–1161 (1994).
- [29] Yeong-Cherng Liang, Yu-Hao Yeh, Paulo E M F Mendonça, Run Yan Teh, Margaret D Reid, and Peter D Drummond. “Quantum fidelity measures for mixed states”. *Reports on Progress in Physics* **82**, 076001 (2019).
- [30] A. Uhlmann. “The “transition probability” in the state space of a $*$ -algebra”. *Reports on Mathematical Physics* **9**, 273–279 (1976).
- [31] Howard Barnum, Carlton M. Caves, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher. “Noncommuting Mixed States Cannot Be Broadcast”. *Physical Review Letters* **76**, 2818–2821 (1996).
- [32] Andreas Elben, Benoît Vermersch, Rick Van Bijnen, Christian Kokail, Tiff Brydges, Christine Maier, Manoj K. Joshi, Rainer Blatt, Christian F. Roos, and Peter Zoller. “Cross-Platform Verification of Intermediate Scale Quantum Devices”. *Physical Review Letters* **124**, 010504 (2020).
- [33] Sumeet Khatri and Mark M. Wilde. “Principles of Quantum Communication Theory: A Modern Approach” (2024). arXiv:2011.04672 [quant-ph].
- [34] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, Ll. Masanes, A. Acín, and F. Verstraete. “Discriminating States: The Quantum Chernoff Bound”. *Physical Review Letters* **98**, 160501 (2007).
- [35] Michael Nussbaum and Arleta Szkoła. “The Chernoff lower bound for symmetric quantum hypothesis testing”. *The Annals of Statistics* **37**, 1040–1057 (2009).
- [36] O. Klein. “Zur quantenmechanischen Begründung des zweiten Hauptsatzes der Wärmelehre”. *Zeitschrift für Physik* **72**, 767–775 (1931).
- [37] Göran Lindblad. “Completely positive maps and entropy inequalities”. *Communications in Mathematical Physics* **40**, 147–151 (1975).
- [38] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti. “An initialization strategy for addressing barren plateaus in parametrized quantum circuits”. *Quantum* **3**, 214 (2019).
- [39] B. Rosgen and J. Watrous. “On the hardness of distinguishing mixed-state quantum computations”. In 20th Annual IEEE Conference on Computational Complexity (CCC’05). Pages 344–354. (2005).
- [40] A Yu Kitaev. “Quantum computations: algorithms and error correction”. *Russian Mathematical Surveys* **52**, 1191–1249 (1997).
- [41] Giuliano Benenti and Giuliano Strini. “Computing the distance between quantum channels: usefulness of the Fano representation”. *Journal of Physics B: Atomic, Molecular and Optical Physics* **43**, 215508 (2010).
- [42] Piotr Gawron, Dariusz Kurzyk, and Łukasz Paweł. “QuantumInformation.jl—A Julia package for numerical computation in quantum information theory”. *PLOS ONE* **13**, e0209358 (2018).
- [43] Wojciech Bruzda, Valerio Cappellini, Hans-Jürgen Sommers, and Karol Życzkowski. “Random quantum operations”. *Physics Letters A* **373**, 320–324 (2009).
- [44] Karol Życzkowski and Hans-Jürgen Som-

- mers. “Induced measures in the space of mixed quantum states”. *Journal of Physics A: Mathematical and General* **34**, 7111 (2001).
- [45] Reinhard F. Werner. “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. *Physical Review A* **40**, 4277–4281 (1989).
- [46] Cécilia Lancien and Andreas Winter. “Approximating quantum channels by completely positive maps with small Kraus rank”. *Quantum* **8**, 1320 (2024).
- [47] Scott Aaronson. “Shadow tomography of quantum states”. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. Pages 325–338. Los Angeles CA USA (2018). ACM.
- [48] Hsin-Yuan Huang, Richard Kueng, and John Preskill. “Predicting many properties of a quantum system from very few measurements”. *Nature Physics* **16**, 1050–1057 (2020).
- [49] G. Mauro D’Ariano, Matteo G. A. Paris, and Massimiliano F. Sacchi. “Quantum Tomography”. In Peter W. Hawkes, editor, *Advances in Imaging and Electron Physics*. Volume 128, pages 205–308. Elsevier (2003).
- [50] M Guță, J Kahn, R Kueng, and J A Tropp. “Fast state tomography with optimal error bounds”. *Journal of Physics A: Mathematical and Theoretical* **53**, 204001 (2020).
- [51] Andrew Arrasmith, M. Cerezo, Piotr Czarnik, Lukasz Cincio, and Patrick J. Coles. “Effect of barren plateaus on gradient-free optimization”. *Quantum* **5**, 558 (2021).
- [52] R. Penrose. “A generalized inverse for matrices”. *Mathematical Proceedings of the Cambridge Philosophical Society* **51**, 406–413 (1955).

