universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„High-dimensional and multipartite quantum metrology and entanglement theory“

verfasst von / submitted by

Simon Morelli, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2022 / Vienna 2022

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:	Assoz. Prof. Mag. Dr. Marcus Huber, Privatdoz.
Betreut von / Supervisor:	Mag. Dr. Nicolai Friis

"Can you tell us what that means?"
"I'm not altogether sure. Let's be straight here. If we find something we can't understand we like to call it something you can't understand, or indeed pronounce. ..."

Douglas Adams, So long, and thanks for all the fish, p. 135

1 Abstract

To realize the ambitious promises of rapidly advancing quantum technologies, it is indispensable to create the theoretical framework for situations likely to be encountered in realistic scenarios and identify strategies taking into account limitations of current technologies. In this spirit, we investigate several key areas in the theoretical description of quantum technologies.

First, we turn to quantum metrology, which plays a crucial role for noisy quantum devices, where it facilitates improved passive error correction via the estimation of unknown parameters and noise channels. In this context, we consider bosonic systems for which we identify easily implementable and robust Bayesian estimation strategies that are relevant for near-future experimental implementations. We further apply techniques from Bayesian metrology for the distribution of high-dimensional entangled states to multiple parties via noisy channels and the subsequent probabilistic conversion of these states to desired target states using stochastic local operations and classical communication. Such state-conversion protocols can be enhanced by embedded channel-estimation routines at no additional cost in terms of the number of copies of the distributed states, allowing the efficient distribution of entanglement in noisy networks.

Entanglement characterization between many copies of mixed states hence becomes essential to fully access the potential of the distributed states. Moreover, we show that multiple copies of partially separable states can unlock genuine multipartite entanglement, making the ability to control and jointly locally manipulate multiple copies of quantum states a valuable resource.

We then turn to the description of the involved measurements. Here it is important to acknowledge that realistic measurements only approximately correspond to those intended. To verify entanglement with imperfect control over the measurement devices, one still wants to exploit as much information as available over the performed measurements. We formalize this through an operational notion of inaccuracy that can be estimated directly in the lab and compute tight corrections to standard entanglement witnesses due to any given level of measurement inaccuracy for two systems of arbitrary dimensions.

Finally we developed a three-dimensional model for the state space of a qutrit. This model, although clearly not in one-to-one correspondence to the actual eight-dimensional state space, still captures many of its fundamental geometric and algebraic properties and thus provides a helpful tool for studying higher-dimensional quantum systems.

2 Kurzfassung

Um die ehrgeizigen Ziele im Feld der zügig voranschreitenden Quantentechnologien zu realisieren, ist es notwendig einen theoretischen Rahmen für realistische Szenarien zu schaffen und Strategien zu entwickeln, welche die Einschränkungen gegenwärtiger Technologien berücksichtigen. In diesem Sinne untersuchen wir verschiedene zentrale Fragestellungen in der theoretischen Beschreibung von Quantentechnologien.

Zuallererst wenden wir uns der Quantenmetrologie zu, welche eine Schlüsselrolle in Apparaturen mit störenden äußeren Einflüssen spielt. Dort erleichtert sie passive Fehlerkorrektur durch das Abschätzen unbekannter Parameter und Rauschen in Quantenkanälen. Für bosonische Systeme ermitteln wir leicht implementierbare und robuste Bayessche Strategien zur Abschätzung von Parametern, welche relevant für experimentelle Anwendungen in naher Zukunft sind. Weiters verwenden wir Methoden der Bayesschen Metrologie zur Verteilung von hochdimensional verschränkten Zuständen an mehrere Parteien mittels imperfekter Quantenkanäle. Die verteilten Zustände werden probabilistisch mit lokalen Operationen und klassischer Kommunikation in die Zielzustände umgewandelt. Solche Protokolle können ohne zusätzliche Kosten in der Anzahl an verteilten Zuständen mit eingebauter Kanalabschätzungen verbessert werden, um eine effiziente Verteilung von verschränkten Zuständen in Netzwerken mit unerwünschtem Rauschen zu erlauben.

Die Charakterisierung der Verschränkung von mehreren Kopien gemischter Zustände wird somit notwendig um auf das gesamte Potential der verteilten Zustände zugreifen zu können. Wir zeigen, dass genuine Vielteilchenverschränkung durch mehrere Kopien von partiell separablen Zuständen aktiviert werden kann, womit die Fähigkeit zur lokalen gemeinsamen Kontrolle und Manipulation multipler Zustandskopien zu einer kostbaren Ressource wird.

Danach wenden wir uns der Beschreibung der beteiligten Messungen zu. Dabei ist es wichtig zu beachten, dass reale Messungen nur näherungsweise den beabsichtigten Messungen entsprechen. Um Verschränkung trotz unvollständiger Kontrolle über die Messapparatur nachzuweisen, ist es vorteilhaft alle verfügbaren Informationen über die ausgeführten Messungen zu verwenden. Wir formalisieren diesen Ansatz durch eine operative Definition der Messungenauigkeit, welche direkt im Labor ermittelt werden kann, und berechnen scharfe Korrekturen von kanonischen Verschränkungszeugen für jedes Niveau von Messungenauigkeiten für zwei Systeme beliebiger Dimensionen.

Schlussendlich entwickeln wir ein dreidimensionales Modell des Zustandsraumes eines Qutrits. Obwohl keine treue Darstellung des achtdimensionalen Zustandsraumes, zeigt das Modell dennoch viele relevante geometrische und algebraische Eigenschaften und dient somit als nützliches Werkzeug zur Erforschung höher-dimensionaler Quantensysteme.

3 Acknowledgements

First and foremost I want to thank my two supervisors Nicolai Friis and Marcus Huber. During the last three years I had the opportunity to learn so much from both of you, going way beyond solely scientific knowledge. Even though at times it must have been demanding for you to take care of your scientific offspring while busy with your natural one, I always knew that I could rely on your full support. I am extremely fortunate to have had two such amazing supervisors.

Thank you Nico for giving me all the freedom I could dream of for my PhD, supporting every decision I made and encouraging me to pursue the research that interested me the most. I could always ask for your help and feedback, but I never felt pressured or pushed towards tasks and projects. Your valuable advice helped me in many situations and even if my PhD is about to end I hope that also in the future I can ask you for your advice if I need it.

Thank you Marcus for welcoming me to your group, where my research could thrive in a relaxed and amicable environment. I highly appreciate you for always valuing my scientific input and that you always treated me as an equal. I doubt that I would ever have pursued an academic path if you would not have believed in me.

I am grateful to all the Quitters, present and past (although there is no such thing as a past Quitter: once a Quitter, always a Quitter) for making the last years such an incredible experience. You are the best academic family imaginable, thank you for letting me be part of it and to walk beside you on our scientific paths.

An especially warm thank you to all my fellow PhD siblings, Fabien Clivaz, Jessica Bavaresco, Paul Appel, Emanuel Schwarzhans and Phil Taranto, without you it wouldn't have been the same. Sharing the same experiences during our PhD made the hard ones bearable and the nice ones extraordinary.

I would like to thank Andreas Winter and Michalis Skotiniotis for hosting me in Barcelona and all of GIQ for welcoming me with open arms. I had a wonderful time with you that I will never forget.

I want to thank all my collaborators, your expertise and curiosity makes working with you a privilege I'm happy to have. You motivate me to tackle the challenges of scientific research and am looking forward to continuing doing so together. Especially I want to thank Jens Siewert, whom I owe my enthusiasm for the Bloch representation. You never grew tired of discussing science with me, I learnt so much from you - and hope to continue doing so in the future.

I am exceptionally grateful to Hannah Tschenett for making the lonesome days of home office so much better. You are always there for me to lift my mood and turn even the

worst day into a precious moment. I am so lucky to have you in my life.

I finally would like to acknowledge the Austrian Science Found (FWF), for founding my PhD through the Stand-Alone project P 31339-N27 and the START project Y879-N27.

Publications

Peer-reviewed

(not included)

Dimensionally sharp inequalities for the linear entropy
Simon Morelli, Claude Klöckl, Christopher Eltschka, Jens Siewert, and Marcus Huber,
Linear Algebra and its Applications 584, 294-325 (2020)
DOI: 10.1016/j.laa.2019.09.008
arXiv: 1903.11887

Bayesian parameter estimation using Gaussian states and measurements
Simon Morelli, Ayaka Usui, Elizabeth Agudelo, and Nicolai Friis,
Quantum Science and Technology 6, 025018 (2021)
DOI: 10.1088/2058-9565/abd83d
arXiv: 2009.03709

The shape of higher-dimensional state space: Bloch-ball analog for a qutrit
Christopher Eltschka, Marcus Huber, Simon Morelli, and Jens Siewert,
Quantum 5, 485 (2021)
DOI: 10.22331/q-2021-06-29-485
arXiv: 2012.00587

Preprints

Activation of genuine multipartite entanglement: beyond the single-copy paradigm of entanglement characterisation
Hayata Yamasaki, Simon Morelli, Markus Miethlinger, Jessica Bavaresco, Nicolai Friis, and Marcus Huber,
arXiv: 2106.01372

Metrology-assisted entanglement distribution in noisy quantum networks
Simon Morelli, David Sauerwein, Michalis Skotiniotis, and Nicolai Friis,
arXiv: 2110.15627

Entanglement detection with imprecise measurements
Simon Morelli*, Hayata Yamasaki*, Marcus Huber, and Armin Tavakoli,
arXiv: 2202.13131

Contents

4 Introduction

In the past years there has been substantial progress in the quantum information sciences, both on the theoretical side as on the technological front. Quantum computers potentially providing an exponential speed-up in computation time with respect to the best known classical algorithms for certain problems [1] and provably secure cryptography are just the most known promises of this new technological era. To exchange quantum information and increase the power of smaller computation units, individual quantum computers can be connected into a network, which is envisioned to one day realize a form of quantum internet [2–5]. Shared entanglement within a quantum network will be a crucial resource for various tasks and the distribution of entanglement between different nodes is one of the central challenges to the creation of such a network [6–9]. Through quantum teleportation entanglement provides the means to reliably exchange quantum information, allowing quantum computers to share computational tasks and access exponentially growing Hilbert-space dimensions [5]. Multipartite entanglement facilitates the distribution of cryptographic keys and allows quantum keys shared between multiple users [10–13]. Furthermore it gives an advantage in complex metrological tasks such as the synchronization of atomic clocks [14]. However, there are substantial obstacles in the distribution, characterization and certification of entanglement. Potential near-future strategies are therefore strongly limited by current technologies and it becomes indispensable to create the theoretical framework for realistic scenarios likely to be encountered in future applications.

Decoherence of quantum systems through the interaction with the environment makes it generally difficult to store and transmit quantum information. This leads to a decay rate for entanglement distribution that severely limits the distribution of entangled states over long distances [15]. Further, the no-cloning theorem forbids copying unknown quantum states, making in-built error correction an essential but demanding task. For this it is necessary to establish an accurate noise model and thus quantum channel estimation becomes important for efficient quantum communication. To estimate the parameters describing an unknown quantum channel, a probe system is sent through the channel and is subsequently being measured. We investigate parameter estimation in continuous-variable quantum systems and develop strategies motivated by technological feasibility to estimate parameters encoding three paradigmatic unitary transformations. We further examine how the techniques for channel estimation can be embedded into the distribution of highly entangled states within a noisy network. The key idea here is to use states for estimation that are usually discarded in the following probabilistic manipulation of the distributed states. Also high-dimensional entanglement can help to overcome noise in

quantum communication [16]. We examine the state space of higher-dimensional systems and create an intuitive three-dimensional model of a qutrit, for which we visualize the effects of unitary transformations and three paradigmatic quantum channels. Characterizing entanglement becomes essential to fully access the potential of the distributed states within a quantum network. We investigate the entanglement properties between many copies of mixed states and show that multiple copies of partially separable states can unlock genuine multipartite entanglement. We finally turn to the certification of entanglement. As entanglement is a crucial ingredient for many applications in quantum communication, computation and cryptography, the involved parties usually want to verify the properties of the used states. We assume imperfect control over the measurement devices and investigate the performance of well-known entanglement witnesses in such a realistic scenario.

The research articles constituting this thesis are a collection of theoretical investigations at the forefront of quantum information. We investigate several key areas of quantum information and quantum metrology, making a contribution to distinct areas of the field.

5 Joint framework

This section is intended as a general framework of the theory used in the presented research papers. It is meant as an overview over some of the studied topics. A more specific elaboration of the topics can, where needed, be found in the actual research articles and for a more complete and detailed treatment of the topics we refer to the indicated literature.

5.1 Quantum states and their representation

A quantum system is associated to a complex Hilbert space $\mathcal{H}$, a pure state $|\psi\rangle$ of the system is described by vectors of length 1, where two vectors represent the same state if they differ only by a global phase. Equivalently pure states can be identified with subspaces of $\mathcal{H}$ of complex dimension 1, called *rays*.

Let $\mathcal{B}(\mathcal{H})$ denote the set of bounded operators acting on $\mathcal{H}$. Mixed states are represented by the subset of normalized positive semidefinite operators $\rho \in \mathcal{B}(\mathcal{H})$. A pure state $|\psi\rangle$ is identified with the projector $|\psi\rangle\langle\psi| \in \mathcal{H} \otimes \mathcal{H}^*$, the outer product of the state $|\psi\rangle$ with its Hermitian conjugate $\langle\psi|$, belonging to the dual space $\mathcal{H}^*$. Every mixed state ρ can be decomposed into a convex combination of projectors $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $p_i > 0$ and $\sum_i p_i = 1$, and we denote the set of all such decompositions as $\mathcal{D}(\rho)$. Hence the quantum states form a convex set, that is exactly the convex hull of all pure states.

For finite-dimensional d-level quantum systems the associated Hilbert space is $\mathcal{H} = \mathbb{C}^d$

and a general state is represented by a complex positive semidefinite $d \times d$-matrix with trace equal 1, called a *density matrix*.

Equipped with the Hilbert-Schmidt inner product $\langle \rho, \sigma \rangle = \mathrm{Tr}(\rho^\dagger \sigma)$, the set of complex $d \times d$-matrices forms a finite-dimensional Hilbert space and hence one can find a basis for it. Following Bertlmann and Krammer [17], a *practical* basis $\{\lambda_i\}_{i=0}^{d^2-1}$ for the Hilbert space $(\mathbb{C}^d)^2$ with $\dim(\mathcal{H}) = d$ satisfies the following properties

i. the identity matrix is included, $\lambda_0 = \mathbb{1}_d$,

ii. the other $d^2 - 1$ elements are traceless matrices,

iii. the basis elements are mutually orthogonal, $\mathrm{Tr}(\lambda_i^\dagger \lambda_j) = d\ \delta_{ij}$.

We call a basis satisfying these criteria a *Bloch basis*. Every state ρ can be represented in a Bloch basis in the following way

$$\rho = \frac{1}{d}\left(\mathbb{1}_d + \sum_{i=1}^{d^2-1} b_i \lambda_i\right) = \frac{1}{d}\left(\mathbb{1}_d + \vec{\mathbf{b}} \cdot \vec{\mathbf{\Gamma}}\right), \tag{1}$$

where $\vec{\mathbf{b}} \in \mathbb{C}^{d^2-1}$ is the Bloch vector and $\vec{\mathbf{\Gamma}}$ the vector of basis elements. The entries of the Bloch vector are $b_i = \langle \lambda_i \rangle = \mathrm{Tr}(\rho \lambda_i)$ and satisfy $\sum_{i=1}^{d} b_i^* b_i = \|\vec{\mathbf{b}}\|^2 \leq d - 1$. Here it is important to note that not every matrix that can be written in this way is positive, i.e. not every vector with norm smaller than or equal to $\sqrt{d-1}$ represents a state. By choosing a Bloch basis we can identify the quantum state space with a subspace of $\mathbb{C}^{(d^2-1)}$ of (real) dimension $(d^2 - 1)$. This becomes obvious by choosing a Hermitian Bloch basis, as the entries of the Bloch vector are then real. It is always possible to choose a Hermitian Bloch basis, for $d > 2$ it is however not possible to find such a set of Hermitian and unitary matrices [17].

A composite quantum system is associated to the tensor product of the Hilbert spaces of the individual systems $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. For a bipartite quantum system with associated product Hilbert space it is easy to construct a basis for the whole space from the two local bases, by taking all possible tensor products of the basis elements. More formally, let $\{\lambda_i\}_{i=0}^{d_A^2-1}$ be a basis for the Hilbert space $\mathcal{H}_A$ of dimension d_A and $\{\sigma_j\}_{j=0}^{d_B^2-1}$ for the Hilbert space $\mathcal{H}_B$ of dimension d_B, then $\{\lambda_i \otimes \sigma_j\}_{i,j}$ is a basis for the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ of dimension $d_{AB} = d_A d_B$. It can be easily seen that if both $\{\lambda_i\}_{i=0}^{d_A^2-1}$ and $\{\sigma_j\}_{j=0}^{d_B^2-1}$ are Bloch bases, so is $\{\lambda_i \otimes \sigma_j\}_{i,j}$. Every state $\rho_{AB} \in \mathcal{H}_{AB}$ can then be written as

$$\rho_{AB} = \frac{1}{d_{AB}}\left(\mathbb{1}_{d_A} \otimes \mathbb{1}_{d_B} + \sum_{i=1}^{d_A^2-1} a_i \lambda_i \otimes \mathbb{1}_{d_B} + \sum_{j=1}^{d_B^2-1} b_j \mathbb{1}_{d_B} \otimes \sigma_j + \sum_{i,j \geq 1} c_{ij} \lambda_i \otimes \sigma_j\right). \tag{2}$$

This concept can straightforwardly be generalized to more parties.

5.2 Entanglement theory

In this subsection we shortly revise the relevant topics that appear in this cumulative thesis. For a more extensive presentation of entanglement theory the reader is referred to Refs. [18, 19].

5.2.1 Multi-party entanglement

A pure state $|\psi_{AB}\rangle \in \mathcal{H}_{AB}$ of a bipartite quantum system is called *separable*, if it can be written in the form $|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\phi_B\rangle$, where $|\phi_A\rangle \in \mathcal{H}_A$ and $|\phi_B\rangle \in \mathcal{H}_B$. A mixed bipartite state ρ_{AB} is *separable*, if there exists a decomposition into pure separable states $\rho_{AB} = \sum_i p_i |\phi_A^i\rangle\langle\phi_A^i| \otimes |\phi_B^i\rangle\langle\phi_B^i|$. One can easily check that the condition is equivalent to $\rho_{AB} = \sum_i q_i \rho_A^i \otimes \rho_B^i$. If a state is not separable, it is called *entangled*.