A The Composite Parametrization of Isometries

In this section, we derive the composite parametrization for isometries. As the name suggests, it is obtained from the composite parametrization of the unitary group $\mathcal{U}(d)$ introduced in Ref. [14, 15].

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces with dimensions d_1 and d_2 , respectively. Let $\mathcal{L}(\mathcal{H})$ be the set of linear operators mapping $\mathcal{H}$ into itself, and $\text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$ the set of isometries from $\mathcal{H}_1$ to $\mathcal{H}_2$ with $d_1 \leq d_2$. Furthermore, let $\{|i\rangle_\ell\}_{i=0}^{d_\ell-1}$ be the computational basis of $\mathcal{H}_\ell$.

Any isometry $V \in \text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$ with $d_1 \leq d_2$ can be written as

$$V = U \mathbb{1}_{d_2 \times d_1}, \quad (29)$$

where $U \in \mathcal{U}(d_2)$, and $\mathbb{1}_{d_2 \times d_1} \in \text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the $d_2 \times d_1$ matrix consisting of the first d_1 columns of $\mathbb{1}_{d_2 \times d_2}$. It can be written as

$$\mathbb{1}_{d_2 \times d_1} = \sum_{i=0}^{d_1-1} |i\rangle_2 \langle i|_1. \quad (30)$$

Note that not all basis vectors $|i\rangle_2$ of $\mathcal{H}_2$ need to appear in this sum. By (29), the action $VAV^\dagger$ of V on $A \in \mathcal{L}(\mathcal{H}_1)$ can be interpreted as first canonically embedding A into $\mathcal{L}(\mathcal{H}_2)$ and subsequently applying the unitary U .

We can use the composite parametrization of unitary matrices [14, 15] to write U as

$$U = \left[\prod_{m=0}^{d_2-2} \left(\prod_{n=m+1}^{d_2-1} \Lambda_{m,n} \right) \right] \left[\prod_{l=0}^{d_2-1} e^{iP_l \lambda_l} \right], \quad (31)$$

with

$$P_n := |n\rangle_2 \langle n|_2, \quad (32)$$

$$Y_{m,n} := -i|m\rangle_2 \langle n|_2 + i|n\rangle_2 \langle m|_2, \quad 0 \leq m < n \leq d_2 - 1, \quad (33)$$

$$\Lambda_{m,n} := e^{iP_n \lambda_{n,m}} e^{iY_{m,n} \lambda_{m,n}}, \quad (34)$$

and $\lambda_{m,n} \in [0, 2\pi]$ for $m \geq n$ and $\lambda_{m,n} \in [0, \pi/2]$ for $m < n$.

Using (29) and (31) we compute

$$\prod_{l=0}^{d_2-1} e^{iP_l \lambda_{ll}} \mathbb{1}_{d_2 \times d_1} = \left(\sum_{l=0}^{d_2-1} e^{i\lambda_{l,l}} |l\rangle_2 \langle l|_2 \right) \left(\sum_{k=0}^{d_1-1} |k\rangle_2 \langle k|_1 \right) \quad (35)$$

$$= \sum_{k=0}^{d_1-1} e^{i\lambda_{k,k}} |k\rangle_2 \langle k|_1, \quad (36)$$

and observe that only the first d_1 phases $\lambda_{k,k}$ are relevant for the isometry. This $d_2 \times d_1$ -dimensional matrix is explicitly given by

$$\prod_{l=0}^{d_2-1} e^{iP_l \lambda_{ll}} \mathbb{1}_{d_2 \times d_1} = \begin{pmatrix} e^{i\lambda_{0,0}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & e^{i\lambda_{d_1-1,d_1-1}} \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \quad (37)$$

Next, we calculate

$$\Lambda_{m,n} = \mathbb{1}_{d_2 \times d_2} + (c_{m,n} - 1)P_m + (e_{n,m}c_{m,n} - 1)P_n - e_{n,m}s_{m,n}|n\rangle_2 \langle m|_2 + s_{m,n}|m\rangle_2 \langle n|_2, \quad (38)$$

where we abbreviated $\sin \lambda_{m,n} = s_{m,n}$, $\cos \lambda_{m,n} = c_{m,n}$, and $e^{i\lambda_{n,m}} = e_{n,m}$. In matrix notation, this amounts to

$$\Lambda_{m,n} = \begin{pmatrix} \mathbb{1}_{m \times m} & 0 & 0 & 0 & 0 \\ 0 & c_{m,n} & 0 & s_{m,n} & 0 \\ 0 & 0 & \mathbb{1}_{(n-m-1) \times (n-m-1)} & 0 & 0 \\ 0 & -e_{n,m}s_{m,n} & 0 & e_{n,m}c_{m,n} & 0 \\ 0 & 0 & 0 & 0 & \mathbb{1}_{(d_2-n-1) \times (d_2-n-1)} \end{pmatrix}. \quad (39)$$

By inspection, we see that $\Lambda_{m,n}$ acts trivially from the left on a matrix of the form (37) if $m \geq d_1$. Thus, we can write the isometry (29) as

$$V = \left[\prod_{m=0}^{d_2-2} \left(\prod_{n=m+1}^{d_2-1} \Lambda_{m,n} \right) \right] \left[\prod_{l=0}^{d_2-1} e^{iP_l \lambda_{ll}} \right] \mathbb{1}_{d_2 \times d_1} \quad (40)$$

$$= \left[\prod_{m=0}^{d_1-1} \left(\prod_{n=m+1}^{d_2-1} \Lambda_{m,n} \right) \right] \left[\prod_{l=0}^{d_1-1} e^{iP_l \lambda_{ll}} \right] \mathbb{1}_{d_2 \times d_1}. \quad (41)$$

Regarding the parameter count in (41), we note that the first term on the right-hand side gives $\sum_{m=0}^{d_1-1} 2(d_2 - m - 1) = 2d_1d_2 - d_1^2 - d_1$ because each $\Lambda_{m,n}$ introduces two degrees of freedom. The second term on the right-hand side gives an additional d_1 free parameters. Thus, we find that the isometry $V \in \text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$ in (41) has $2d_1d_2 - d_1^2$ free real parameters. A simple argument shows that this is indeed the number of free real parameters of a general isometry in $\text{Iso}(\mathcal{H}_1, \mathcal{H}_2)$. Consequently, we cannot eliminate more parameters from (41).

The parameters of the isometry (41) can be conveniently collected in the matrix

$$(\lambda_{m,n})_{m,n} = \begin{pmatrix} \lambda_{0,0} & \dots & \lambda_{0,d_1-1} & \dots & \dots & \lambda_{0,d_2-1} \\ \vdots & \ddots & \vdots & & & \vdots \\ \lambda_{d_1-1,0} & \dots & \lambda_{d_1-1,d_1-1} & \dots & \dots & \lambda_{d_1-1,d_2-1} \\ \vdots & & \vdots & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_{d_2-1,0} & \dots & \lambda_{d_2-1,d_1-1} & 0 & \dots & 0 \end{pmatrix}, \quad (42)$$

The diagonal entries $\lambda_{n,n}$ correspond to global phases in the respective subspaces. The entries $\lambda_{m,n}$ in the upper triangular part represent rotations in the subspaces spanned by $|n\rangle_2$ and $|m\rangle_2$, and $\lambda_{n,m}$ in the lower triangular part correspond to relative phases in these subspaces. Except for the diagonal entries $\lambda_{k,k}$, these are the same parameters needed to parameterize a general mixed state $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$ of rank $d_1 \leq d_2$ (in addition to $d_1 - 1$ required mixing probabilities; cf. [14]).