The notion of separability for states of a multipartite quantum system has to be specified with respect to a *partition* of the parties. This was not necessary for bipartite systems since there exists only one non-trivial partition of two parties. A partition of a set $\mathcal{M}$ is a collection of non-empty subsets $\mathcal{A}_i \subset \mathcal{M}$, such that $\mathcal{A}_i \cap \mathcal{A}_j = \varnothing$ for $i \neq j$ and $\bigcup_i \mathcal{A}_i = \mathcal{M}$. A pure quantum state $|\psi_{[N]}\rangle$ of an N-partite system with associated Hilbert space $\mathcal{H}_{[N]} = \bigotimes_{i=1}^N \mathcal{H}_i$ is called k-separable, if there exists a partition $\{\mathcal{A}_1, \ldots, \mathcal{A}_k\}$ of $[N] = \{1, \ldots, N\}$ such that $|\psi_{[N]}\rangle = \bigotimes_{i=1}^k |\phi_{\mathcal{A}_i}\rangle$, where $|\phi_{\mathcal{A}_i}\rangle \in \bigotimes_{j\in\mathcal{A}_i} \mathcal{H}_j$. In the special case of $k = N$ the state is called *fully separable* and for the minimal case $k = 2$ *biseparable*. A pure state is called *multipartite entangled* if it is not k-separable for any $k \geq 2$, i.e. it is not biseparable.

When extending the notion of separability to mixed states, one can either pin the definition to a specific partition or only on the cardinality k of the involved partitions of a state. This results in two different notions of separability for mixed states. A mixed N-partite state $\rho_{[N]}$ is called *partition separable* with respect to the partition $\{\mathcal{A}_1, \ldots, \mathcal{A}_k\}$, if there exists a decomposition into pure states separable in the same partition $\{\mathcal{A}_1, \ldots, \mathcal{A}_k\}$. If a mixed state ρ is not partition separable with respect to any partition it is called *multipartite entangled*. A mixed N-partite state $\rho_{[N]}$ is called k-separable, if there exists a decomposition into k-separable pure states. A mixed state ρ that is not k-separable with respect to any partition is *genuinely multipartite entangled*. The k-separable states form a convex set, where the set of partition separable states is a strict subset, as all partition separable states (with partition length k) are also k-separable. In fact the set of k-separable states is the convex hull of the set of partition separable states (with partition length k).

5.2.2 Entanglement detection

To decide whether a given state is entangled is generally an NP-hard problem in the Hilbert-space dimension [20, 21]. Every mixed state can be decomposed into pure states, but the decomposition is not unique. In fact there exists a continuum of different decompositions for a given state, and in principle one has to look at all of them to exclude the possibility for a separable one. There are however sufficient conditions to certify entanglement shared between different parties. Two paradigmatic examples are positive but not completely positive maps and entanglement witnesses.

A linear map $\Lambda \in \mathcal{B}(\mathcal{H})$ acting on the space of operators of a Hilbert space $\mathcal{B}(\mathcal{H})$ is called *positive*, if $\Lambda(\rho)$ is a positive operator for all positive operators ρ. A map Λ is called *completely positive* (CP), if $\Lambda_A \otimes \mathbf{id}_B(\rho_{AB}) \geq 0$ for all positive operators $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and arbitrary dimension of $\mathcal{H}_B$. A completely positive map is positive, but the converse does not hold in general. Let Λ be a positive but not completely positive map, then for every separable state ρ_{SEP} it holds that

$$\Lambda_A \otimes \mathbf{id}_B(\rho_{SEP}) = \sum_i p_i \Lambda(|\phi_A^i\rangle\langle\phi_A^i|) \otimes |\phi_B^i\rangle\langle\phi_B^i| \geq 0. \tag{3}$$

So if a state is not positive after a positive but not completely positive map is applied to one subsystem, it already follows that the state is entangled. In fact, a bipartite state is separable exactly if it is positive under every positive map applied to one subsystem. Transposition is the most paradigmatic example of a positive but not completely positive map. Every separable state is positive under partial transposition (PPT). No pure entangled state is positive under partial transposition (NPPT), but in general the converse does not hold for mixed states, that is there exists entangled states that are positive under partial transposition. However, in dimensions $d_A = d_B = 2$ and $d_A = 2$, $d_B = 3$ the PPT-criterion is a necessary and sufficient condition for separability [22]. Another example for a positive but not completely positive map is the *Choi map* [23], defined as $\Lambda_d(\rho) = 2\,\mathrm{Diag}(\rho) + \sum_{j=1}^{d-2} X_d^j\,\mathrm{Diag}(\rho)(X_d^j)^\dagger - \rho$, where $\mathrm{Diag}(\rho)$ denotes the diagonal matrix with the same diagonal entries as ρ and $(X_d)_{i,j} = \delta_{i,j-1}$ (indices $\mod d$) is the shift operator.

An observable W is called an *(linear) entanglement witness*, if $\mathrm{Tr}(\rho_{SEP}W) \geq 0$ for every separable state and there exists a state ρ such that $\mathrm{Tr}(\rho W) < 0$. Since all separable states have positive expectation values for such an entanglement witness, a negative expectation value already implies entanglement. On the contrary, a positive expectation value does not allow one to conclude whether the state is entangled or separable. So entanglement witnesses are observables dividing the state space into two regions, one of them containing the subset of separable states. For every entangled state ρ there exists an observable W such that $\mathrm{Tr}(\rho W) < 0$ and $\mathrm{Tr}(\rho_{SEP}W) \geq 0$ for all separable states.

The *Choi-Jamiołkowski isomorphism* establishes a one-to-one correspondence between maps $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ to operators $E \in \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ via $\mathcal{E}(\rho) = \mathrm{Tr}_A(E\rho^T \otimes \mathbb{1}_B)$ and the inverse $E = \sum_{i,j} |i\rangle\langle j| \otimes \mathcal{E}(|i\rangle\langle j|)$ [19]. This connects positive maps to entanglement witnesses, as the following holds

i. The map $\mathcal{E}$ is completely positive, iff E is positive semidefinite.

ii. The map $\mathcal{E}$ is positive but not completely positive, iff E is an entanglement witness.

5.2.3 Entanglement characterization

Entangled states shared between different parties are a key resource in quantum information and they cannot be created nor can their number be increased by local operations assisted by classical communication (LOCC). We have seen in the previous section how states can be entangled with respect to different partitions. But also for the same partitions states can be entangled in different ways and to a different extent.

For pure states of a bipartite system their entanglement is characterized by their *Schmidt coefficients*. Every such state $|\psi_{AB}\rangle$ can be written as $|\psi_{AB}\rangle = \sum_{i=0}^{r-1} \sqrt{\lambda_i} |\nu_A^i\rangle \otimes |\phi_B^i\rangle$, where $\{|\nu_A^i\rangle\}$ and $\{|\phi_B^i\rangle\}$ are orthonormal bases and λ_i are positive real numbers satisfying $\sum_{i=0}^{r-1} \lambda_i = 1$. This decomposition is known as the *Schmidt decomposition* and $r \leq \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\}$ is called the *Schmidt rank* of the state $|\psi_{AB}\rangle$. A Schmidt rank greater than 1 is a necessary and sufficient condition for the state to be entangled. Writing the Schmidt coefficients as a vector $v^{\downarrow}(|\psi\rangle)$ with entries λ_i in decreasing order, the *Nielsen majorisation theorem* gives a necessary and sufficient condition for the possibility of a transformation between two given states via LOCC. A state $|\chi\rangle$ can be obtained by LOCC from the state $|\psi\rangle$ if and only if $v^{\downarrow}(|\psi\rangle)$ is *majorised* by $v^{\downarrow}(|\chi\rangle)$, that is $\sum_{i=1}^{k} v^{\downarrow}(|\psi\rangle)_i \leq \sum_{i=1}^{k} v^{\downarrow}(|\chi\rangle)_i$ for all $k \leq d$ [24]. This introduces a partial order on the set of bipartite pure states.

Since not all states are comparable in a partial order, the Schmidt decomposition does not allow us to quantify entanglement. On the other side it is clear that maximally entangled states of the form $|\psi_{ME}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ are the maximal elements in this partial order (for fixed dimension d) and comparable to all other states, thus they are a universal resource. This idea is used in the following two definitions, that can be directly extended to mixed states. The *entanglement cost* $\mathbf{E}_C(\rho)$ is defined as the minimal rate at which m copies of maximally entangled two-qubit states can be used to create n copies of the state ρ via LOCC

$$\mathbf{E}_C(\rho) = \inf_{LOCC} \lim_{n\to\infty} \frac{m}{n}. \tag{4}$$

The infimum is over all *asymptotically exact* LOCC transformation, that is the output comes arbitrary close to the target in the asymptotic limit. The *entanglement of distillation* $\mathbf{E}_D(\rho)$ is defined as the maximal distillation rate at which n copies of ρ can be transformed into m copies of maximally entangled qubit states via LOCC

$$\mathbf{E}_D(\rho) = \sup_{LOCC} \lim_{n\to\infty} \frac{m}{n}, \tag{5}$$

where again the supremum is over all asymptotically exact LOCC transformations. These definitions have a clear interpretation, but lack a practical way of actually computing them. For pure states it holds that

$$\mathbf{E}_C(|\psi\rangle) = \mathbf{E}_D(|\psi\rangle) = S(\mathrm{Tr}_B(|\psi\rangle\langle\psi|)) = -\mathrm{Tr}(\rho_A \log_2 \rho_A), \tag{6}$$

where $S(\cdot)$ denotes the von Neumann entropy of a state and Tr_B the partial trace over subsystem B and $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$ is the reduced state of subsystem A. This gives rise to a new entanglement measure, the *entanglement of formation*, defined as the convex roof of the von Neumann entropy

$$\mathbf{E}_F(\rho) = \inf_{\mathcal{D}(\rho)} \sum_i p_i S(\mathrm{Tr}_B(|\psi_i\rangle\langle\psi_i|)). \tag{7}$$

For mixed states the situation becomes more complicated. PPT states are *undistillable*, that means that no number of copies is sufficient to create a single maximally entangled qubit state, i.e. $\mathbf{E}_D(\rho) = 0$ [25]. However, there exists entangled states that are PPT for which clearly $\mathbf{E}_C(\rho) > 0$. This means that in general the entanglement of distillation and the entanglement cost are not equal, but only $\mathbf{E}_D(\rho) \leq \mathbf{E}_C(\rho)$ holds.

There exist various entanglement measures and monotones [26]. One of the most prominent entanglement measures (or monotone) is the *concurrence* $C(\cdot)$ [19, 27]. For pure states the concurrence is defined as $C(|\psi\rangle) = \sqrt{2[1 - \mathrm{Tr}(\rho_A)]}$. For mixed states this definition is extended via the *convex-roof construction*, i.e. the infimum of the concurrence of all decompositions $C(\rho) = \inf_{\mathcal{D}(\rho)} \sum_i p_i C(|\psi_i\rangle)$. This measure is faithful, i.e. $\mathbf{E}(\rho) = 0$ iff ρ is separable, and convex $\mathbf{E}(\sum_i p_i \rho_i) \leq \sum_i p_i \mathbf{E}(\rho_i)$. Further it is invariant under local unitary (LU) transformations and non-increasing under LOCC. For two-qubit systems there exists an analytical expression and the concurrence can be written as $C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$, where λ_i are the decreasing eigenvalues of $\sqrt{\sqrt{\rho}(\sigma_Y \otimes \sigma_Y)\rho^*(\sigma_Y \otimes \sigma_Y)\sqrt{\rho}}$. The notion of concurrence can be generalized to the multipartite setting [28]. For pure states on $\bigotimes_{i=1}^N \mathcal{H}_{A_i}$ it is defined as $C_{GME}(|\psi\rangle) = \min_{\mathcal{A}\subset[N]} \sqrt{2[1 - \mathrm{Tr}(\rho_{\mathcal{A}}^2)]}$. Via the convex-roof construction it can again be extended to mixed states $C_{GME}(\rho) = \inf_{\mathcal{D}(\rho)} \sum_i p_i C_{GME}(|\psi_i\rangle)$. It can be shown that $C_{GME}(\rho) = 0$ iff ρ is biseparable and with this to satisfy the above mentioned requirements.

With the previously defined entanglement measures any two pure bipartite states can be compared. However, for more parties there exist inequivalent entanglement classes

with different properties. Two pure states $|\psi\rangle$ and $|\phi\rangle$ are said to belong to the same entanglement class, if they can be transformed into each other via stochastic LOCC (SLOCC). This is equivalent to the existence of a local invertible operator A, such that $|\psi\rangle = A|\phi\rangle$ [29]. For three qubits there exists two classes of entangled states, pure three-party entangled states belong either to the GHZ-class or the W-class [30], but already for four qubits there exists a continuum of different entanglement classes [31].

5.3 Quantum channels

In this section we very briefly discuss quantum channels and three paradigmatic examples thereof. For a more detailed treatment the reader is referred to Refs. [32, 33].

When transmitting or storing quantum information, the system carrying the information inevitably interacts with the environment in a way that is not fully controlled. This introduces noise into the system, which can be modelled as a joint unitary evolution U of the system in state ρ and the environment in state ρ_{E} and tracing out the environment system. Such state transformations $\Lambda(\rho) = \mathrm{Tr}_{\mathrm{E}}(U\rho \otimes \rho_{\mathrm{E}} U^\dagger)$ are called *quantum channels* and are the most general maps from states to states, thus describing all transformations allowed by quantum mechanics. Mathematically, quantum channels correspond to completely positive, trace-preserving (CPTP) maps $\Lambda(\cdot)$ acting on density operators. Every CPTP map $\Lambda(\cdot)$ acting on finite-dimensional density matrices can be written as a *Kraus-operator* decomposition $\Lambda(\rho) = \sum_j K_j \rho K_j^\dagger$, where the K_j are the *Kraus operators* on $\mathcal{H}$ satisfying $\sum_j K_j^\dagger K_j = \mathbb{1}$.

The following are exemplary models for noise channels in quantum information. The *depolarizing channel* $\Lambda(\rho) = \frac{p}{d}\mathbb{1} + (1-p)\rho$ simply probabilistically leaves the state unchanged or replaces it with the maximally mixed state. For qubits, its Kraus oparators are $K_0 = \sqrt{1-\frac{3p}{4}}\mathbb{1}$ and $K_i = \frac{\sqrt{p}}{2}\sigma_i$ with the Pauli matrices σ_i, $i = X, Y, Z$. The action of this channel can be easily visualized in the Bloch ball by a contraction towards the center, uniformly shrinking the set that can be attained after this channel. The maximally mixed state is the only invariant state under the action of this channel.

The *phase damping* or *dephasing channel* $\Lambda(\rho) = \int_{2\pi} p(\theta) U(\theta) \rho U^\dagger(\theta) d\theta$ describes the effect of a unitary evolution with random parameter θ distributed according to $p(\theta)$. Such noise arises when a system evolves for an unknown time and has the effect that the relative phases between eigenstates of the Hamiltonian giving raise to the unitary evolution are lost. Physically this corresponds to a loss of quantum information without a loss of energy. For a qubit system and $U(\theta) = \exp(i\sigma_Z\theta/2)$, the Kraus operators are $K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix}$ and $K_1 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{pmatrix}$. In the Bloch representation this corresponds to a contraction towards the z-axis, resulting in a ellipsoid. Points on the z-axis remain unchanged, as diagonal states

are invariant under the action of this channel.

The *amplitude damping channel* describes the effect of energy dissipation from the system. For qubits this channel has Kraus operators $K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix}$ and $K_1 = \begin{pmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{pmatrix}$. K_0 leaves the groundstate $|0\rangle$ unchanged and reduces the amplitude of the excited state $|1\rangle$, whereas K_0 transforms $|1\rangle$ into $|0\rangle$, which corresponds to loosing a quantum of energy to the environment. In the Bloch sphere the action of this channel is to shrink the volume towards the north pole, representing the ground state as the only fixed point under the action of this channel.

5.4 Quantum metrology

Quantum metrology is the study of precision measurements of physical parameters using quantum systems and exploiting distinct quantum features such as entanglement and nonclassicality. In many scenarios, quantum estimation techniques potentially provide better sensitivity and faster increase in precision with the number or average energy of the probe systems than their classical counterpart [34–38].

5.4.1 Quantum parameter estimation

A key discipline in quantum metrology is parameter estimation. The unknown parameter describes a given transformation and can therefore usually not be accessed and measured directly, but only by the impact of the transformation on another system. For the estimation one has to rely on data obtained from measurements of a probe system. The typical procedure in parameter estimation consists of preparing a suitable system in a specific state and subsequently letting it undergo the transformation encoding the unknown parameter. After the transformation, the probe system carries information about the parameter and by measuring the probe one can access this information. Certain probe states and measurements result in a more efficient estimation of the parameter than others. The exploration of the probe and measurement space to identify optimal strategies is hence of great importance.

By exploiting properties intrinsic to quantum systems, quantum metrology can achieve a better estimation precision at a fixed amount of invested resources. More precisely, the uncertainty of purely classical strategies decreases at most with a rate of $1/\sqrt{n}$, where n denotes the number of probes used, often called the *standard quantum limit*. In comparison, the estimation precision for a general quantum strategy can scale at the so-called *Heisenberg limit* of $1/n$ as $n \to \infty$ [39].

Depending on the transformation encoding the unknown parameter, a suitable probe system has to be chosen. In the standard parameter-estimation scenario, the probe system prepared in the state ρ undergoes a transformation Λ_θ that imprints information about

the parameter θ on the system, afterwards in the state $\rho(\theta)$. The transformation encoding the unknown parameter is described by a completely positive and trace-preserving map, also known as a quantum channel. We assume that the form of the channel is known, except for one continuous parameter. In most applications the unknown transformation is assumed to be a unitary evolution $U_\theta = \mathrm{e}^{-i\theta H}$ with known Hamiltonian H, transforming the probe according to $\rho(\theta) = U_\theta \rho U_\theta^\dagger$. If there is a discrete set of possible channels, the task to determine the channel acting on the system is called channel discrimination, whereas if the set of possible channels is not restricted at all or only to a large subclass of channels we talk about channel estimation. In this sense parameter estimation can be seen as a special case of channel estimation. General quantum measurements are described by a positive-operator valued measure (POVM), i.e. a discrete or continuous collection of positive operators E_m whose sum or integral evaluates to the identity. Performing a measurement with POVM elements $\{E_m\}$ gives the outcome m with probability $p(m|\theta) = \mathrm{Tr}[E_m \rho(\theta)]$ conditioned on the parameter θ. This procedure is repeated, collecting the outcomes $\mathbf{m} = (m_1, \ldots, m_n)$, where the m_i denote the outcomes of the individual rounds. There are two statistical frameworks that are often used in estimation theory. On the one hand, the frequentist (or local) approach offers a rich theoretical formalism and is especially well adapted to scenarios where the initial uncertainty is already low or the number of probes is very large, but is ill defined for very few probes. The Bayesian framework, on the other hand, is less rigid and works in every scenario, making it a viable option for the high uncertainty and low probe number regime, but relies on specifying a prior, i.e., a probability distribution encoding one's initial information (or belief) about the estimated quantity. Both the frequentist and the Bayesian approach give a framework that provide an estimate of the parameter θ from the statistics on the outcomes $\mathbf{m}$, and both can be equipped with a suitable figure of merit for the expected quality of the estimate.