A.1 Composite Parametrization for the Stinespring Isometry of a Quantum Channel

A general CPTP map $\mathcal{E} : \mathcal{D}(\mathcal{H}_1) \rightarrow \mathcal{D}(\mathcal{H}_2)$ can be written in its Stinespring representation as

$$\mathcal{E}(\rho) = \text{Tr}_{1,A}(V \rho V^\dagger), \quad (43)$$

where $V \in \text{Iso}(\mathcal{H}_1, \mathcal{H}_2 \otimes \mathcal{H}_A \otimes \mathcal{H}_1)$, and the Hilbert space $\mathcal{H}_A$ with $\dim(\mathcal{H}_A) = d_A$ corresponds to an ancilla system. We can choose $d_A = d_2$ because $d_1 d_2$ is the maximal Kraus rank of $\mathcal{E}$ [16]. In this case, the isometry in (41) with $d_2 d_A d_1 = d_1 d_2^2$ has $d_1^2(2d_2^2 - 1)$ degrees of freedom. However, due to the unitary freedom on the space $\mathcal{H}_A \otimes \mathcal{H}_1$ (which has no physical relevance), a general CPTP map can be reduced to $d_1^2(d_2^2 - 1)$ degrees of freedom. Unfortunately, we cannot straightforwardly get rid of the $d_1^2 d_2^2$ redundant parameters in (42) because these degrees of freedom do not coincide one-to-one with the parameters $\lambda_{i,j}$.

Nonetheless, using (41), we can write a Stinespring isometry for $\mathcal{E}$ as

$$V = U_C \mathbb{1}_{d_1 d_2^2 \times d_1} = U_C (|0\rangle_2 \otimes |0\rangle_A \otimes \mathbb{1}_1), \quad (44)$$

with

$$U_C = \left[\prod_{m=0}^{d_1-1} \left(\prod_{n=m+1}^{d_1 d_2^2 - 1} \Lambda_{m,n} \right) \right] \left[\prod_{l=0}^{d_1-1} e^{i P_l \lambda_{ll}} \right]. \quad (45)$$

We thus obtain for any $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$:

$$\mathcal{E}(\rho_1) = \text{Tr}_{1,A}[V \rho_1 V^\dagger] \quad (46)$$

$$= \sum_{k=0}^{d_2-1} \sum_{l=0}^{d_1-1} \langle k|_A \otimes \langle l|_1 V \rho_1 V^\dagger |k\rangle_A \otimes |l\rangle_1 \quad (47)$$

$$= \sum_{k=0}^{d_2-1} \sum_{l=0}^{d_1-1} \langle k|_A \otimes \langle l|_1 U_C |0\rangle_2 \otimes |0\rangle_A \rho_1 \langle 0|_2 \otimes \langle 0|_A U_C^\dagger |k\rangle_A \otimes |l\rangle_1 \quad (48)$$

$$= \sum_{k=0}^{d_2-1} \sum_{l=0}^{d_1-1} G_{k,l} \rho_1 G_{k,l}^\dagger, \quad (49)$$

with the Kraus operators $G_{k,l} = \langle k|_A \otimes \langle l|_1 V = \langle k|_A \otimes \langle l|_1 U_C |0\rangle_2 \otimes |0\rangle_A$.

Note that this representation of a quantum channel $\mathcal{E}$ requires us to tensor the systems $\mathcal{H}_A$ and $\mathcal{H}_2$ from the left onto $\mathcal{H}_1$ to obtain $|0\rangle_2 \otimes |0\rangle_A \otimes \mathbb{1}_1$ in (44). In this case, the parameters in the composite parametrized isometry are reduced to the correct number $d_1^2(2d_2^2 - 1)$. If we tensor the systems $\mathcal{H}_A$ and $\mathcal{H}_2$ from the right onto $\mathcal{H}_1$, we would get $\mathbb{1}_1 \otimes |0\rangle_A \otimes |0\rangle_2$ in (44). Consequently, the $d_1^2 d_2^4$ parameters of the full unitary U_C get reduced only by less than $d_1^2(d_2^2 - 1)^2$, and (45) is no longer valid. The reason is that the reduction of parameters shown above cannot be carried out. This also becomes apparent when numerically optimizing DQNNs using the composite parametrization because the gradient for redundant degrees of freedom in the unitary formalism vanishes. If tensored in the “wrong” order, only some diagonal elements of the parameter matrix (42) are irrelevant for the quantum channel and have vanishing derivative (cf. [7]). If done correctly, only the parameters $\lambda_{i,j}$ in (42) are relevant for the optimization, i.e., have non-vanishing derivative in general. Consequently, the derivative of the cost function only needs to be calculated for those, leading to better computational performance.

B Derivatives of different cost functions

This section presents the derivatives of the different cost functions required for optimizing a DQNN by gradient descent/ascent. In Sec. 4, the resulting gradient matrix is used to update the isometry parameters, e.g., with the ADAM optimizer [18].

For example, consider an extended DQNN consisting of 5 qudits and two perceptrons U_1 and U_2 in the unitary formulation. It comprises one input, one hidden, one output, and two ancilla layers. Using the notation of Sec. 2.2, we have $U_1 \equiv U_1^{(1,2,3)}$ and $U_2 \equiv U_1^{(3,4,5)}$. Let us denote the set of variational parameters in U_m as $\mathcal{S}_m = \{\lambda_{x,y}^{(m)}\}_{x,y}$. Due to the equivalence of the isometry and the unitary picture, this is already a reduced set of parameters (cf. (42)). The network channel is given by

$$\rho_{\text{out}} = \mathcal{E}_{\text{net}}(\rho_{\text{in}}) = \text{Tr}_{1,2,3,4} \left\{ U (|\mathbf{0}\rangle\langle\mathbf{0}| \otimes \rho_{\text{in}}) U^\dagger \right\}, \quad (50)$$

where $U = U_2 U_1$, and $|\mathbf{0}\rangle\langle\mathbf{0}| \equiv |\mathbf{0}\rangle\langle\mathbf{0}|_{5,4,3,2}$ is the initial state of the ancilla, hidden, and output layers.

Gradient optimization utilizes either the numerical or the analytical derivative of the total cost function (11). In the former case, we approximate the cost function gradient by

$$\frac{\partial C_{\text{tot}}}{\partial \lambda_{x,y}^{(m)}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\partial}{\partial \lambda_{x,y}^{(m)}} C(\rho_{\text{tar}}^{(i)}, \rho_{\text{out}}^{(i)}) \quad (51)$$

$$\approx \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{C(\rho_{\text{tar}}^{(i)}, \rho_{\text{out}}^{(i)}(\lambda_{x,y}^{(m)} + \varepsilon)) - C(\rho_{\text{tar}}^{(i)}, \rho_{\text{out}}^{(i)}(\lambda_{x,y}^{(m)} - \varepsilon))}{2\varepsilon} \quad (52)$$

for each $m \in \{1, 2\}$ and relevant x, y (cf. (42)). This method requires choosing a suitable small $\varepsilon > 0$ and is generally only an approximation of the true gradient. Thus, we resort to it only when we cannot compute the cost function's analytical derivative. This is the case for the quantum Chernoff bound (25) and the quantum relative entropy (26).

For the analytic approach, we compute (51) analytically for each $m \in \{1, 2\}$ and relevant x, y . To unclutter the notation in the following, we focus on one input-target pair of the training set (i.e., one term in the sum (51)) and drop the superscript (i) . When evaluating $\partial C(\rho_{\text{tar}}, \rho_{\text{out}})/\partial \lambda_{x,y}^{(m)}$, we necessarily encounter $\partial \rho_{\text{out}}/\partial \lambda_{x,y}^{(m)}$ due to the chain rule. Thus, before moving on, we first take care of this. As shown in [7], we can calculate for our example DQNN

$$\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(1)}} = \text{Tr}_{1,2,3,4} \left\{ U_2 U_1 i [\tilde{Y}_{x,y}^{(1)}, \tilde{\rho}] U_1^\dagger U_2^\dagger \right\}, \quad (53)$$