5.4.2 Frequentist estimation

In the frequentist framework an *estimator* $\hat{\theta}(m)$ assigns an estimate for the value of the parameter θ. The estimator is called *unbiased* if the expected value of the estimator equals the value of the parameter, i.e. $\langle \hat{\theta}(m) \rangle = \theta$ and has variance $\mathrm{V}[\hat{\theta}(m)] = \sum_m p(m|\theta)[\hat{\theta}(m) - \theta]^2$. For an increasing number of measurement rounds n the mean value of the estimates $\hat{\theta}(\mathbf{m}) = \frac{1}{n}\sum_i^n \hat{\theta}(m_i)$ converges towards the expected value of the estimator, which for an unbiased estimator corresponds to the value of the parameter. Also the mean square error (MSE) of the estimates $\frac{1}{n}\sum_i [\hat{\theta}(m_i) - \hat{\theta}(\mathbf{m})]$ converges towards the variance of the estimator $\mathrm{V}[\hat{\theta}(m)]$ and with this the estimation precision as quantified by the variance of the mean, which is given as the inverse of $n\mathrm{V}[\hat{\theta}(m)]$ for independent and identically distributed (i.i.d.) measurement outcomes. The *Cramér-Rao bound* gives a lower bound

on the variance of any unbiased estimator, in fact

$$\mathrm{V}[\hat{\theta}(m)] \geq \frac{1}{I[\rho(\theta)]}, \tag{8}$$

where the *Fisher information* is given by $I[\rho(\theta)] = \sum_m p(m|\theta)\left[\frac{\partial}{\partial\theta}\log p(m|\theta)\right]^2$.

Notice that while the estimation precision depends on the variance of the estimator, the Cramér-Rao bound holds for every unbiased estimator and depends solely on the probe state and the performed measurement. An estimator saturating the Cramér-Rao bound is called *efficient*. However, it is not always possible to find an efficient estimator that is*globally unbiased*, i.e. for all values of the parameter. But this is often not necessary, if the estimated parameter is known sufficiently well it suffices to find a *locally unbiased* efficient estimator. This means that the estimator is unbiased in a sufficiently small region around θ. Therefore this approach is also referred to as *local estimation*. The maximum-likelihood estimator saturates the Cramér-Rao bound asymptotically and is asymptotically unbiased. This means that for a sufficiently large number of probes the precision of the estimator behaves like $n/I[\rho(\theta)]$ in good approximation.

The *quantum Fisher information* $\mathcal{I}[\rho(\theta)]$ is defined as the Fisher information maximised over all POVMs and hence depends only on the probe state ρ and the transformation Λ_θ. Clearly the inequality

$$\mathrm{V}[\hat{\theta}(m)] \geq \frac{1}{\mathcal{I}[\rho(\theta)]}, \tag{9}$$

holds and is known as the *quantum Cramér-Rao bound* [37, 38, 40]. The quantum Fisher information equals the expectation value of the squared symmetric logarithmic derivative (SLD) defined implicitly by $S[\rho(\theta)]\rho(\theta) + \rho(\theta)S[\rho(\theta)] = 2\dot{\rho}(\theta)$ and a projective measurement in the eigenbasis of the SLD is always optimal [41]. Hence this framework already provides the optimal strategy for a given probe state. Different probe states can be compared by their quantum Fisher information, which becomes more accessible to calculate for pure states

$$\mathcal{I}[|\psi(\theta)\rangle] = 4(\langle\dot{\psi}(\theta)|\dot{\psi(\theta)}\rangle - |\langle\dot{\psi}(\theta)|\psi(\theta)\rangle|^2). \tag{10}$$

5.4.3 Bayesian estimation

The Bayesian framework is an alternative approach to parameter estimation. In this framework the parameter θ is treated as a random variable with a given probability distribution function $p(\theta)$, called the *prior distribution function*, or simply prior. It encodes all the information about the parameter θ available prior to the measurement. Using Bayes' law the distribution function is updated conditioned on the measurement outcome

m

$$p(\theta|m) = \frac{p(m|\theta)p(\theta)}{p(m)} = \frac{p(m|\theta)p(\theta)}{\int d\mu p(m|\mu)p(\mu)}. \tag{11}$$

As in the frequentist framework, the likelihood to observe the measurement outcome m given that the parameter equals θ is $p(m|\theta) = \mathrm{Tr}[E_m\rho(\theta)]$. The updated distribution of θ is called the *posterior distribution function*, or just posterior. This updating can subsequently be repeated, where the posterior of one measurement round serves as the prior for the next round. In the end we obtain the distribution of θ conditioned on all the measurement outcomes $\mathbf{m} = (m_1, \ldots, m_n)$.

The estimator $\hat{\theta}(\mathbf{m})$ can now simply be calculated from this final distribution, a canonical choice is simply the mean $\hat{\theta}(\mathbf{m}) = \langle\theta\rangle$. The precision of this estimate is given by the variance of the final distribution $\mathrm{V}_{\mathrm{post}}(\mathbf{m}) = \int d\theta p(\theta|\mathbf{m})\ [\theta - \hat{\theta}(\mathbf{m})]^2$

The choice of the prior $p(\theta)$ introduces a subjective element into the estimation process. However, the subjective margin is not so large. If there is no reason to prefer a certain parameter range, a flat or wide prior is chosen that has little influence on the posterior. Additionally, the significance of the prior becomes smaller with increasing number of measurement rounds. On the other side the choice of the prior allows for flexible parameter regions, e.g. excluding unattainable parameter ranges.

6 Overview and discussion

In this section we briefly introduce the research articles that constitute this thesis, published, under peer review and in preparation. We want to highlight the research questions that lead to the individual projects and present the key ideas and main results of each project, together with a brief discussion.

6.1 Bayesian parameter estimation using Gaussian states and measurements

The determination of unknown parameters is one of the most promising applications of quantum technologies in the near future. For instance, parameter estimation will be essential to counteract noise that arises in the preparation, manipulation and transmission of quantum systems and will thus contribute to the success of other future quantum technologies. We have seen that in an idealised scenario, a quantum sensing device can achieve a decrease in uncertainty scaling with the Heisenberg limit of $1/n$. However, the difficulty lies in the practical implementation. To beat classical strategies and achieve Heisenberg scaling typically requires to create some highly correlated/entangled states and to perform global measurements on them. This extremely demanding task is further complicated by the fact that already small noise levels potentially destroy the advantage compared to classical strategies. Although local estimation based on the quantum Cramér-Rao bound provides an extensive framework for investigating optimal strategies, it is sometimes too rigid. Knowing the ideal measurement and probe state is of little practical value, if those strategies are not implementable or extremely fragile. Additionally, the analysis provided by local estimation is often only valid if the parameter to be estimated is known sufficiently well.

By starting from a set of states and measurements available with current technology and investigating their potential for parameter estimation, in this project we shift the focus from the search for probe states and measurements scaling optimally with energy towards a more experimentally feasible approach. Instead of creating a theoretical framework that has to be adapted and approximated for applications, we identify efficient parameter estimation strategies based solely on technologies that are by now common practice in the laboratory. Bayesian analysis allows us to investigate the performance of these strategies within a less rigid framework than local parameter estimation focusing on the preparability of the probe states instead of their optimal energy scaling. In our case, where the probe system consists of bosonic modes, we limit the set of states to single-mode Gaussian states, as these states combine various advantages. They can be straightforwardly prepared and controlled in the laboratory, have a compact mathematical description in phase space and their intrinsic robustness towards noise makes our analysis still valid for

reasonably noisy environments, thus strengthening the promise of near future applicability of our results. After our probe system undergoes the transformation, a measurement is performed. In our investigation we considered heterodyne and homodyne detection as potential measurements, as their outcomes are Gaussian distributed when measuring a Gaussian state and they share all the advantages previously discussed for Gaussian states. Within this setting, we investigate three paradigmatic cases of continuous-variable quantum metrology: the estimation of displacements, phase rotations, and single-mode squeezing strengths.

This project is a step towards the creation of a universal framework for parameter estimation in continuous variable quantum information with the potential of near future experimental implementation. We provided a comprehensive investigation of Bayesian parameter estimation with single-mode Gaussian states and suitable Gaussian measurements, analysing what can be achieved with practically realizable techniques instead of trying to maximise the quantum Fisher information. This allows us to identify strategies that combine good performance with the potential for straightforward experimental realization. Our results provide practical solutions for reaching uncertainties where local estimation techniques apply, representing an important connection to the respective local estimation problems and thus bridging the gap to regimes where asymptotically optimal strategies can be employed. Besides the relevance for experimental implementations, this investigation also creates a significant reference point for future explorations of more complicated probe states and measurements within the theory of Bayesian estimation. We envisage the results presented as a first step in the exploration of Gaussian probe states and measurements in the framework of Bayesian parameter estimation.

6.2 Metrology-assisted entanglement distribution in noisy quantum networks

Entanglement shared between distant parties is a valuable resource for many tasks in quantum information and quantum metrology. One way of distributing entanglement in a quantum network is via a central node, able to prepare a multipartite system in a highly entangled state and successively distributing it to the individual users. In a realistic scenario the state received will likely not correspond to the target state for the protocol the users wish to perform, but can be converted into the desired state. The most general transformations that the parties can perform are described by stochastic local operations assisted by classical communication. These transform the initial state into the target state probabilistically, where the unsuccessful conversions are usually discarded as they do not belong to the same entanglement class anymore. Sending quantum states over longer distances introduces unknown noise into the transmitted states. Before utilizing the states the noise has to be estimated and counteracted, which can be done by initially sacrificing

a certain number of states. For channels that can be used only for a restricted period of time for transmission or varies continuously over time, e.g. quantum communication via satellite, this could severely limit the transmission rate.

This project investigates the potential of utilizing unsuccessfully converted states for channel estimation. Since after failed conversion the system still carries information about the transmission channel, the idea is to access this information and so estimate the channel, without the necessity of sacrificing useful states. We focus on one-successful-branch protocols, that are almost always optimal for single-copy pure-state transformations. After the state conversion, the output states of the successful branches are saved for further use, while a measurement is performed on the failure branch to estimate the noise channel with a Bayesian estimation strategy. Once the channel is sufficiently well estimated, the effect of the noise can be counteracted on the previously stored output states of the successful branches. Inverting the order of the state conversion and noise counteraction is possible in those situations where the Kraus operators describing the channel and the POVM measurements describing the one-successful-branch protocol commute. The presented strategy operates sequentially, which facilitates its implementation as each copy is processed independently and it does not rely on quantum memories. We provide an example, where we assume a pure qudit state in the GHZ entanglement class is distributed via a noisy network with local dephasing noise. We compare our strategy to 'naive' protocols, that sacrifice a number of probes for channel estimation. Instead of evaluating a specific strategy for such naive protocols, we derive upper bounds for all such strategies via the quantum Fisher information and entropy-based distillation bounds, both for sequential strategies or ones with access to quantum memories. Finally we present a protocol that works also in cases where the noise does not commute with the state conversion. The idea is to continuously link the 'naive' approach to the previously presented strategy, where the estimation precision after each step determines the procedure in the following.

In this article we show that probabilistic state-conversion protocols in noisy networks can be improved by embedded channel-estimation routines at no additional cost in terms of the number of copies of the distributed states. We provide a concrete example, illustrating the usefulness of our approach, and compare it to traditional strategies. This proof of principle is but the first step into the exploration in the direction of entanglement distribution with integrated estimation techniques. As the usefulness of our approach strongly depends on the specific noise model, the identification of viable strategies remains an interesting open challenge in more general scenarios.

6.3 Activation of genuine multipartite entanglement: Beyond the single-copy paradigm of entanglement characterisation

Multipartite entanglement is a central resource for potential applications in quantum information science. Although states shared between different parties can be entangled in numerous ways, one often distinguishes only between genuinely multipartite entangled and biseparable states, where the latter are considered less valuable. Any practical protocol that aims at distributing entanglement will send numerous copies of the same state to the involved parties. Hence it becomes natural to characterize entanglement for more than one copy. At the same time it is known that biseparability of a state is not a tensor-stable property.

We examine the possibility to activate GME from multiple copies of biseparable states and give an explicit example where this is indeed possible. For the family of isotropic GHZ states, which are convex mixtures of GHZ states with maximally mixed states, we show that within a certain parameter range two copies of biseparable states become GME. We further show that for a certain parameter range more than two copies are necessary to observe this effect and construct an upper bound on the number of copies needed for the activation of GME for all states of this family that are biseparable but not partition separable. This leads us to conjecture that the same holds for all biseparable but not partition-separable states, namely that there exists a number of copies that activates GME. We additionally conjecture the existence of a hierarchy of states with k-copy activatable GME, that cannot be activated with a number of copies less than k. This effect has to be distinguished from entanglement distillation. In fact we give an example of a biseparable state that only exhibits bound entanglement across all cuts. Nonetheless two copies of this state become GME.

These results shows that the characterization of multipartite entanglement exhibits some surprising behaviour when we consider multiple copies. We conjecture a hierarchy of states based on the number of copies needed to activate GME. At first sight this hierarchy would make the characterization of entanglement even more complicated, but, if true, in the limit of infinitely many copies the hierarchy would collapse, reducing the characterization of multipartite entanglement to that of detecting partition separability.

6.4 Entanglement detection with imprecise measurements

Certifying entanglement shared between parties is an important task essential for the success of numerous applications in quantum communication and cryptography. To detect entanglement the involved parties usually perform local measurements and compare the obtained outcomes. Although there exists the possibility of certifying entanglement in a

device independent way via the violation of a Bell-inequality, this is often excessive and unnecessary for non malicious scenarios. Assuming full control over the measurement devices, entanglement witnesses provide an practical tool to potentially detect any entanglement shared between parties. However, this idealized scenario falls short of describing the actual situation in the laboratory, where the performed measurements only come close to those intended.

Assuming imperfect control over the measurement devices, we investigate entanglement detection with bounded distrust. As the measurement directions are not arbitrary, we still want to exploit as much information as available over the performed measurements. For this, we first introduce an operational notion of inaccuracy, that allows us to examine the performance of several prominent entanglement witnesses under conditions expected to be encountered in the laboratory. We quantify the correspondence between the target measurements and the actual measurements through their average fidelity, that can be estimated experimentally, and compute tight corrections to the examined witnesses. With this we create a single theoretical framework to describe all ranges of distrust in the measurement devices, at the ends of which sit two previously unconnected scenarios. One is the well-studied situation of perfect control over the measurement apparatus, while the other assumes that only the Hilbert-space dimension is known and the experimenter has no control over the relevant degrees of freedom. Finally we develop a semidefinite programming method to calculate upper bounds on witnesses, both for separable as for entangled states.

In this project we show that already small misalignments can substantially compromise entanglement witnesses. We provide a experimentally determinable notion of inaccuracy and show how to compute tight corrections to a family of entanglement witnesses for arbitrary dimensional two-party systems. Our results open the door to the exploration of similar scenarios with different notions of inaccuracies and provide a starting point for the search of experimentally more robust entanglement witnesses.

6.5 The shape of higher-dimensional state space: Bloch-ball analog for a qutrit

The geometry of the quantum state space has been extensively studied, especially for the most simple system of a qubit. The Bloch ball gives an elegant and intuitive representation of its state space, with the luxury of having a one-to-one correspondence between points in the Bloch ball and states. This simplicity is at the same time also a weakness of the model, as it falls short of representing many properties and characteristics of general quantum state spaces. Quantum state spaces in higher dimensions share many properties with the Bloch ball, but have a much richer structure in general. With the

advance of quantum-information sciences it becomes clear that this richer structure can lead to advantages for different tasks, e.g. error correction, increased noise resistance in entanglement distribution or quantum cryptography. It is hence important to gain a better understanding of higher-dimensional systems and have the tools at hand to represent transformations thereof.

In this project we develop two three-dimensional Bloch models for a qutrit system, which capture many of the fundamental geometric and algebraic properties of the actual eight-dimensional state space and can therefore be used as powerful tools when exploring such systems. By choosing the basis of the representation as the eigenbasis of a state, the convex combination with any other state can be represented by a straight line. Also the action of unitary operations can be visualized, the unitary orbit forms a subset of a sphere about the origin, showing an interesting connection to the action of doubly-stochastic matrices on classical probability distributions. The resulting Birkhoff polytope as orbit of a diagonal state appears as the 'shadow' of the unitary orbit. Finally, we study the action of three paradigmatic noise channels, i.e. the depolarizing channel, the phase-damping channel and the amplitude-damping channel, on the qutrit state space. The action of these channels can be visualized in our model and strongly resembles the known image for qubits.

References

[1] Ashley Montanaro, *Quantum algorithms: an overview,* npj Quantum Information **2**, 15023 (2016), arXiv:1511.04206.

[2] H. Jeff Kimble, *The quantum internet,* Nature **453**, 1023 (2008), arXiv:0806.4195.

[3] Wolfgang Dür, Raphael Lamprecht, and Stefan Heusler, *Towards a quantum internet,* European Journal of Physics **38**, 043001 (2017).

[4] Stephanie Wehner, David Elkouss, and Ronald Hanson, *Quantum internet: A vision for the road ahead,* Science **362**, eaam9288 (2018).

[5] Marcello Caleffi, Angela S. Cacciapuoti, and Giuseppe Bianchi, *Quantum internet,* Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication (2018), arXiv:1805.04360.

[6] Stefan Bäuml and Koji Azuma, *Fundamental limitation on quantum broadcast networks,* Quantum Science and Technology **2**, 024004 (2017), arXiv:1609.03994.

[7] Michael Epping, Hermann Kampermann, and Dagmar Bruß, *Large-scale quantum networks based on graphs,* New Journal of Physics **18**, 053036 (2016), arXiv:1504.06599.

[8] Michael Epping, Hermann Kampermann, and Dagmar Bruß, *Robust entanglement distribution via quantum network coding,* New Journal of Physics **18**, 103052 (2016), arXiv:1605.08384.

[9] Marcello Cacciapuoti, Angela S.and Caleffi, Francesco Tafuri, Francesco Saverio Cataliotti, Stefano Gherardini, and Giuseppe Bianchi, *Quantum Internet: Networking Challenges in Distributed Quantum Computing,* IEEE Network **34**, 137 (2020), arXiv:1810.08421.

[10] Mark Hillery, Vladimír Bužek, and André Berthiaume, *Quantum secret sharing,* Physical Review A **59**, 1829 (1999), arXiv:quant-ph/9806063.

[11] Remigiusz Augusiak and Paweł Horodecki, *Multipartite secret key distillation and bound entanglement,* Physical Review A **80** (2009), arXiv:0811.3603.

[12] Michael Epping, Hermann Kampermann, Chiara Macchiavello, and Dagmar Bruß, *Multi-partite entanglement can speed up quantum key distribution in networks,* New Journal of Physics **19**, 093012 (2017), arXiv:1612.05585.

[13] Matej Pivoluska, Marcus Huber, and Mehul Malik, *Layered quantum key distribution,* Physical Review A **97**, 032312 (2018), arXiv:1709.00377.

[14] Peter Kómár, Erik M. Kessler, Michael Bishof, Liang Jiang, Anders S. Sørensen, Jun Ye, and Mikhail D. Lukin, *A quantum network of clocks,* Nature Physics **10**, 582 (2014), arXiv:1310.6045.

[15] Fernando G. S. L. Brandão and Michał Horodecki, *An area law for entanglement from exponential decay of correlations,* Nature Physics **9**, 721 (2013), arXiv:1504.06599.

[16] Sebastian Ecker, Frédéric Bouchard, Lukas Bulla, Florian Brandt, Oskar Kohout, Fabian Steinlechner, Robert Fickler, Mehul Malik, Yelena Guryanova, Rupert Ursin, and Marcus Huber, *Overcoming Noise in Entanglement Distribution,* Physical Review X **9**, 041042 (2019), arXiv:1904.01552.

[17] Reinhold A. Bertlmann and Philipp Krammer, *Bloch vectors for qudits,* Journal of Physics A: Mathematical and Theoretical **41**, 235303 (2008), arXiv:0806.1174.

[18] Dagmar Bruß, *Characterizing entanglement,* Journal of Mathematical Physics **43**, 4237 (2002), arXiv:quant-ph/0110078.

[19] Otfried Gühne and Géza Tóth, *Entanglement detection,* Physics Reports **474**, 1 (2009), arXiv:0811.2803.

[20] Leonid Gurvits, *Classical Deterministic Complexity of Edmonds' Problem and Quantum Entanglement,* in *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03 (ACM, New York, NY, USA, 2003) pp. 10–19.

[21] Leonid Gurvits, *Classical complexity and quantum entanglement,* Journal of Computer and System Sciences **69**, 448 (2004), special Issue on STOC 2003, arXiv:quant-ph/0303055.

[22] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki, *Separability of mixed states: necessary and sufficient conditions,* Physics Letters A **223**, 25 (1996), arXiv:quant-ph/9605038.

[23] Fabien Clivaz, Marcus Huber, Ludovico Lami, and Gláucia Murta, *Genuine-multipartite entanglement criteria based on positive maps,* Journal of Mathematical Physics **58**, 082201 (2017), arXiv:1609.08126.

[24] Michael A. Nielsen, *Conditions for a Class of Entanglement Transformations,* Physical Review Letters **83**, 436 (1999), arXiv:quant-ph/9811053.

[25] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki, *Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?* Physical Review Letters **80**, 5239 (1998), arXiv:quant-ph/9801069.

[26] Martin B. Plenio and Shashank Virmani, *An introduction to entanglement measures,* Quantum Information & Computation **7**, 1 (2007), arXiv:quant-ph/0504163.

[27] William K. Wootters, *Entanglement of Formation of an Arbitrary State of Two Qubits,* Physical Review Letters **80**, 2245 (1998), arXiv:quant-ph/9709029.

[28] Zhi-Hao Ma, Zhi-Hua Chen, Jing-Ling Chen, Christoph Spengler, Andreas Gabriel, and Marcus Huber, *Measure of genuine multipartite entanglement with computable lower bounds,* Physical Review A **83**, 062325 (2011), arXiv:1101.2001.

[29] David Sauerwein, Nolan R. Wallach, Gilad Gour, and Barbara Kraus, *Transformations among Pure Multipartite Entangled States via Local Operations are Almost Never Possible,* Physical Review X **8**, 031020 (2018), arXiv:1711.11056.

[30] Wolfgang Dür, Guifré Vidal, and Juan Ignacio Cirac, *Three qubits can be entangled in two inequivalent ways,* Physical Review A **62**, 062314 (2000), arXiv:quant-ph/0005115.

[31] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, *Four qubits can be entangled in nine different ways,* Physical Review A **65**, 052112 (2002), arXiv:quant-ph/0109033.

[32] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).

[33] Filippo Caruso, Vittorio Giovannetti, Cosmo Lupo, and Stefano Mancini, *Quantum channels and memory effects,* Reviews of Modern Physics **86**, 1203 (2014), arXiv:1207.5435.

[34] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, *Quantum Metrology,* Physical Review Letters **96**, 010401 (2006), arXiv:quant-ph/0509179.

[35] Matteo G. A. Paris, *Quantum estimation for quantum technology,* International Journal of Quantum Information **7**, 125 (2009), arXiv:0804.2981.

[36] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, *Advances in quantum metrology,* Nature Photonics **5**, 222 (2011), arXiv:1102.2318.

[37] Géza Tóth and Iagoba Apellaniz, *Quantum metrology from a quantum information science perspective,* Journal of Physics A: Mathematical and Theoretical **47**, 424006 (2014), arXiv:1405.4878.

[38] Rafał Demkowicz-Dobrzański, Marcin Jarzyna, and Janek Kołodyński, *Quantum limits in optical interferometry,* Progress in Optics **60**, 345 (2015), arXiv:1405.7703.

[39] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit,* Science **306**, 1330 (2004), arXiv:quant-ph/0412078.

[40] Nicolai Friis, Davide Orsucci, Michalis Skotiniotis, Pavel Sekatski, Vedran Dunjko, Hans J. Briegel, and Wolfgang Dür, *Flexible resources for quantum metrology,* New Journal of Physics **19**, 063044 (2017), arXiv:1610.09999.

[41] Samuel L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states,* Physical Review Letters **72**, 3439 (1994).

7 Bayesian parameter estimation using Gaussian states and measurements

The following research article was published in *Quantum Science and Technology, Volume 6, Number 2* the 5th March 2021 with the DOI 10.1088/2058-9565/abd83d.

7.1 Contribution

I contributed significantly to the creation of the theoretical framework of this research project, especially the application of the Bayesian formalism to the intended parameter estimation setting. The analytical investigation for the examples of displacement (Section III.) and squeezing estimation (Section V.) were carried out by me, as well as the numerical calculations for the squeezing estimation, where I wrote the code for the numerical analysis. Further I wrote the main text of the manuscript and took care of the revision.

Bayesian parameter estimation using Gaussian states and measurements

Simon Morelli,[1,*] Ayaka Usui,[2,†] Elizabeth Agudelo,[1,‡] and Nicolai Friis[1,§]

[1] *Institute for Quantum Optics and Quantum Information - IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria*

[2] *Quantum Systems Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan*

(Dated: February 2, 2022)

Bayesian analysis is a framework for parameter estimation that applies even in uncertainty regimes where the commonly used local (frequentist) analysis based on the Cramér-Rao bound is not well defined. In particular, it applies when no initial information about the parameter value is available, e.g., when few measurements are performed. Here, we consider three paradigmatic estimation schemes in continuous-variable quantum metrology (estimation of displacements, phases, and squeezing strengths) and analyse them from the Bayesian perspective. For each of these scenarios, we investigate the precision achievable with single-mode Gaussian states under homodyne and heterodyne detection. This allows us to identify Bayesian estimation strategies that combine good performance with the potential for straightforward experimental realization in terms of Gaussian states and measurements. Our results provide practical solutions for reaching uncertainties where local estimation techniques apply, thus bridging the gap to regimes where asymptotically optimal strategies can be employed.

I. INTRODUCTION

Quantum sensing devices hold the promise of outperforming their classical counterparts. However, since classical strategies can achieve arbitrary precision, provided that sufficiently many independent probes are used, the advantage of quantum sensing devices does not lie in the achievable precision. Instead, quantum strategies provide a faster increase in precision with n, the number of probes. In an idealised quantum sensing scenario, the estimation precision can in principle scale at the so-called Heisenberg limit (HL) of $1/n$ as $n \to \infty$. In contrast, classical strategies can at most achieve a precision scaling of $1/\sqrt{n}$, the so-called standard quantum limit (SQL).

In the context of quantum optics, which we are interested in here, the possibility of preparing states with uncertain photon number means that the number of probes is uncertain. Therefore, the scaling usually refers to resources such as the mean photon number or mean energy of the probe systems. Nevertheless, general quantum strategies can result in a quadratic scaling advantage and thus outperform 'classical' strategies using the same resources. However, two important factors have to be considered.

First, preparing optimal or at least close to optimal probes and carrying out the corresponding joint measurements can be complicated and technologically demanding. Moreover, in the presence of uncorrelated noise the scaling advantage with increasing n persists only up to a certain point, beyond which only a (potentially high) constant advantage remains [1–3]. Even if one disregards any additional costs that might incur from trying to combat noise [4, 5], overheads from complex preparation procedures and the resulting low probe state fidelities may thus invalidate the expected benefits. Consequently, it is important to identify estimation strategies that can outperform 'classical' approaches while being feasibly implementable as well as robust against noise. For instance, for estimation problems in continuous-variable (CV) systems, Gaussian states and measurements are generally considered to be comparably easily implementable. They allow achieving the HL for many scenarios within the local, also called 'frequentist', paradigm, including the local estimation of phases, displacements, squeezing and others [6–15].

Second, many of these insights are based on the Cramér-Rao bound (CRB). The CRB applies for estimation with unbiased estimators. It provides a lower bound for the precision via the inverse Fisher information (FI). Estimators that are unbiased locally (i.e., for specific parameter values) are readily available, but profiting from their unbiasedness requires precise prior information on the estimated parameter. The 'local' approach is therefore only well-justified when the number of independent probes is sufficiently large (hence 'frequentist'), in such a case, the CRB provides the asymptotically achievable limit on scaling. However, when the available number of probes is limited (some authors [16–18] refer to 'limited data' in this context) then local estimation is not well defined. Resulting pathologies can lead to scaling seemingly better than the HL [19, 20] and even to an unbounded FI for finite average photon numbers [21]. The available prior information also has to be carefully considered when calculating the CRB. For instance, for phase estimation with $N00N$-states, a growing (average) photon number n implicitly assumes that the prior interval is narrowing with $2\pi/n$. If this is not accounted for, part of the scaling advantage comes from the increasing prior information, as pointed out in Refs. [22, 23].

This motivates the study of Bayesian estimation ap-

* simon.morelli@oeaw.ac.at
† ayaka.usui@oist.jp
‡ elizabeth.agudelo@oeaw.ac.at
§ nicolai.friis@univie.ac.at

proaches for quantum sensing, which we consider here. In Bayesian estimation, one's initial knowledge of the parameter is described by a probability distribution (the prior) which is updated as more measurement data becomes available. The Bayesian approach is valid for an arbitrary number of probes and can in this sense be considered to be more rigorous than local estimation, at the cost of introducing a dependence on the prior. However, the influence of the prior vanishes for larger number of measurements, since the prior knowledge becomes less and less relevant with growing amount of measurement data. In practice, one may pursue a hybrid strategy, where initial Bayesian estimation is employed to sufficiently narrow down the possible range of the parameter before switching to a local estimation strategy with many repetitions.

Here, we consider Bayesian estimation scenarios for quantum optical fields. While much progress has been made for CV parameter estimation within the local paradigm, in particular, regarding the calculation of the quantum Fisher information (QFI) [6–9, 11–15] and the associated optimal strategies achieving the CRB [24–29], CV parameter estimation in the Bayesian setting is much less explored. There, recent work has provided insight into Bayesian estimation with discrete [30] and CV systems using some specific probe states, including coherent states [16–18, 31], $N00N$ states [16, 32], and single-photon states [33]. Determining efficient and practically realizable strategies for Bayesian estimation in quantum optical systems can thus be considered an important link in the development of quantum sensing technologies, which this paper aims to establish.

Within the Bayesian paradigm, the additional freedom represented by the choice of the prior exacerbates the difficulty of determining optimal estimation strategies, making it all the more necessary to identify practically realizable strategies that can also be easily adapted. Here, in particular, we are interested in identifying strategies for Bayesian estimation considering Gaussian states and Gaussian measurements. Gaussian states not only permit an elegant mathematical description in phase space, but are also especially easy to realise experimentally and are by now broadly used [34, 35]. Gaussian measurements, i.e., homodyne or heterodyne detection, have been shown to outperform number detection for few repetitions [17] and to be more robust against noise [27, 36, 37] than photon number detection or 'on/off' detection—which discriminates only between the absence or presence of photons.

To broadly investigate the performance of Gaussian states and measurements in Bayesian metrology, we consider three paradigmatic problems: the estimation of phase-space displacements, phase estimation, and the estimation of single-mode squeezing. For each task, we provide practically realisable strategies based on single-mode Gaussian states combined with homodyne or heterodyne detection that allow efficiently narrowing the prior to the point where local estimation strategies may take over. To set the stage for this investigation, we briefly review the method of Bayesian estimation and relevant concepts of Gaussian quantum optics in Sec. II. In Sec. III, we focus on the estimation of displacements for Gaussian priors, and provide analytical results for the achievable precision using single-mode Gaussian states for both homodyne and heterodyne detection. In Secs. IV and V, we proceed with similar investigations of Bayesian estimation of phases and squeezing parameters, where we compare the performance of squeezing and displacement of the probe system. Finally, we discuss our results and provide an outlook and conclusions in Sec. VI.

II. FRAMEWORK

In this section, we provide a brief overview of the relevant concepts in Bayesian estimation (Sec. II.A) and Gaussian quantum optics (Sec. II.B), before we present our results in the following sections. For a more extensive overview of classical Bayesian estimation theory we refer to [38–40], while more details on local and Bayesian estimation in the quantum setting can be found, e.g., in the appendix of [41].

II.A. Bayesian quantum parameter estimation

II.A.1. The Bayesian estimation scenario

The framework of Bayesian parameter estimation revolves around updating initially available information (or a previously held belief) based on new measurement data via Bayes' theorem, as we will explain in the following. The initial knowledge of the estimated parameter θ is encoded in a probability distribution $p(\theta)$ called the *prior distribution function* or 'prior' for short. It captures all our beliefs (system properties, expertise) and information (prior experimental data) about the system under investigation. When a measurement is performed on the system, the probability $p(m|\theta)$ to observe the measurement outcome m in a system characterised by the parameter θ is called the *likelihood*, and can be calculated from the properties of the model used to describe the system and the measurement. Combined with the prior $p(\theta)$, the likelihood leads one to expect the outcome m with probability

$$p(m) = \int d\theta \; p(m|\theta)\, p(\theta), \tag{1}$$

where the integral is over the support of the prior and it is to be understood as a sum in case of a discrete parameter. The conditional probability that the estimated parameter equals θ, given that measurement outcome m was observed, can then be calculated via Bayes' law, i.e.,

$$p(\theta|m) = \frac{p(m|\theta)\, p(\theta)}{p(m)}. \tag{2}$$

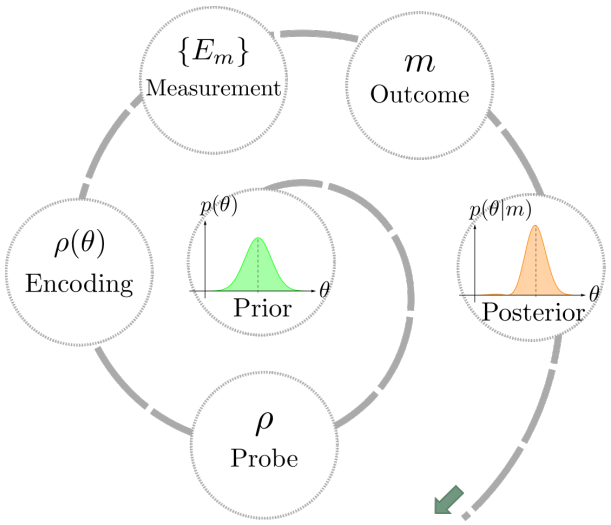

FIG. 1. Bayesian Quantum Parameter Estimation. In Bayesian estimation scenarios, prior information encoded in a probability distribution $p(\theta)$ is updated based on available measurement data such as observing a particular measurement outcome m, resulting in a posterior conditional probability distribution $p(\theta|m)$. In quantum parameter estimation, the measurement procedure consists of preparing the system in a probe state ρ on which the parameter θ is encoded by a suitable transformation. The measurement is represented by a positive-operator valued measure (POVM) with elements E_m representing the possible outcomes m.

The function $p(\theta|m)$ is called the *posterior distribution* of the system parameter, after we have updated our belief with newly available data. The updating procedure, illustrated in Fig. 1, can be repeated arbitrary many times, where the posterior of one step serves as the prior in the next step and the measurement procedure leading to $p(m|\theta)$ can in principle also be adapted from step to step.

After concluding the measurements, the posterior distribution represents a complete description of all available information about the parameter. Nevertheless, it is often desirable (even if not strictly necessary) to nominate an estimator $\hat{\theta}$ and a suitable variance to express the result of the estimation procedure. While the estimator assigns a specific value for θ to any prior or posterior, the variance quantifies the associated uncertainty in the estimate. For parameters $\theta \in \mathbb{R}$, the canonical choice for an estimator is the mean value of the posterior distribution

$$\hat{\theta}(m) = \langle\theta\rangle = \int d\theta\; p(\theta|m)\,\theta. \tag{3}$$

In this case, a valid figure of merit for the confidence in this estimate is the variance of the posterior

$$V_{\text{post}}(m) = \int d\theta\; p(\theta|m)\left[\theta - \hat{\theta}(m)\right]^2. \tag{4}$$

A wide posterior with large variance suggests there is still high uncertainty in our belief about the parameter, whereas a narrow distribution with small variance indicates high confidence in our estimator. Since the variance of the posterior generally depends on the measurement outcome, a good figure of merit for the expected confidence in the estimate provided by a particular measurement strategy is the average variance of the posterior,

$$\bar{V}_{\text{post}} = \int dm\; p(m)\, V_{\text{post}}(m), \tag{5}$$

which we will use here to quantify the precision of the estimation process. However, note that in some cases, the mean and mean square error variance above need to be replaced by more appropriate quantifiers. For instance, in the case that the parameter in question is a phase, where $\theta = -\pi$ and $\theta = \pi$ are identified, $\hat{\theta}(m)$ and $V_{\text{post}}(m)$ can be replaced by suitable alternatives, as we will discuss in in Sec. IV. In any given setting, the task is then to determine estimation strategies that provide sufficiently high precision.