$$\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(2)}} = \text{Tr}_{1,2,3,4} \left\{ U_2 i [\tilde{Y}_{x,y}^{(2)}, U_1 \tilde{\rho} U_1^\dagger] U_2^\dagger \right\}, \quad (54)$$

where $[\cdot, \cdot]$ denotes the commutator, $\tilde{\rho} = |\mathbf{0}\rangle\langle\mathbf{0}| \otimes \rho_{\text{in}}$, and

$$\tilde{Y}_{x,y}^{(m)} = \begin{cases} U_m^\dagger Y_{x,y}^{(m)} U_m & , x < y \\ P_x^{(m)} & , x = y \\ U_m^\dagger P_x^{(m)} U_m & , x > y \end{cases} \quad (55)$$

with $P_x^{(m)}$ and $Y_{x,y}^{(m)}$ as in (32) and (33), respectively. Generalizing the calculation for (53) and (54) to larger networks with more (unitary) perceptrons U_m is straightforward.

The remainder of this section deals with evaluating $\partial C(\rho_{\text{tar}}, \rho_{\text{out}})/\partial \lambda_{x,y}^{(m)}$ for the cost functions presented in Sec. 3. It is important to keep in mind that only ρ_{out} depends on the variational parameters $\lambda_{x,y}^{(m)}$. Hence, $\partial \rho_{\text{out}}/\partial \lambda_{x,y}^{(m)} \neq 0$ in general, but $\partial \rho_{\text{tar}}/\partial \lambda_{x,y}^{(m)} \equiv 0$. Throughout, we furthermore assume $\rho_{\text{tar}} \neq \rho_{\text{out}}$ as this can become problematic for the gradient of certain cost functions. This condition can be implemented in the training algorithm: Before computing the gradient for any optimization iteration, check if $\rho_{\text{out}} = \rho_{\text{tar}}$. If yes, the network is already optimal for this training state, and we exclude this instance from the present iteration of the gradient calculation.

B.1 Derivative of D_{HS}

A straightforward calculation using (18) yields

$$\frac{\partial D_{\text{HS}}(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \sqrt{\text{Tr}((\rho_{\text{out}} - \rho_{\text{tar}})^2)} \quad (56)$$

$$= \frac{1}{2D_{\text{HS}}(\rho_{\text{tar}}, \rho_{\text{out}})} \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \text{Tr}((\rho_{\text{out}} - \rho_{\text{tar}})^2) \quad (57)$$

$$= \frac{1}{D_{\text{HS}}(\rho_{\text{tar}}, \rho_{\text{out}})} \text{Tr} \left((\rho_{\text{out}} - \rho_{\text{tar}}) \frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \right). \quad (58)$$

B.2 Derivative of D_{Tr}

The derivative of the trace distance cost function (19) is

$$\frac{\partial D_{\text{Tr}}(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{1}{2} \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \text{Tr} \left(\sqrt{(\rho_{\text{out}} - \rho_{\text{tar}})^2} \right) \quad (59)$$

$$= \frac{1}{4} \text{Tr} \left(\left(\sqrt{(\rho_{\text{out}} - \rho_{\text{tar}})^2} \right)^{-1} \frac{\partial}{\partial \lambda_{x,y}^{(m)}} (\rho_{\text{out}} - \rho_{\text{tar}})^2 \right) \quad (60)$$

$$= \frac{1}{2} \text{Tr} \left(\left(\sqrt{(\rho_{\text{out}} - \rho_{\text{tar}})^2} \right)^{-1} (\rho_{\text{out}} - \rho_{\text{tar}}) \frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \right). \quad (61)$$