The precision of the estimation procedure generally depends on the shape of the prior, which can in principle be an arbitrarily complicated distribution. Uninformative priors generally influence the outcome less than narrow priors, so one should always be careful which amount of information should be encoded in the prior. However, the influence of the prior on the final estimate generally reduces with increasing number of measurements, and can be argued to become irrelevant asymptotically, see, e.g., [38, Chapter 13]. Consequently, encoding one's knowledge only approximately using a family of probability distributions with only few degrees of freedom can help to facilitate a more straightforward evaluation of the performance of the chosen strategy, while preserving its qualitative features.

For instance, a class of probability distributions is said to be conjugate to a given likelihood function, if priors from within this class result in posterior distributions that belong to that class as well. Choosing the prior to be conjugate to the likelihood in this way makes the updating particularly easy, since this only requires the parameters to be updated to define the posterior distribution uniquely within the chosen class of probability functions, instead of requiring an entirely new calculation to determine the posterior. Gaussian distributions are self-conjugate with respect to the mean, e.g. for Gaussian likelihood functions encoding the parameter to be estimated in their mean, the class of conjugate priors are Gaussian distributions as well. The following proposition is a well known result in statistical theory [38–40, 42].

Proposition 1. *Let the likelihood be Gaussian distributed,* $p(m|\theta) = \mathcal{N}_m\big(\bar{m}(\theta), \tilde{\sigma}^2\big) \propto \mathcal{N}_\theta\big(\bar{\theta}(m), \sigma^2\big)$*, where* $\bar{\theta}(m)$ *is the mean of the distribution in* θ*, the parameter to be estimated. Then a Gaussian prior is the natural conjugate, i.e., if the prior is Gaussian distributed with* $p(\theta) = \mathcal{N}_\theta(\mu_0, \sigma_0^2)$*, the posterior distribution* $p(\theta|m)$ *is also Gaussian with mean value* $\mu_p = \big[\sigma^2\mu_0 + \sigma_0^2\bar{\theta}(m)\big]/(\sigma_0^2 + \sigma^2)$ *and variance* $\sigma_p^2 = (\sigma^2\sigma_0^2)/(\sigma_0^2 + \sigma^2)$*.*

II.A.2. Bayesian estimation using quantum systems

The framework of Bayesian estimation can easily be applied to a quantum setting, as illustrated in Fig. 1. In this case the parameter θ one is interested in estimating is encoded by a transformation that can generally be a completely positive and trace-preserving (CPTP) map. However, in many cases, including those we study here, the transformation is considered to be a unitary U_θ that acts on an initially prepared probe state, represented by a density operator ρ. The resulting encoded state is then given by $\rho(\theta) = U_\theta \rho U_\theta^\dagger$. The measurement of the encoded state can then be represented by a positive operator-valued measure (POVM) with elements $E_m \geq 0$, whose integral (or sum in case of a discrete set of possible measurement outcomes m) evaluates to the identity on the Hilbert space of the probe, i.e., $\int dm\, E_m = \mathbb{1}$. In the quantum case the likelihood is then given by $p(m|\theta) = \mathrm{Tr}\big[E_m \rho(\theta)\big]$.

In local estimation scenarios with unbiased estimators $\hat{\theta}$, the CRB gives a lower bound for the variance of the estimator in terms of the inverse Fisher information $I\big[p(m|\theta)\big]$, that is, $V(\hat{\theta}) \geq I[p(m|\theta)]^{-1}$. Here, the Fisher information depends only on the likelihood function and is given by

$$I\big[p(m|\theta)\big] = \int dm\, p(m|\theta) \Big[\tfrac{\partial}{\partial\theta} \log p(m|\theta)\Big]^2. \tag{6}$$

In the asymptotic limit of infinite sample size, the CRB is always tight, since it is saturated by the maximum likelihood estimator, which becomes unbiased in this limit, see e.g., [43]. Any local estimation problem can thus be reduced to determining an estimation strategy with a likelihood $p(m|\theta)$ corresponding to as large a FI as possible. In the quantum setting, this leaves us with the task of determining suitable probe states ρ and measurements $\{E_m\}_m$. The optimisation of the FI over all POVMs can be carried out analytically, leading to the quantum Fisher information (QFI) $\mathcal{I}\big[\rho(\theta)\big]$, and the corresponding quantum CRB [27, 44], $V(\hat{\theta}) \geq 1/\mathcal{I}\big[\rho(\theta)\big]$. The QFI can be expressed in terms of the Uhlmann fidelity $\mathcal{F}(\rho_1, \rho_2) = \big(\mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\big)^2$ as

$$\mathcal{I}\big[\rho(\theta)\big] = \lim_{d\theta\to 0} 8\, \frac{1 - \sqrt{\mathcal{F}\big[\rho(\theta), \rho(\theta + d\theta)\big]}}{d\theta^2}. \tag{7}$$

For the Bayesian estimation scenario, a similar bound exists. The *Van Trees inequality* bounds the average variance from below according to

$$\bar{V}_{\mathrm{post}} \geq \frac{1}{I\big[p(\theta)\big] + \bar{I}\big[p(m|\theta)\big]}, \tag{8}$$

where $I\big[p(\theta)\big] = \int d\theta\, p(\theta) \big[\frac{\partial}{\partial\theta} \log p(\theta)\big]^2$ is the FI of the prior and $\bar{I}\big[p(m|\theta)\big] = \int d\theta\; I\big[p(m|\theta)\big] p(\theta)$ is the average FI of the likelihood [45, 46]. This inequality is often referred to as the Bayesian Cramér-Rao bound, see, e.g., [47]. In contrast to the CRB in the local scenario, this bound is not tight, which means there might not exist a strategy achieving the equality.

In a Bayesian quantum estimation problem, the Van Trees inequality can be modified to a Bayesian version of the quantum CRB by noting that the FI is bounded from above by the QFI, $\mathcal{I}\big[\rho(\theta)\big] \geq I\big[p(m|\theta)\big]$. Moreover, if the parameter to be estimated is encoded by a unitary transformation U_θ, the QFI is independent of θ. Consequently, the average FI can be bounded by the QFI to obtain the Bayesian quantum CRB

$$\bar{V}_{\mathrm{post}} \geq \frac{1}{I\big[p(\theta)\big] + \mathcal{I}\big[\rho(\theta)\big]}, \tag{9}$$

which gives a lower bound for the average variance for all possible POVMs [41]. As before with Eq. (8), this bound is not tight.

While well-known methods for constructing optimal POVMs for fixed probe states exist for local estimation, optimization of the probe state and measurements for Bayesian estimation has to be carried out on a case-by-case basis and is typically challenging. At the same time, states and measurements that are optimal for a given prior may require complicated preparation procedures while generally no longer being optimal after even a single update. Consequently, it is of interest to devise measurement strategies for Bayesian estimation that are easily realizable and provide 'good' performance for different priors. Here, we provide and examine such strategies for a range of estimation problems in quantum optical scenarios.

II.B. Gaussian quantum optics

As we established before, we are interested in the analysis of scenarios where probe states are quantum states of the electromagnetic field. In particular, our goal is studying the performance of Gaussian states. To set the stage for this investigation, we will here briefly summarize the relevant concepts of Gaussian quantum optics. For a more extensive treatment of CV systems and Gaussian quantum optics we refer the reader to the Refs. [48, 49] and for the particular context of quantum information processing cf. Refs. [50–55]. Multimode optical fields can be represented as collections of bosonic modes. We consider a CV system that consists of N bosonic modes, i.e., N quantum harmonic oscillators. To each mode, labelled k, one associates a pair of annihilation and creation operators, $\hat{a}_k$ and $\hat{a}_k^\dagger$, respectively. These mode operators satisfy the bosonic commutation relations $[\hat{a}_k, \hat{a}_l^\dagger] = \delta_{kl}$. The mode operators can be combined into the quadrature operators $\hat{q}_k = (\hat{a}_k + \hat{a}_k^\dagger)/\sqrt{2}$ and $\hat{p}_k = i(\hat{a}_k^\dagger - \hat{a}_k)/\sqrt{2}$. These operators correspond to the generalized position and momentum observables for the mode k. They have continuous spectra, and eigenbases $\{|q\rangle\}_{q\in\mathbb{R}}$ and $\{|p\rangle\}_{p\in\mathbb{R}}$, respectively. In the simplec-

tic form [56], the quadrature operators are collected in one single vector $\hat{\mathbf{x}} = (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_N, \hat{p}_N)^T$.

The state of such an N-mode system is described by a density operator $\rho \in \mathcal{D}(\mathcal{H}^{\otimes N})$, a positive (semi-definite) and unit trace operator. Alternatively, the state of the system can be represented by its Wigner function $W(\mathbf{x})$ [57], i.e., a quasiprobability distribution in the $2N$-dimensional phase space with real coordinates $q_i, p_i \in \mathbb{R}$, collected in a vector $\mathbf{x} = (q_1, p_1, \ldots, q_N, p_N)^T$.

II.B.1. Gaussian states

In the cases where the Wigner function of the state is a multivariate Gaussian distribution of the form

$$W(\mathbf{x}) = \frac{\exp[-(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{\Gamma}^{-1} (\mathbf{x}-\bar{\mathbf{x}})]}{\pi^N \sqrt{\det(\mathbf{\Gamma})}}, \tag{10}$$

the states are called *Gaussian*. Gaussian states are fully characterized by its vector of first moments $\bar{\mathbf{x}} = \mathrm{Tr}(\hat{\mathbf{x}}\rho)$ and its covariance matrix $\sigma = (\sigma_{ij}) = \frac{1}{2}\mathbf{\Gamma}$. The real and symmetric $2N \times 2N$ covariance matrix collects the second moments $\sigma_{ij} = \langle\{\hat{\mathbf{x}}_i - \langle\hat{\mathbf{x}}_i\rangle, \hat{\mathbf{x}}_j - \langle\hat{\mathbf{x}}_j\rangle\}\rangle/2$. Examples for Gaussian states include the vacuum state, thermal states as used, e.g., to describe black-body radiation, or coherent states modelling the photon distribution in a laser. The full description via the vector of first moments and the covariance matrix allows one to completely and compactly capture an important class of familiar states in an infinite-dimensional Hilbert space via a finite number of degrees of freedom.

In this paper we investigate the performance of single-mode Gaussian states for Bayesian parameter estimation. More specifically, we consider coherent and displaced-squeezed states. Coherent states are the right-eigenstates of the annihilation operator $\hat{a}_k$ such that $\hat{a}_k |\alpha\rangle_k = \alpha |\alpha\rangle_k$ and form a basis in the Hilbert space $\mathcal{H}_k$. They result from applying the displacement operator of the coherent amplitude $\alpha \in \mathbb{C}$,

$$\hat{D}_k(\alpha) = \exp\left(\alpha \hat{a}_k^\dagger - \alpha^* \hat{a}_k\right), \tag{11}$$

to the vacuum $|0\rangle_k$, such that $|\alpha\rangle_k = \hat{D}_k(\alpha)|0\rangle_k$. Coherent states are states with the same covariance matrix as the vacuum state. For a single-mode coherent state $|\alpha\rangle_k$, the first moment is $\bar{\mathbf{x}} = \sqrt{2}\big[\Re(\alpha), \Im(\alpha)\big]^T$ and the second moment is the identity matrix divided by 2, meaning that the variance both in $\hat{q}_k$ and $\hat{p}_k$ equals 1/2, saturating the uncertainty relation in a balanced way.

Coherent states are not the only states saturating the uncertainty relation. Indeed, squeezed states are a larger class of states with this property, while allowing for unbalanced variances of the two canonical quadratures for each mode, c.f. Fig. 2. Squeezed states are obtained by the action of the squeezing operator,

$$\hat{S}_k(\xi) = \exp\left[\frac{1}{2}(\xi^* \hat{a}_k^2 - \xi \hat{a}_k^{\dagger 2})\right], \tag{12}$$

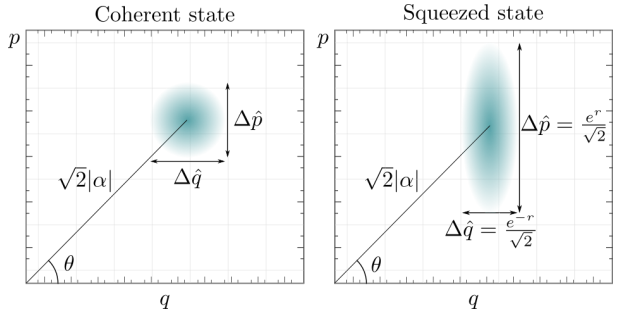

FIG. 2. Contours of the Wigner functions for single-mode Gaussian states. The Wigner functions are given by Gaussian distributions of the form Eq. (10), and are characterised by a complex displacement α, a real squeezing strength r and a squeezing angle φ. The illustration compares a displaced vacuum state ($r = 0$) on the left-hand side and a squeezed displaced state with $r > 0$ and $\varphi = 0$ on the right-hand side. The width of the latter Wigner function is reduced in the $\hat{q}$-quadrature and increased in the $\hat{p}$-quadrature with respect to the coherent state.

on the vacuum $|0\rangle_k$. The states $\hat{S}_k(\xi)|0\rangle_k$ are characterized by a complex parameter $\xi = re^{i\varphi}$, where $r \in \mathbb{R}$ is the so-called squeezing strength, and $\varphi \in [0, 2\pi)$ is the squeezing angle.

Every pure single-mode Gaussian state has minimal uncertainty and can be generated by the combined action of squeezing and displacement operators on the vacuum state. Such states are therefore entirely specified by their displacement parameter $\alpha \in \mathbb{C}$, their squeezing strength $r \in \mathbb{R}$, and their squeezing angle $\varphi \in [0, 2\pi)$. If squeezing is restricted to a real parameter only, then also a phase rotation

$$\hat{R}_k(\theta) = \exp\left(-i\theta \hat{a}_k^\dagger \hat{a}_k\right), \tag{13}$$

is needed to describe the most general pure single-mode Gaussian state. The vector of first moments of such a displaced squeezed state $|\alpha, re^{i\varphi}\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle = \hat{D}(\alpha)\hat{R}(\varphi/2)\hat{S}(r)|0\rangle$ is given by $\bar{\mathbf{x}} = \sqrt{2}\big[\Re(\alpha), \Im(\alpha)\big]^T$ and its covariance matrix is

$$\sigma = \frac{1}{2}\begin{pmatrix} \cosh 2r - \cos\varphi \, \sinh 2r & \sin\varphi \, \sinh 2r \\ \sin\varphi \, \sinh 2r & \cosh 2r + \cos\varphi \, \sinh 2r \end{pmatrix}. \tag{14}$$

A unitary transformation is called Gaussian, if it maps Gaussian states into Gaussian states. This class of unitary operations is generated by Hamiltonians that are (at most) second order polynomials of the mode operators. Notice that every single-mode Gaussian unitary operation can be decomposed into displacement, rotation, and squeezing operations. In addition to having a relatively straightforward theoretical description, Gaussian states and Gaussian transformations are also especially relevant in practice, since they are typically easy to produce and manipulate experimentally [34, 35].

II.B.2. Gaussian measurements

Any measurement can be described by a positive-operator valued measure (POVM). In CV quantum information, it is common to use continuous POVMs, that is, POVMs that are continuous sets of operators and a continuous range of measurement outcomes. A measurement is called Gaussian if it gives a Gaussian distribution of outcomes whenever it is applied to a Gaussian state. Gaussian measurements that are frequently considered in the context of CV quantum information are homodyne [58, 59] and heterodyne detection [60]. Homodyne detection corresponds to the measurement of a mode quadrature, for example $\hat{q}$. In this case, the POVM consists of projectors onto the quadrature basis, $\{|q\rangle\langle q|\}_{q\in\mathbb{R}}$. For heterodyne detection the POVM elements are projectors onto coherent states $\{\frac{1}{\pi}|\beta\rangle\langle\beta|\}_{\beta\in\mathbb{C}}$. Moreover, we note that it has recently been shown that every bosonic Gaussian observable can be considered as a combination of (noiseless and noisy) homodyne and heterodyne detection [61].

III. DISPLACEMENT ESTIMATION

We now consider Bayesian estimation of displacements using Gaussian states and Gaussian measurements. That is, we assume a displacement operator $\hat{D}(\alpha)$ as in Eq. (11) acts on our system, initially prepared in a Gaussian probe state. We then want to estimate the unknown displacement parameter $\alpha = \alpha_\mathrm{R} + i\alpha_\mathrm{I}$, with $\alpha_\mathrm{R}, \alpha_\mathrm{I} \in \mathbb{R}$. To this end, we focus on estimation strategies based on heterodyne and homodyne detection. These measurements are covariant under the action of displacement in the sense that the probability distribution obtained by displacing the probe state gives the same probability distribution translated by the displacement parameter in the parameter space [62]. Without loss of generality, we can therefore assume that the initial probe state has not been displaced from the origin, i.e., that our probe state is a squeezed vacuum state $|\xi\rangle = \hat{S}(\xi)|0\rangle$ with $\hat{S}(\xi)$ defined in Eq. (12). We further assume that our prior knowledge of the displacement is encoded in a Gaussian distribution of width σ_0 that is centered around α_0, i.e.,

$$p(\alpha) = \frac{1}{2\pi\sigma_0^2}\exp\left(-\frac{|\alpha-\alpha_0|^2}{2\sigma_0^2}\right). \tag{15}$$

Our goal is then to examine the performance of the estimation strategies based on heterodyne and homodyne detection, including the respective asymptotic behaviour, both in the limit of high photon numbers and of repeated measurements, and compare the respective results.