If $A = \sqrt{(\rho_{\text{out}} - \rho_{\text{tar}})^2}$ is singular, we use the Moore-Penrose inverse [52] instead of A^{-1} . In more detail, in (60) we used

$$\text{Tr} \left(\frac{\partial}{\partial \lambda} \sqrt{A} \right) = \text{Tr} \left(\frac{1}{2} \frac{\partial A}{\partial \lambda} (A^{1/2})^{-1} \right) \quad (62)$$

for $\lambda = \lambda_{x,y}^{(m)}$ and $A = (\rho_{\text{out}} - \rho_{\text{tar}})^2$. To evaluate this, we can utilize the power series expansion for the matrix square root, given by

$$A^{1/2} = \sum_{n=0}^{\infty} (-1)^n \binom{1/2}{n} (\mathbb{1} - A)^n, \quad (63)$$

which is convergent if the spectrum of A satisfies $\text{spec}(A) \subseteq \mathcal{D}(1, 1) \subset \mathbb{C}$, where $\mathcal{D}(1, 1)$ denotes a disk with radius 1 and centered at 1 in $\mathbb{C}$. If $\text{spec}(A) \subseteq (0, 1]$, i.e., A is non-singular, the inverse of (63) is

$$A^{-1/2} = \sum_{n=0}^{\infty} (-1)^n \binom{-1/2}{n} (\mathbb{1} - A)^n. \quad (64)$$

For singular A we can use the Moore-Penrose pseudoinverse to define (64). We can then calculate

$$\text{Tr} \left(\frac{\partial}{\partial \lambda} \sqrt{A} \right) = \text{Tr} \left(\frac{\partial}{\partial \lambda} \sum_{n=0}^{\infty} (-1)^n \binom{1/2}{n} (\mathbb{1} - A)^n \right) \quad (65)$$

$$= \text{Tr} \left(\sum_{n=0}^{\infty} (-1)^n \binom{1/2}{n} \sum_{k=0}^{n-1} (\mathbb{1} - A)^k \frac{\partial (\mathbb{1} - A)}{\partial \lambda} (\mathbb{1} - A)^{n-k-1} \right) \quad (66)$$

$$= \text{Tr} \left(\frac{\partial A}{\partial \lambda} \sum_{n=1}^{\infty} (-1)^{n+1} \binom{1/2}{n} n (\mathbb{1} - A)^{n-1} \right) \quad (67)$$

$$= \text{Tr} \left(\frac{1}{2} \frac{\partial A}{\partial \lambda} A^{-1/2} \right), \quad (68)$$

where in the third equality we used that the trace is cyclic, and a simple index shift in combination with the identity $\binom{1/2}{n+1} (n+1) = \frac{1}{2} \binom{-1/2}{n}$ yields the last line.

B.3 Derivative of F_1 and D_1

The derivative of the fidelity cost function (21) takes the form

$$\frac{\partial F_1(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \text{Tr} \left(\sqrt{\sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}} \right)^2 \quad (69)$$

$$= 2 \sqrt{F_1(\rho_{\text{tar}}, \rho_{\text{out}})} \text{Tr} \left(\frac{\partial}{\partial \lambda_{x,y}^{(m)}} \sqrt{\sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}} \right) \quad (70)$$

$$= \sqrt{F_1(\rho_{\text{tar}}, \rho_{\text{out}})} \text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \sqrt{\rho_{\text{tar}}} \left(\sqrt{\sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}} \right)^{-1} \sqrt{\rho_{\text{tar}}} \right), \quad (71)$$

where in the last line we used (62) with $A = \sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}$. This is valid as one can show that $\text{spec}(\sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}) \subset [0, 1]$.

For the Bures distance cost function (22), we obtain

$$\frac{\partial D_1(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \sqrt{2 \left(1 - \sqrt{F_1(\rho_{\text{tar}}, \rho_{\text{out}})} \right)} \quad (72)$$

$$= \frac{1}{D_1(\rho_{\text{tar}}, \rho_{\text{out}})} \frac{\partial}{\partial \lambda_{x,y}^{(m)}} \left(1 - \sqrt{F_1(\rho_{\text{tar}}, \rho_{\text{out}})} \right) \quad (73)$$

$$= \frac{-1}{2D_1(\rho_{\text{tar}}, \rho_{\text{out}}) \sqrt{F_1(\rho_{\text{tar}}, \rho_{\text{out}})}} \frac{\partial F_1(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} \quad (74)$$

$$= \frac{-1}{2D_1(\rho_{\text{tar}}, \rho_{\text{out}})} \text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \sqrt{\rho_{\text{tar}}} \left(\sqrt{\sqrt{\rho_{\text{tar}}} \rho_{\text{out}} \sqrt{\rho_{\text{tar}}}} \right)^{-1} \sqrt{\rho_{\text{tar}}} \right), \quad (75)$$

where we used (71) in the last line.