III.A. Heterodyne measurement

Let us first consider heterodyne detection, where the measurement is described by the POVM $\{\frac{1}{\pi}|\beta\rangle\langle\beta|\}_{\beta\in\mathbb{C}}$. The probability to obtain the measurement outcome β, given a displacement of α, is

$$p(\beta|\alpha) = \tfrac{1}{\pi}\mathrm{Tr}\left[|\beta\rangle\langle\beta|\hat{D}(\alpha)|\xi\rangle\langle\xi|\hat{D}^\dagger(\alpha)\right] = \tfrac{1}{\pi}\mathcal{F}\big(|\beta-\alpha\rangle,|\xi\rangle\big). \tag{16}$$

Here, $\mathcal{F}\big(\rho_1,\rho_2\big)$ is the Uhlmann fidelity of the states ρ_1 and ρ_2 (defined in Sec. II.A.2), which reduces to $\mathcal{F}\big(|\psi\rangle,|\phi\rangle\big) = |\langle\psi|\phi\rangle|^2$ for pure states. For two Gaussian states, the fidelity can be written in terms of the respective first moments $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$, and second moments $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ (cf. [7]) as

$$\mathcal{F}(\rho_1,\rho_2) = \frac{2\exp[-(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)^T(\mathbf{\Gamma}_1+\mathbf{\Gamma}_2)^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)]}{\sqrt{|\mathbf{\Gamma}_1+\mathbf{\Gamma}_2|+(1-|\mathbf{\Gamma}_1|)(1-|\mathbf{\Gamma}_2|)}-\sqrt{(1-|\mathbf{\Gamma}_1|)(1-|\mathbf{\Gamma}_2|)}}. \tag{17}$$

For simplicity we now assume that our probe state is squeezed only along one fixed direction, i.e., $\varphi = 0$. This simplifies the following calculation considerably. In particular, this allows us to write the fidelity, the likelihood, and posterior distribution as products of the corresponding distributions for the real and imaginary part of the displacements, respectively. In contrast, for the general case of probe states squeezed along arbitrary directions, the resulting formulas are unwieldy and complicated, but qualitatively yield the same behaviour as for $\varphi = 0$. We therefore refrain from presenting these calculations here.

In our case, we have $\rho_1 = |\beta-\alpha\rangle\langle\beta-\alpha|$ and $\rho_2 = |\xi\rangle\langle\xi|$, for which the first moments are

$$\bar{\mathbf{x}}_{\beta-\alpha} = \sqrt{2}\begin{pmatrix}\Re[\beta-\alpha]\\ \Im[\beta-\alpha]\end{pmatrix} = \sqrt{2}\begin{pmatrix}\beta_\mathrm{R}-\alpha_\mathrm{R}\\ \beta_\mathrm{I}-\alpha_\mathrm{I}\end{pmatrix} \quad\text{and}\quad \bar{\mathbf{x}}_\xi = \begin{pmatrix}0\\0\end{pmatrix},$$

while the second moments are represented by

$$\mathbf{\Gamma}_{\beta-\alpha} = \mathbb{1}_2 \quad\text{and}\quad \mathbf{\Gamma}_\xi = \begin{pmatrix}\mathrm{e}^{-2r} & 0\\ 0 & \mathrm{e}^{2r}\end{pmatrix},$$

respectively. Accordingly, $p(\beta|\alpha)$ from Eq. (16) becomes

$$\begin{aligned} p(\beta|\alpha) &= \frac{\exp\left[-\frac{\mathrm{e}^r(\beta_\mathrm{R}-\alpha_\mathrm{R})^2+\mathrm{e}^{-r}(\beta_\mathrm{I}-\alpha_\mathrm{I})^2}{\cosh r}\right]}{\pi\cosh r}\\ &= p(\beta_\mathrm{R}|\alpha_\mathrm{R})\,p(\beta_\mathrm{I}|\alpha_\mathrm{I}), \end{aligned} \tag{18}$$

where the distributions $p(\beta_i|\alpha_i)$ for $i = \mathrm{R}, \mathrm{I}$ are given by

$$p(\beta_i|\alpha_i) = \frac{\sqrt{2}\exp\left[-\frac{2(\beta_i-\alpha_i)^2}{1+\mathrm{e}^{\mp 2r}}\right]}{\sqrt{\pi(1+\mathrm{e}^{\mp 2r})}}. \tag{19}$$

Here and in the following equations, the upper and lower signs in $\pm$ and $\mp$ correspond to the subscripts $i = \mathrm{R}$ and $i = \mathrm{I}$, respectively, i.e., for $i = \mathrm{R}$, the respective upper signs apply, while the lower signs apply for $i = \mathrm{I}$. With this expression for the likelihood and with the prior from Eq. (15), one can use Bayes' law [Eq. (2)] to calculate the posterior distribution, the estimators and the (average) variance. This allows one to evaluate the average variance for different estimation scenarios. We rely on such

an approach in the next sections. However, in the special case where both prior and likelihood are Gaussian, these two quantities are conjugate to each other. Following Prop. 1, the posterior is therefore also Gaussian, and we can write down the mean and variance of the posterior directly by inspecting the likelihood and the prior. That is, by noting that $\sigma^2 = (1 + e^{\mp 2r})/4$, $\mu_0 = \alpha_{0,i}$, and $\bar{\theta}(m) = \beta_i$, Prop. 1 provides the mean and variance of the distributions $p(\alpha_i|\beta_i)$. Again using subscripts $i = \mathrm{R}, \mathrm{I}$ to denote real and imaginary parts, respectively, the means are

$$\hat{\alpha}_i(\beta_i) = \frac{4\beta_i\sigma_0^2 + \alpha_{0,i}(1 + e^{\mp 2r})}{4\sigma_0^2 + 1 + e^{\mp 2r}}, \tag{20}$$

which we choose as estimators for the real and imaginary part of the parameter α, and the variances are

$$\mathrm{Var}[p(\alpha_i|\beta_i)] = \left[\frac{1}{\sigma_0^2} + 2(1 \pm \tanh r)\right]^{-1}. \tag{21}$$

We then define the total variance of the posterior $p(\alpha|\beta)$ for the complex parameter α as

$$\mathrm{Var}[p(\alpha|\beta)] = \int d\alpha\; p(\alpha|\beta)\,|\alpha - \hat{\alpha}(\beta)|^2. \tag{22}$$

Because the real and imaginary parts become independent, we can further write the total variance as the sum of the variances of the two independent estimation parameters, i.e.,

$$\mathrm{Var}[p(\alpha|\beta)] = \mathrm{Var}[p(\alpha_\mathrm{R}|\beta_\mathrm{R})] + \mathrm{Var}[p(\alpha_\mathrm{I}|\beta_\mathrm{I})]. \tag{23}$$

After inserting Eq. (21) twice, the latter expression is independent of β and therefore it already represents the average total variance $\bar{V}_\mathrm{post}$ we are interested in determining.

Moreover, it depends only on the variance σ_0^2 of the prior and the squeezing strength r of the probe state. For a fixed prior, the average posterior variance of both coordinates from Eq. (23) is minimized for $r = 0$, that is, when there is no squeezing of the probe state. We thus have

$$\bar{V}_\mathrm{post}(r) \geq \bar{V}_\mathrm{post}(r = 0) = \frac{2\sigma_0^2}{1 + 2\sigma_0^2}. \tag{24}$$

However, squeezing can help to reduce the variance in one coordinate, but this reduction comes at the cost of increasing the variance of the other coordinate with respect to the case where $r = 0$. Irrespective of the squeezing strength, we observe that the variances for both phase space coordinates decrease with respect to the prior, but only slightly. When one is interested in reducing the variance in only one of the coordinates, say α_R, one may note that the variance decreases monotonically for increasing r. Nevertheless, even as $r \to \infty$ the variance of the posterior is still bounded from below by $(\sigma_0^{-2} + 4)^{-1}$. This residual variance originates in the intrinsic uncertainty of the coherent-state basis associated with the POVM representing heterodyne detection. That is, no matter which measurement outcome is obtained, the precision with which the parameter is identified is limited by the width of the variance of the coherent state corresponding to this outcome.

Although coherent states already minimize the product of uncertainties, one can overcome this limitation by considering measurement bases that consist of states with a lower variance in the desired parameter (e.g., in α_R) than that of a coherent state, at the expense of a larger variance in the respective other quadrature. For instance, one may choose a basis of squeezed coherent states to reduce the uncertainty of the measurement basis in one coordinate. In this regard, a homodyne measurement in the quadrature $\hat{q}$, which we will consider next, can be thought of as a limiting case of a measurement in a basis of infinitely squeezed coherent states.

III.B. Homodyne measurement

For homodyne detection with respect to the quadrature $\hat{q}$, the POVM is $\{|\,q\,\rangle\!\langle\,q\,|\}_{q\in\mathbb{R}}$. As before, we begin by considering a squeezed vacuum state $|\,\xi\,\rangle$ as probe state to estimate the unknown displacement α. The prior distribution of α is again assumed to be Gaussian with mean α_0 and variance σ_0^2. The probability to obtain outcome q after a displacement α is given by

$$p(q|\alpha) = |\langle q|\hat{D}(\alpha)|\xi\rangle|^2 = \frac{\exp\left[-\frac{2(\alpha_\mathrm{R} - \frac{q}{\sqrt{2}})^2}{\cosh 2r - \cos\varphi \sinh 2r}\right]}{\sqrt{\pi(\cosh 2r - \cos\varphi \sinh 2r)}}. \tag{25}$$

Note that, here, the likelihood does not depend on the imaginary part α_I of the displacement. This is expected, since homodyne detection in one quadrature is completely 'blind' to the orthogonal quadrature. Therefore, the mean and variance for the imaginary part of the displacement parameter remain unchanged with respect to the prior, and we can focus entirely on the real part.

Since, once again the likelihood is a Gaussian distribution in the measurement outcomes (here, in q), and thus proportional to a Gaussian distribution $\mathcal{N}_{\alpha_\mathrm{R}}(\langle\alpha_\mathrm{R}\rangle, \sigma^2)$ in the estimated parameter with mean $\langle\alpha_\mathrm{R}\rangle = q/\sqrt{2}$ and variance $\sigma^2 = (\cosh 2r - \cos\varphi \sinh 2r)/4$, we can infer from Prop. 1 that the posterior is a Gaussian distribution with mean

$$\hat{\alpha}_\mathrm{R} = \frac{2\sqrt{2}\,\sigma_0^2 q + \alpha_{0,\mathrm{R}}(\cosh 2r - \cos\varphi \sinh 2r)}{4\sigma_0^2 + \cosh 2r - \cos\varphi \sinh 2r}, \tag{26}$$

and variance

$$\mathrm{Var}[p(\alpha_\mathrm{R}|q)] = \frac{\sigma_0^2(\cosh 2r - \cos\varphi \sinh 2r)}{4\sigma_0^2 + \cosh 2r - \cos\varphi \sinh 2r}. \tag{27}$$

The variance of the posterior distribution depends on the squeezing strength r and the squeezing angle φ. Both

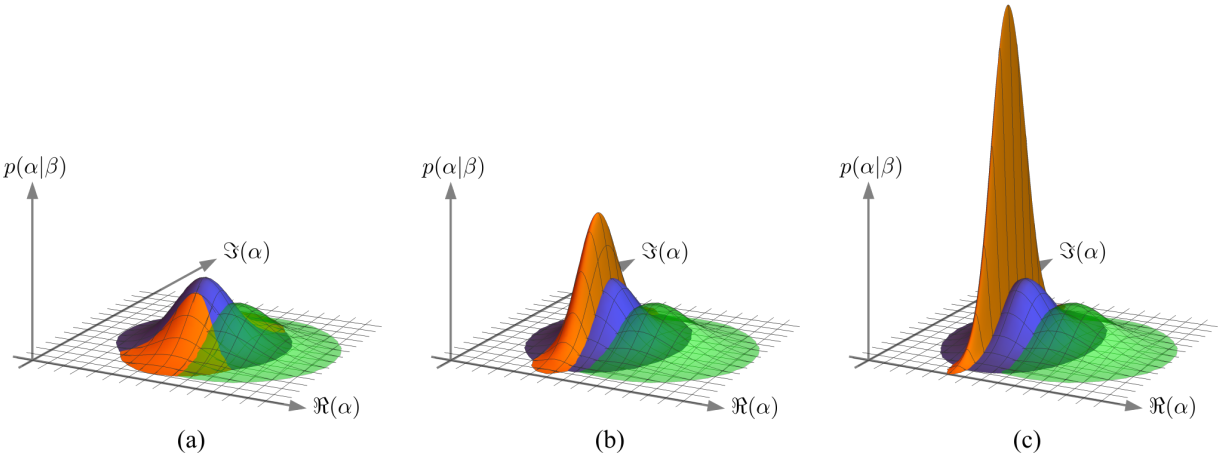

FIG. 3. Displacement estimation using heterodyne and homodyne detection. The images show the same Gaussian prior (green) with initial standard deviation $\sigma_0 = 0.5$, and posterior distributions obtained for heterodyne (blue) and homodyne detection (orange) for different squeezing of the probe state, ranging from $r = 0$ in (a), $r = 1$ in (b), to $r = 2$ in (c). The posterior distributions of the displacement parameter α given measurement outcome q are Gaussian as well.

parameter hence provide room for optimization of the estimation procedure. However, while increasing r can be demanding experimentally and also comes at an increased energy cost for preparing the probe state, the relative angle φ between the directions of measurement and squeezing can be varied freely without any particular practical or energetic restriction. The variance is minimised for $\varphi = 2n\pi$ and without loss of generality we choose $\varphi = 0$. For this choice, the average variance of the posterior for the chosen quadrature $\hat{q}$ is

$$\bar{V}^{\hat{q}}_{\text{post}} = \text{Var}[p(\alpha_{\text{R}}|q)] \overset{\varphi=0}{=} \left(\frac{1}{\sigma_0^2} + 4\text{e}^{2r}\right)^{-1}, \qquad (28)$$

whereas the average total variance (again, for $\varphi = 0$) is $\bar{V}_{\text{post}} = \bar{V}^{\hat{q}}_{\text{post}} + \sigma_0^2$. Fig. 3 shows a sample of different posterior distributions obtained by measurements with probe states with different squeezing. We observe that, whereas the marginal probability in $\hat{p}$ remains unchanged as the initial squeezing increases, the marginal probability in $\hat{q}$ becomes narrower. We further note that for $r = 0$ we recover the results obtained by Personick [63].

III.C. Comparison of measurement strategies

Let us now interpret and compare the results for Gaussian displacement estimation with heterodyne and homodyne measurements. For homodyne detection, squeezing in the probe state results in an average posterior variance in $\hat{q}$, given by Eq. (28), that rapidly decreases to 0 as the squeezing strength r increases. While the posterior variance in $\hat{q}$ can thus be arbitrarily close to zero in the homodyne detection scenario, this comes at the cost of not reducing the variance in $\hat{p}$ at all. We thus have $\lim_{r\to\infty} \bar{V}^{\text{homodyne}}_{\text{post}} = \sigma_0^2$. Comparing this with the result for heterodyne detection in Eq. (24), we see that $\bar{V}^{\text{homodyne}}_{\text{post}} \geq \bar{V}^{\text{heterodyne}}_{\text{post}}(r = 0)$ for priors with variance $\sigma_0^2 \geq 1/2$, independently of the squeezing strength used with the homodyne detection. However, for more narrow priors, homodyne detection supplemented by squeezed probe states can outperform heterodyne detection in terms of the total variance only if the squeezing is strong enough, i.e., when $r > -\frac{1}{2}\ln(1 - 2\sigma_0^2)$.

However, when we focus on the estimation of only one of the quadratures, here quadrature $\hat{q}$, then homodyne detection outperforms heterodyne detection for all prior widths and for all squeezing strengths, even if different squeezing strengths are compared for the two detection methods. That is, the limit of $r \to \infty$ for heterodyne detection in Eq. (21) coincides with the homodyne detection case where $r = 0$ in Eq. (28), and we thus find

$$\bar{V}^{\hat{q},\,\text{homodyne}}_{\text{post}} \leq \frac{\sigma_0^2}{1 + 4\sigma_0^2} \leq \bar{V}^{\hat{q},\,\text{heterodyne}}_{\text{post}}. \qquad (29)$$

We can also compare these results to more general measurement strategies. For a Gaussian prior (in a single parameter), the Fisher information of the prior (see Sec. II.A.2) evaluates to $I[p(\alpha_{\text{R}})] = 1/\sigma_0^2$. At the same time, the QFI for a single-mode Gaussian state is bounded by $\mathcal{I}(\rho) \leq 4\text{e}^{2r}$ (cf. Eq. (15) and subsequent text in Ref. [7]). With this, the Van Trees inequality in the form of Eq. (9) reads

$$\bar{V}^{\hat{q}}_{\text{post}} \geq \left(\frac{1}{\sigma_0^2} + 4\text{e}^{2r}\right)^{-1}. \qquad (30)$$

This shows that the combination of single-mode squeezing and homodyne detection is the optimal strategy for Bayesian estimation of one coordinate of displacement

(or displacement radius with known phase) with a single-mode Gaussian probe state.

Finally, let us consider repeated measurements, which can easily be accommodated within the framework of conjugate priors. In particular, we know that the posterior is of the same form as the prior, i.e., both are normal distributions. Since the posterior distribution is used as the prior for the next measurement round, we obtain a recursive formula for the average variance, given by

$$\sigma_{m+1}^2 = \frac{\sigma_m^2 \operatorname{Var}[p(q|\alpha)]}{\sigma_m^2 + \operatorname{Var}[p(q|\alpha)]}, \tag{31}$$

where σ_m is the variance of round m. Since $\operatorname{Var}[p(q|\alpha)] = \mathrm{e}^{-2r}/4$ depends only on the squeezing of the probe state, this term is constant for the same probe state. Solving the recursive equation gives

$$\sigma_m^2 = \left(\frac{1}{\sigma_0^2} + 4m\,\mathrm{e}^{2r}\right)^{-1}. \tag{32}$$

Moreover, we note that repeated measurements include the possibility of a sequential measurement strategy that provides information about both components of the displacement. For instance, the squeezing in the probe states and the direction of the homodyne measurement can be tailored towards estimating the real part in one half of the estimation rounds, while the remaining rounds are used to estimate the imaginary part. We conclude this section by noting that already a quite simple setup, consisting of (limited) squeezing in the probe states combined with homodyne detection, can provide accurate information for Bayesian estimation of displacements.

IV. PHASE ESTIMATION

We now come to the paradigmatic case of phase estimation, which we want to examine within the framework of Bayesian estimation using Gaussian states and measurements. Historically, phase estimation has been closely associated with interferometry [64], but nowadays, phase estimation is usually considered in a broader context. In particular, Bayesian phase estimation has been studied for a variety of applications, see, e.g., [65–67]. While there are some studies identifying optimal estimation strategies using Gaussian states and measurements [15, 68, 69], these operate within the local estimation paradigm and hence fall outside of the Bayesian phase estimation framework we consider here. We therefore focus on a special case of Bayesian phase estimation, where there is no prior information on the phase and local estimation hence cannot be employed in a meaningful way. For such cases, we wish to identify simple strategies based on Gaussian states and measurements that can efficiently narrow the prior down to the point where local estimation can take over.

Specifically, we consider a phase estimation scenario where a phase rotation operator as in Eq. (13) is applied to a single-mode Gaussian probe state. We consider the phase $\theta \in [-\pi, \pi)$ to be entirely unknown initially, such that the prior is a uniform distribution on the chosen interval, i.e., $p(\theta) = 1/2\pi$.

In the following sections, we then study the performance of heterodyne and homodyne detection in this estimation scenario, and we adapt the specific probe states to the respective measurements. In particular, we note that, although the optimal probe state (at fixed average energy) for local phase estimation is a single-mode squeezed state, this is not necessarily the case for Bayesian estimation.

IV.A. Heterodyne measurement

For Gaussian phase estimation with heterodyne measurements, we consider probe states that are squeezed with strength $r = |\xi|$ before being displaced, i.e., probe states of the form $\hat{D}(\alpha)\hat{S}(r\mathrm{e}^{i\varphi})\,|0\rangle$, where $r \geq 0$ and $\varphi \in [0, 2\pi)$. Whereas the most general Gaussian single-mode probe states are determined by arbitrary complex values α and ξ, i.e., displacement and squeezing with arbitrary strength along arbitrary directions, the rotational symmetry of the phase estimation problem with heterodyne measurements allows one to fix one of these directions. Without loss of generality, we therefore choose $\alpha = |\alpha|$ to be real and positive. More specifically, we assume that the displacement is strictly non-zero, $\alpha > 0$, since the vacuum state is rotationally invariant, and not even a squeezed vacuum state can be used to distinguish between rotations around θ and $\theta + \pi$.

For the squeezing direction, it is then quite intuitive to see that squeezing along the quadrature $\hat{p}$ ($\varphi = \pi$, $\xi = -r < 0$) is optimal for single-mode phase estimation when $\alpha > 0$ and when heterodyne measurements are used. That is, when the variance of the Gaussian state is initially reduced along the quadrature $\hat{p}$, the Wigner function becomes concentrated along the $\hat{q}$-quadrature, decreasing the variance in the phase of the initial state, and hence also decreasing the variance in the phase of the encoded state $\rho(\theta)$. When applying the heterodyne measurement, the probability for obtaining an outcome β whose phase matches the unknown phase θ is thus increased. Conversely, probe states that are squeezed along the same direction as the initial displacement have an increased phase variance and are therefore less useful for phase estimation. In the remainder of this section, we therefore focus on probe states of the form $\hat{D}(\alpha)\hat{S}(-r)\,|0\rangle$.

However, since the calculations and results for arbitrary values of r are still quite unwieldy, we first consider the simple case where the probe state is not squeezed at all but just a coherent state $|\alpha\rangle$ (Sec. IV.A.1). Then we present the results for squeezing along the optimal direction, $\xi = -r < 0$, with respect to the displacement $\alpha > 0$ (Sec. IV.A.2).

IV.A.1. Coherent states & heterodyne detection

Here, the probe state is $|\alpha\rangle$ with $\alpha > 0$. The action of the phase rotation operator $\hat{R}(\theta)$ [Eq. (13)] results in the encoded state $\hat{R}(\theta)|\alpha\rangle = |e^{-i\theta}\alpha\rangle$. The likelihood to obtain outcome $\beta \in \mathbb{C}$, given that the phase has the value θ, is given by

$$p(\beta|\theta) = \tfrac{1}{\pi}|\langle\beta|e^{-i\theta}\alpha\rangle|^2 = \tfrac{1}{\pi}e^{-|e^{i\theta}\beta-\alpha|^2}. \tag{33}$$

Writing $\beta = |\beta|e^{-i\phi_\beta}$ and $|e^{i\theta}\beta-\alpha|^2 = \alpha^2+|\beta|^2-2\alpha|\beta|\cos(\theta-\phi_\beta)$, we can express the (unconditional) probability to obtain outcome β as

$$p(\beta) = \int_{-\pi}^{\pi} d\theta\, p(\theta)\, p(\beta|\theta) = \frac{e^{-(\alpha^2+|\beta|^2)}}{\pi} I_0(2\alpha|\beta|), \tag{34}$$

where $I_0(x)$ is the modified Bessel function of the first kind. Using Bayes' law, the posterior is given by

$$p(\theta|\beta) = \frac{p(\theta)\,p(\beta|\theta)}{p(\beta)} = \frac{e^{2\alpha|\beta|\cos(\theta-\phi_\beta)}}{2\pi\, I_0(2\alpha|\beta|)}. \tag{35}$$

Since we are considering a parameter with a range whose endpoints $\pm\pi$ are identified, it is useful to consider estimators and variances that are invariant under shifts by 2π. For the estimator we therefore choose $\hat{\theta}(\beta) = \arg\langle e^{i\theta}\rangle_{p(\theta|\beta)}$. As we discuss in more detail in Appendix A.1, the estimator evaluates to

$$\hat{\theta}(\beta) = \arg\Big[\int_{-\pi}^{\pi} d\theta\, p(\theta|\beta) e^{i\theta}\Big] = \phi_\beta, \tag{36}$$

and hence corresponds to the phase ϕ_β of the measurement outcome β.

To evaluate the performance of this estimation strategy, we calculate the average variance of the posterior as done in the above sections. However, instead of an expression such as in Eq. (4), we now use a covariant variance that is invariant under shifts by 2π, by taking the average of $\sin^2\big[\theta-\hat{\theta}(\beta)\big]$ rather than of $\big(\theta-\hat{\theta}(\beta)\big)^2$.[1] Specifically, we calculate

$$V_{\text{post}}(\beta) = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\, \sin^2\big[\theta-\hat{\theta}(\beta)\big] = \frac{{}_0F_1(2;\alpha^2|\beta|^2)}{2\,I_0(2\alpha|\beta|)\,\Gamma(2)}, \tag{37}$$

where ${}_0F_1(a;z)$ is the confluent hypergeometric function and $\Gamma(z)$ is the Euler gamma function. Despite the complicated form of the posterior and the variance, the average variance then simply becomes

$$\bar{V}_{\text{post}} = \int d^2\beta\; p(\beta)\, V_{\text{post}}(\beta) = \frac{1-e^{-|\alpha|^2}}{2|\alpha|^2}, \tag{38}$$

as we discuss in more detail in Appendix A.1. In terms of the average photon number $n = |\alpha|^2$, which is proportional to the average energy of the probe state, the average variance of the posterior hence scales as $1/n$ as $n \to \infty$, as can be expected for 'classical' probe states such as the coherent states considered here.

IV.A.2. Displaced squeezed states & heterodyne detection

Let us now consider probe states that are squeezed with strength r before being displaced, i.e., probe states of the form $\hat{D}(\alpha)\hat{S}(-r)|0\rangle$, where we assume $\alpha, r \in \mathbb{R}$ with $\alpha > 0$ and $r > 0$ as mentioned. For the heterodyne measurement, the likelihood to obtain outcome β given the phase θ is given by

$$\begin{aligned} p(\beta|\theta) &= \tfrac{1}{\pi}|\langle\beta|\hat{R}(\theta)\hat{D}(\alpha)\hat{S}(-r)|0\rangle|^2 \\ &= \tfrac{1}{\pi}\mathcal{F}\big(|e^{i\theta}\beta\rangle, |\alpha,-r\rangle\big). \end{aligned} \tag{39}$$

For the fidelity of the two Gaussian states, we can again refer to Eq. (17), where $\rho_1 = |e^{i\theta}\beta\rangle\langle e^{i\theta}\beta|$ and $\rho_2 = |\alpha,-r\rangle\langle\alpha,-r|$, for which the first moments are

$$\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_{e^{i\theta}\beta} = \sqrt{2}\begin{pmatrix}\Re(e^{i\theta}\beta)\\ \Im(e^{i\theta}\beta)\end{pmatrix} \text{ and } \bar{\mathbf{x}}_2 = \bar{\mathbf{x}}_{\alpha,-r} = \sqrt{2}\begin{pmatrix}\Re(\alpha)\\ \Im(\alpha)\end{pmatrix}.$$

The second moments of these states are represented by

$$\mathbf{\Gamma}_1 = \mathbf{\Gamma}_{e^{i\theta}\beta} = \mathbb{1}_2 \quad \text{and} \quad \mathbf{\Gamma}_2 = \mathbf{\Gamma}_{\alpha,-r} = \begin{pmatrix}e^{2r} & 0\\ 0 & e^{-2r}\end{pmatrix},$$

respectively. Since $\det\Gamma_1 = \det\Gamma_2 = 1$ and $\det(\Gamma_1+\Gamma_2) = 4\cosh^2(r)$, we then have

$$p(\beta|\theta) = \frac{\exp\Big[-\frac{e^{-r}\Re^2(e^{i\theta}\beta-\alpha)+e^{r}\Im^2(e^{i\theta}\beta-\alpha)}{\cosh r}\Big]}{\pi\cosh r}. \tag{40}$$

As we explain in more detail in Appendix A.1.II, the (unconditional) probability to obtain outcome β can then be written as an infinite sum of Bessel functions of the first kind by using the Jacobi-Anger expansion, which results in

$$\begin{aligned} p(\beta) = \frac{e^{-\alpha^2(1-\tanh r)-|\beta|^2}}{\pi\cosh r} \sum_{\substack{m_1,m_2\\=-\infty}}^{\infty} e^{-im_1\pi}\, I_{-2m_1-m_2}(-2\alpha|\beta|) \\ \times\, I_{m_1}(-|\beta|^2\tanh r)\, I_{m_2}(2\alpha|\beta|\tanh r). \end{aligned} \tag{41}$$

[1] We note here that the chosen variance is invariant also under shift of the estimator by integer multiples of π, not just shift by even multiples of π. In principle, one could also use quantifiers for the width of the distribution that depend only on $|\langle e^{i\theta}\rangle_{p(\theta|\beta)}|$, such as the Holevo phase variance [70], which are completely independent of the value of the estimator. The choice we make here is motivated by the better comparison with the homodyne detection scenario in Sec. IV.B, where the phase can only be resolved within an interval of length π.

Using Bayes' law, the posterior can then be obtained directly as $p(\theta|\beta) = p(\beta|\theta)/\big[2\pi\, p(\beta)\big]$ with the likelihood from Eq. (40) and $p(\beta)$ as in Eq. (41). Similarly, we can use the Jacobi-Anger expansion to evaluate $\langle e^{i\theta}\rangle = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta) e^{i\theta}$. As shown explicitly in Appendix A.1.II, one finds $\Im\big(\langle e^{i(\theta-\phi_\beta)}\rangle\big) = 0$, and the estimator is hence given by

$$\hat{\theta}(\beta) = \phi_\beta \text{ or } \phi_\beta + \pi, \tag{42}$$

i.e., the estimate either corresponds to the phase ϕ_β of the measurement outcome β, or is shifted by π.

To see if squeezing improves the estimation, we calculate the variance of the posterior, $V_{\text{post}}(\beta) = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\, \sin^2[\theta - \hat{\theta}(\beta)]$, and its average, and compare the latter with the corresponding value obtained for coherent probe states. Specifically, we obtain the expression (see Appendix A.1.II for more details)

$$\begin{aligned}\bar{V}_{\text{post}} = \frac{e^{-\alpha^2(1-\tanh r)}}{\cosh r} \sum_{\substack{n_2,n_3\\=-\infty}}^{\infty} \int_0^{\infty} d|\beta|\, |\beta|\, e^{-|\beta|^2}\, I_{n_2}(-|\beta|^2 \tanh r) \\ \times I_{n_3}(2\alpha|\beta|\tanh r)\tfrac{1}{2}(-1)^{n_2}\Big[2I_{-2n_2-n_3}(-2\alpha\,|\beta|) \\ - I_{2-2n_2-n_3}(-2\alpha\,|\beta|) - I_{-2-2n_2-n_3}(-2\alpha\,|\beta|)\Big].\end{aligned} \tag{43}$$

Unfortunately, the analytical solution of the integral and double-sum in Eq. (43) is unknown. We have therefore numerically evaluated the average variance $\bar{V}_{\text{post}}$ for different values of α and r. As illustrated by the sample plots in Fig. 4 (a), for any fixed displacement, squeezing improves the estimation precision as measured by the average variance beyond the value achievable by displacements alone, where the latter is represented by Eq. (38). This is in agreement with the intuition provided by the Wigner function of the probe states: Squeezing along the $\hat{p}$-quadrature ($\xi = -r < 0$) of a coherent state displaced along the $\hat{q}$ axis ($\alpha > 0$) leads to a concentration of the Wigner function around the $\hat{q}$-axis, that is rotated around the origin by the phase rotation, visually resembling a clock dial. Increased squeezing narrows the width of this 'dial', making it more likely to obtain measurement outcomes β whose phase matches the phase to be estimated.

However, when considering constraints on the average energy of the probe state, here represented by the average photon number $n = |\alpha|^2 + \sinh^2 r$, squeezing is only beneficial in certain regimes. For relatively strong squeezing such as $r = 1$ or $r = 1.25$, the average variance is larger for squeezed-displaced states than for purely displaced states with the same average photon number, as illustrated in Fig. 4 (b). This can be understood from the fact that the average photon number required for a squeezing of $r = 1.25$ is sufficient for a coherent state that is displaced more than 2 standard deviations from the origin and hence already provides a clear phase reference. For smaller squeezing, such as for $r = 0.75$, there is a regime of small photon numbers where the combination of squeezing and displacement can outperform pure

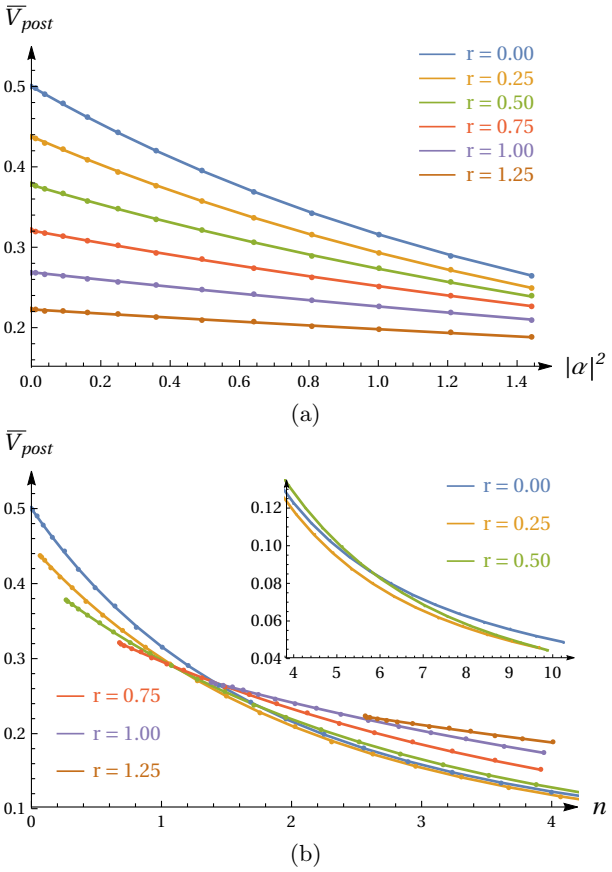

FIG. 4. Bayesian phase estimation with single-mode Gaussian probes and heterodyne measurements. (a) The average variance $\bar{V}_{\text{post}}$ from Eq. (43) is shown for different values of $\alpha \geq 0$ and $r \geq 0$ as a function of α^2. The line on the top represents the average variance for purely displaced probe states ($r = 0$) from Eq. (38). The lines below indicate results of numerically evaluating Eq. (43) for different values of α for fixed values of r from $r = 0.25$ to $r = 1.25$ (top to bottom, starting at the second line from the top). (b) The average variance $\bar{V}_{\text{post}}$ is shown as a function of the average photon number $n = |\alpha|^2 + \sinh^2 r$. The lines do not start at $n = 0$ because the nonzero values of r give rise to non-zero average energies even for $\alpha = 0$. The inset shows how the lines for $r = 0$, $r = 0.25$, and $r = 0.5$ continue as n increases.

displacement. This can also be readily understood, while such a squeezed vacuum state already has a standard deviation $\Delta\hat{p}(|\xi\rangle) = e^{-r}/\sqrt{2}$ less than half of that of a coherent state, a coherent state with the same average energy is displaced by only $\sqrt{2}\alpha = \sqrt{2\sinh^2 r} \approx 1.64\Delta\hat{q}(|\alpha\rangle)$. However, for larger n (already around $n \approx 1.41$) pure displacements become better, see Fig. 4 (b). Finally, we see that for even smaller values of r, such as for $r = 0.5$, there is only a specific range of values for n where purely coherent probes are more efficient, while low squeezing ($r = 0.25$) added to the displacement outperforms pure displacement for the entire range of n that we have explored numerically. At the same time, in terms of the

difference between the average variances achieved, e.g., for $r = 0$ and $r = 0.25$, the advantage obtained from using a slightly squeezed state seems to be at least an order of magnitude smaller than the average variances achieved (in the explored parameter range).

IV.B. Homodyne measurement

Here, we consider Bayesian phase estimation with single-mode Gaussian probe states combined with homodyne measurements in the quadrature $\hat{q}$. Since this kind of measurement provides no information on the complementary quadrature $\hat{p}$, it cannot distinguish between phases phases of θ and $-\theta$. Thus, we restrict the range of θ to $[0, \pi]$, and the prior distribution is given by $p(\theta) = 1/\pi$.