B.4 Derivative of F_2 and D_2

To take the derivative of the fidelity F_2 , we write it as

$$F_2(\rho_{\text{tar}}, \rho_{\text{out}}) = \frac{\mathcal{A}}{\mathcal{B}} \quad (76)$$

where

$$\mathcal{A} = \text{Tr}(\rho_{\text{out}} \rho_{\text{tar}}), \quad (77)$$

$$\mathcal{B} = \max \left\{ \text{Tr}(\rho_{\text{out}}^2), \text{Tr}(\rho_{\text{tar}}^2) \right\}. \quad (78)$$

Thus, we have

$$\frac{\partial F_2(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{\frac{\partial \mathcal{A}}{\partial \lambda_{x,y}^{(m)}} \mathcal{B} - \mathcal{A} \frac{\partial \mathcal{B}}{\partial \lambda_{x,y}^{(m)}}}{\mathcal{B}^2}. \quad (79)$$

The derivative of $\mathcal{A}$ is easily evaluated to be

$$\frac{\partial \mathcal{A}}{\partial \lambda_{x,y}^{(m)}} = \text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \rho_{\text{tar}} \right). \quad (80)$$

By writing

$$\max \left\{ \text{Tr}(\rho_{\text{out}}^2), \text{Tr}(\rho_{\text{tar}}^2) \right\} = \frac{1}{2} \left(\text{Tr}(\rho_{\text{out}}^2) + \text{Tr}(\rho_{\text{tar}}^2) + \left| \text{Tr}(\rho_{\text{out}}^2) - \text{Tr}(\rho_{\text{tar}}^2) \right| \right) \quad (81)$$

$$= \frac{1}{2} \left(\text{Tr}(\rho_{\text{out}}^2 + \rho_{\text{tar}}^2) + \sqrt{\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2)^2} \right), \quad (82)$$

we see that $\mathcal{B}$ is not differentiable at $\rho_{\text{out}} = \rho_{\text{tar}}$. For all other cases, we can compute

$$\frac{\partial \mathcal{B}}{\partial \lambda_{x,y}^{(m)}} = \text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \rho_{\text{out}} \right) \left(1 + \frac{\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2)}{\sqrt{\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2)^2}} \right) \quad (83)$$

$$= \text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} \rho_{\text{out}} \right) \left(1 + \text{sign} \left(\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2) \right) \right). \quad (84)$$

Inserting (80) and (84) into (79) finally yields after some simplifications

$$\frac{\partial F_2(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{\text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} (\rho_{\text{tar}} - \rho_{\text{out}} F_2(\rho_{\text{tar}}, \rho_{\text{out}}) (1 + \text{sign}(\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2)))) \right)}{\max \{ \text{Tr}(\rho_{\text{out}}^2), \text{Tr}(\rho_{\text{tar}}^2) \}}. \quad (85)$$

For the D_2 distance (24), the derivative evaluates to

$$\frac{\partial D_2(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} = \frac{-1}{D_2(\rho_{\text{tar}}, \rho_{\text{out}})} \frac{\partial F_2(\rho_{\text{tar}}, \rho_{\text{out}})}{\partial \lambda_{x,y}^{(m)}} \quad (86)$$

$$= \frac{-\text{Tr} \left(\frac{\partial \rho_{\text{out}}}{\partial \lambda_{x,y}^{(m)}} (\rho_{\text{tar}} - \rho_{\text{out}} F_2(\rho_{\text{tar}}, \rho_{\text{out}}) (1 + \text{sign}(\text{Tr}(\rho_{\text{out}}^2 - \rho_{\text{tar}}^2)))) \right)}{D_2(\rho_{\text{tar}}, \rho_{\text{out}}) \max \{ \text{Tr}(\rho_{\text{out}}^2), \text{Tr}(\rho_{\text{tar}}^2) \}}. \quad (87)$$