IV.B.1. Coherent states & homodyne detection

As before in Sec. IV.A, we start with the case where the probe state is a coherent state, $\hat{D}(\alpha)|0\rangle = |\alpha\rangle$ for $\alpha > 0$. The likelihood to obtain outcome $q \in \mathbb{R}$ can be written as

$$p(q|\theta) = |\langle q|e^{-i\theta}\alpha\rangle|^2 = \int_{-\infty}^{\infty} dp\, W(q,p)\,, \tag{44}$$

where $W(q,p)$ is the Wigner function of the rotated coherent state $|e^{-i\theta}\alpha\rangle$. The latter can be obtained from Eq. (10) by noting that $\mathbf{\Gamma}_{e^{-i\theta}\alpha} = \mathbb{1}_2$ and $\bar{\mathbf{x}}_{e^{-i\theta}\alpha} = \sqrt{2}\alpha(\cos\theta, -\sin\theta)^T$. With this, one finds that

$$p(q|\theta) = \tfrac{1}{\sqrt{\pi}} e^{-\left(q-\sqrt{2}\alpha\cos\theta\right)^2}. \tag{45}$$

Further noting that the range of θ is $[0, \pi]$, the (unconditional) probability to obtain q can be expressed as

$$\begin{aligned} p(q) &= \int_0^{\pi} d\theta\, p(\theta)\, p(q|\theta) \\ &= \tfrac{1}{\sqrt{\pi}} e^{-q^2-\alpha^2} \sum_{m=-\infty}^{\infty} I_{2m}(2\sqrt{2}q\alpha) I_m(-\alpha^2), \end{aligned} \tag{46}$$

as we show in detail in Appendix A.1.III. Using Bayes' law, the posterior $p(\theta|q) = p(q|\theta)/\left[\pi\, p(q)\right]$ is then just obtained by inserting $p(q|\theta)$ and $p(q)$ from Eqs. (44) and (46), respectively. In Appendix A.1.III we also explicitly calculate the circular moment, which we find to be given by

$$\begin{aligned} \langle e^{i\theta}\rangle &= \tfrac{1}{M} \sum_{n=-\infty}^{\infty} I_{2n+1}(2\sqrt{2}q\alpha) I_n(-\alpha^2) \\ &\quad + i\tfrac{2}{M\pi} \sum_{m,n=-\infty}^{\infty} \tfrac{I_m(-\alpha^2) I_{2n}(2\sqrt{2}q\alpha)(1-4m^2-4n^2)}{(2n-2m-1)(2n-2m+1)(2n+2m+1)(2n+2m-1)}, \end{aligned} \tag{47}$$

where

$$M := \sum_{m=-\infty}^{\infty} I_{2m}(2\sqrt{2}q\alpha) I_m(-\alpha^2)\,. \tag{48}$$

As we see, already the expression for the estimator $\hat{\theta}(q) = \arctan\left[\Im(\langle e^{i\theta}\rangle)/\Re(\langle e^{i\theta}\rangle)\right]$ for a coherent probe state is sufficiently more complicated than its counterpart in the case of heterodyne measurements [cf. Eq. (36)]. We therefore resort to a numerical evaluation of the variance $V_{\text{post}}(q) = \int_0^{\pi} d\theta\, p(\theta|q)\, \sin^2\left[\theta - \hat{\theta}(q)\right]$ and the average variance $\bar{V}_{\text{post}} = \int_{-\infty}^{\infty} dq\, p(q)\, V_{\text{post}}(q)$ already for the case of coherent probe states. The results for $\bar{V}_{\text{post}}$ as a function of $n = \alpha^2$ are shown in Fig. 5, together with the corresponding average variance for squeezed probe states, which we will briefly discuss next.

IV.B.2. Displaced squeezed states & homodyne detection

In the present section, we consider a squeezed and displaced probe state, $\hat{D}(\alpha)\hat{S}(re^{i\varphi})|0\rangle$ for $\alpha \geq 0$ and $\varphi \in [0, 2\pi)$. While the optimal squeezing angle for heterodyne measurements is $\varphi = \pi$, the optimal φ for homodyne measurements depends on the phase θ.

Since the homodyne measurement informs us of the value of the quadrature $\hat{q}$, the squeezing direction of the probe state is optimal, when the rotated probe state $\hat{R}(\theta)\hat{D}(\alpha)\hat{S}(re^{i\varphi})|0\rangle = \hat{R}(\theta)\hat{D}(\alpha)\hat{R}(\varphi/2)\hat{S}(r)|0\rangle$ is squeezed along the $\hat{q}$-quadrature such that its Wigner function is elongated along the $\hat{p}$-quadrature. Thus, for any fixed θ, the optimal squeezing angle satisfies $\theta + \frac{\varphi}{2} = m\pi$ for $m \in \mathbb{Z}$, i.e. $\varphi = 2(m\pi - \theta)$. However, since we consider a flat prior and there is hence no initial information on θ available, we leave the squeezing angle as a variable for the following calculations.

For the homodyne measurement, the likelihood to obtain outcome q given the phase θ can again be obtained by integrating the Wigner function from Eq. (10) over the $\hat{p}$-quadrature as in Eq. (44). To this end, we note that the vector of first moments is again $\bar{\mathbf{x}} = \sqrt{2}\alpha(\cos\theta, -\sin\theta)^T$, while the covariance matrix is given by Eq. (14) but with $\varphi \to \varphi + 2\theta$. Accordingly, we find the likelihood

$$\begin{aligned} p(q|\theta) &= |\langle q|\hat{R}(\theta)\hat{D}(\alpha)\hat{S}(re^{i\varphi})|0\rangle|^2 \\ &= \frac{\exp\left[-\frac{\left(x-\sqrt{2}\alpha\cos\theta\right)^2}{\Gamma_{qq}(r,\varphi+2\theta)}\right]}{\sqrt{\pi\Gamma_{qq}(r,\varphi+2\theta)}}, \end{aligned} \tag{49}$$

where $\Gamma_{qq}(r,\varphi) = \cosh(2r) - \cos(\varphi)\sinh(2r)$. The (unconditional) probability $p(q)$ to obtain q is $p(q) = \int_0^{\pi} d\theta\, p(\theta)\, p(q|\theta)$. However, as anticipated from the already complicated form of $p(q)$ for purely displaced probe states, the integration of $p(q)$ from Eq. (49) turns out to be a formidable obstacle and we have not found a closed analytical expression for it. From this point onward, we hence proceed by numerically evaluating $p(q)$, the posterior $p(\theta|q)$, the estimator, the variance, and the average variance for different displacement strengths (r) and angles (φ) as well as for different displacements α. In particular, we plot the resulting average variance $\bar{V}_{\text{post}}$ as a function of $|\alpha|^2$ and as a function of the average

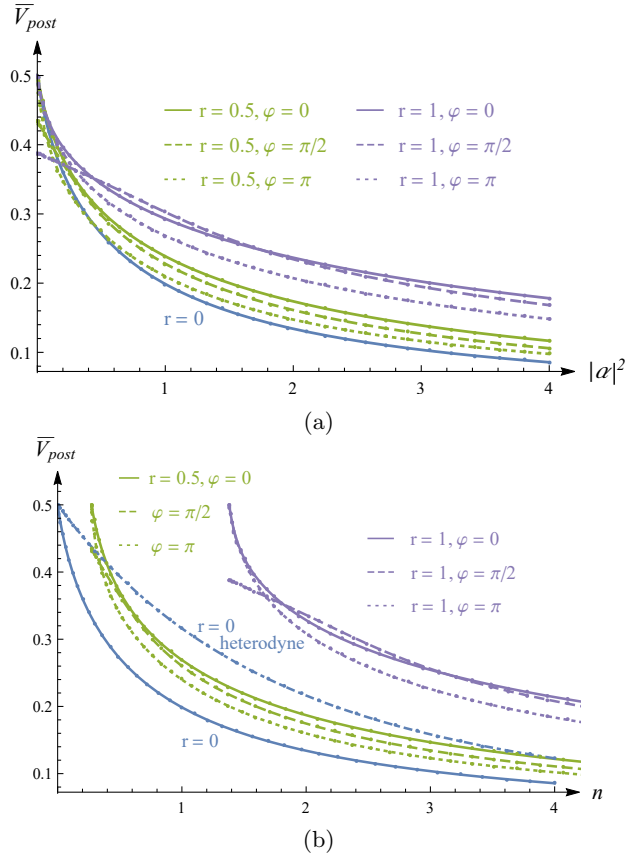

FIG. 5. Bayesian phase estimation with single-mode Gaussian probes. (a) The average variance $\bar{V}_{\text{post}}$ is shown as a function of $|\alpha|^2$, i.e., the energy invested in displacing the probe state. Each curve corresponds to varying values of $\alpha \geq 0$, but fixed squeezing strength r from $r = 0$ (blue), over $r = 0.5$ (green), to $r = 1$ (purple), and fixed squeezing angle φ, from $\varphi = 0$ (solid), over $\varphi = \pi/2$ (dashed), to $\varphi = \pi$ (dotted). Curves for $\varphi = 3\pi/2$ are identical to those for $\varphi = \pi/2$. (b) The average variance $\bar{V}_{\text{post}}$ is shown as a function of the average photon number $n = \alpha^2 + \sinh^2 r$ of the probe state. The colour-coding is the same as in (a), but the lines do not start at $n = 0$ because the nonzero values of r give rise to non-zero average energies even for $\alpha = 0$. In addition, (b) shows $\bar{V}_{\text{post}}$ for a coherent probe state ($r = 0$) and heterodyne detection from Eq. (38) as a blue dashed-dotted curve.

photon number $n = |\alpha|^2 + \sinh^2 r$ in Figs. 5 (a) and (b), respectively.

We first observe that the average variance for the vacuum state ($\alpha = 0 = r$) is 1/2, the same value as for the flat prior. Indeed, any non-zero squeezing appears to improve upon this probe state. However, for increasing displacements, squeezing seems to have a detrimental effect compared to purely displaced states with the same α as seen in Fig. 5 (a), where the average variance of purely displaced states is the smallest except in a regime of small α. When comparing probe states at fixed average energy, it becomes even more clear that squeezing of the probe states in combination with homodyne detection results in strictly worse performance relative to purely displaced probe states. Moreover, a comparison with the combination of coherent probe states and heterodyne detection suggests that coherent probe states and homodyne detection outperform any strategy for Bayesian phase estimation (with flat priors) using Gaussian states and heterodyne detection. However, we note that homodyne detection does not allow us to distinguish between phases shifted by π. If one wishes to explore the full range from $[-\pi, \pi)$, heterodyne detection should be chosen instead.

V. SQUEEZING ESTIMATION

In this section we present a Bayesian estimation strategy for estimating the squeezing strength r of a squeezing operation $\hat{S}(\xi)$, where $\xi = re^{i\varphi}$, as defined in Eq. (12). The squeezing angle φ is assumed to be known. We make this simplifying assumption here, since the investigation of the Bayesian estimation of the single parameter r alone is already computationally demanding, which would only be exacerbated by considering a two-parameter estimation problem.

Optimal covariant measurement strategies for variants of this estimation problem have been presented in [71, 72]. However, the corresponding optimal POVMs may be sufficiently more difficult to realize practically than the Gaussian measurements we consider here. Moreover, we will focus on investigating the performance of different probe states using solely homodyne detection. This is motivated by the findings of the previous sections, namely, that Gaussian strategies for Bayesian single-parameter estimation based on homodyne detection typically outperform those based on heterodyne detection. As we have previously mentioned, this may be a consequence of the intrinsic uncertainties of the coherent states corresponding to the outcomes of the heterodyne measurement. This intuition is also backed up by similar observations made in [6, 72], as well as tentative numerical comparisons we have made. The aim of this section is hence to identify practically realizable strategies for estimating the squeezing strength based on single-mode Gaussian states and homodyne detection. Nevertheless, we should mention here that heterodyne detection should not be disregarded entirely, since there may be scenarios, such as the simultaneous estimation of squeezing strength and angle, where such a strategy could prove to be advantageous.

In the remainder of this section, we consider a general pure Gaussian probe state $\hat{D}(\alpha)\hat{S}(\chi)\,|0\rangle$, where we write the complex variables $\alpha = \alpha_{\text{R}} + i\alpha_{\text{I}}$ for $\alpha_{\text{R}}, \alpha_{\text{I}} \in \mathbb{R}$ and $\chi = se^{i\psi}$, with vector of first moments $\bar{\mathbf{x}}$ and covariance matrix σ. The squeezing transformation that is to be estimated can be represented by a symplectic matrix M,

$$M = \begin{pmatrix} \cosh r - \cos\varphi \, \sinh r & \sin\varphi \, \sinh r \\ \sin\varphi \, \sinh r & \cosh r + \cos\varphi \, \sinh r \end{pmatrix}, \quad (50)$$

such that the moments of the Wigner function change

according to $\bar{\mathbf{x}} \mapsto M\bar{\mathbf{x}}$ and $\sigma \mapsto M\sigma M^T$ under this transformation. Since we assume the direction of the unknown squeezing to be known, we may choose our reference frame accordingly and set $\varphi = 0$ and $r \in \mathbb{R}$ without loss of generality.

Although homodyne detection is not a covariant measurement (cf. definition in Sec. III), it is still a Gaussian measurement (cf. definition in Sec. II.B.2). Consequently, the likelihood $p(q|r)$ is a Gaussian distribution given by

$$p(q|r) = |\langle q|\hat{S}(r)|\alpha,\chi\rangle|^2 = \frac{\exp\left(\frac{-\mathrm{e}^{2r}(\sqrt{2}\alpha_{\mathrm{R}}\mathrm{e}^{-r}-q)^2}{\cosh 2s-\cos\psi\ \sinh 2s}\right)}{\mathrm{e}^{-r}\sqrt{\pi(\cosh 2s-\cos\psi\ \sinh 2s)}}. \tag{51}$$

The parameter we wish to estimate is not the mean of the likelihood, but is encoded in both the variance and the mean of $p(q|r)$. This makes an analytical treatment of this problem extremely difficult, especially since the function $\exp(\exp(r))$ is known to have a nonelementary antiderivative.

V.A. Vacuum probe state

In the present scenario, the only case where the likelihood of Eq. (51) permits an analytical treatment is the vacuum probe state, i.e., when $\alpha = 0$ and $\chi = 0$, where the likelihood becomes

$$p(q|\delta) = \frac{\exp\left(-\frac{q^2}{2\delta^2}\right)}{\delta\sqrt{2\pi}}, \tag{52}$$

with $\delta := \mathrm{e}^{-r}/\sqrt{2}$. This allows us to use the theory of conjugate priors (see Sec. II.A.1). For normal distributions with unknown standard deviation δ, the conjugate priors are gamma distributions. However, since this special case does not provide a promising strategy for the problem at hand, we omit the calculation here and refer the interested reader to Appendix A.2.

Instead of analysing this scenario further, we argue that the vacuum state and even the whole class of squeezed vacuum states perform rather poorly as probes. For probes of this kind the vector of first moments remains unchanged by the transformation and so the parameter has to be estimated solely by the change of the covariance matrix. The most likely measurement outcomes close to the origin are therefore generally very inconclusive. This reasoning is backed up by tentative numerical explorations, suggesting poor performance for any squeezed vacuum states. Since this strategy does not appear to perform reasonably well, we explore the class of coherent probe states instead in the next section, before considering more general single-mode probe states in Sec. V.C.

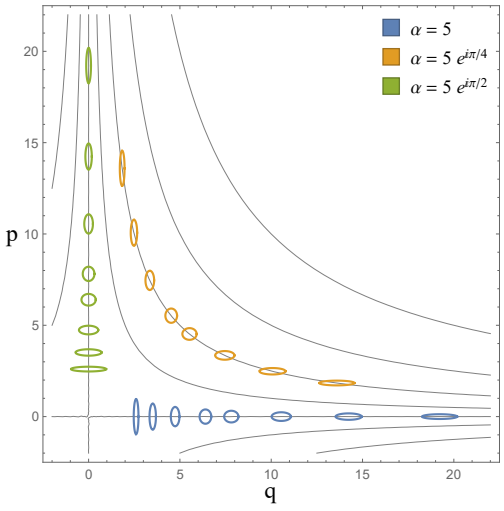

FIG. 6. Coherent probe states for squeezing estimation. The figure shows cross sections of the Wigner functions of coherent probe states with displacements $\alpha = 5$ (blue), $\alpha = 5\ \mathrm{e}^{i\pi/4}$ (orange) and $\alpha = 5\ \mathrm{e}^{i\pi/2}$ (green) after the encoding (squeezing) with strength $r = (-1,\ -0.7,\ -0.4,\ -0.1,\ 0.1,\ 0.4,\ 0.7,\ 1)$ has been applied. The axes show the phase space coordinates q and p. While the shape of the Wigner function can be seen to change with varying squeezing strengths, the mean values $\langle\,\hat{q}\,\rangle$ and $\langle\,\hat{p}\,\rangle$ can be seen to move along hyperbolic trajectories (grey lines).

V.B. Coherent probe states

For coherent probe states, the parameter r is encoded both in the mean and the variance of the likelihood, see Eq. (51). This makes the estimation more efficient, as probes encoded with different values of the parameter become more distinguishable.

Under the influence of a squeezing transformation with unknown strength the mean of our probe state moves along hyperbolic trajectories in phase space, as illustrated in Fig. 6. To simplify our analysis, we pick a trajectory corresponding to a straight line for our estimation. All states with purely real or imaginary displacement lie on such a trajectory (e.g., the states whose Wigner functions are shown in blue and green in Fig. 6) and without loss of generality we assume a positive (real) displacement in $\hat{q}$ together with a homodyne detection in $\hat{q}$. Now the distinguishability of the states with respect to a measurement in $\hat{q}$ is maximal, since the measurement direction is always parallel to the change of the probes mean, ensuring a globally stable measurement procedure. This would not hold for the other hyperbolic trajectories, where the optimal direction of the homodyning (tangential to the curve) would depend on the location on the curve, i.e., the unknown squeezing strength.

With these justified simplifications, our scenario now only has one degree of freedom in the probe preparation, i.e., the displacing amplitude, and none in the measure-

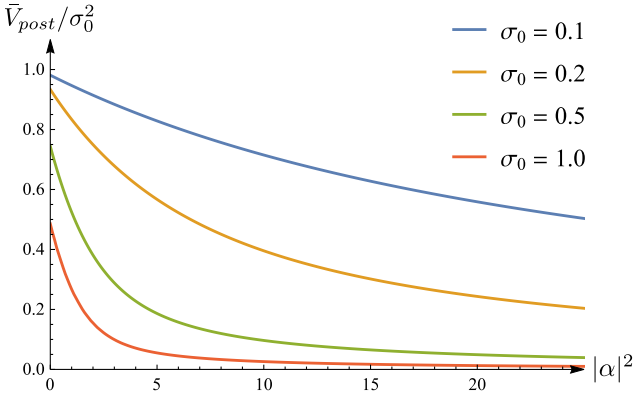

FIG. 7. Ratio between posterior and prior variance for squeezing-strength estimation with coherent probe. The plot shows the quotient of the average variance of the posterior and the variance of the prior plotted against the displacement of the probe state. Different lines show different prior variances. The rate at which we acquire knowledge about the squeezing parameter r decreases for increasing knowledge of that parameter. The prior is a normal distribution with mean $r_0 = 1$ and variance σ_0^2.

ment basis.

In Fig. 7 we show numerical results, indicating already a remarkably good performance of this estimation strategy.

V.C. Displaced-squeezed probe states

To improve our method further, we reduce the uncertainty in the $\hat{q}$-quadrature direction in a similar fashion as in Sec. III for displacement estimation, i.e., we reduce the uncertainty of the probe in the direction we are interested in by squeezing it beforehand. Fig. 8 (a) illustrates this in phase space. Fig. 8 (b) shows how the performance of the estimation is improved by increasing the initial squeezing of the probe and compares the results to the Van Trees bound of Eq. (9). There, the prior is taken to be a normal distribution with variance $\sigma_0^2 = 1$, such that $I\big[p(r)\big] = 1$, and the QFI is optimized over all single-mode Gaussian states with fixed average photon number n, which yields $\mathcal{I}\big[\rho(\theta)\big] = 2(2n+1)^2$, see [6, Eqs. (16) and (18)]. This inequality gives a lower bound on the average posterior variance, but it is unclear if there exists strategies that can saturate it. In Fig. 8 (c), the use of squeezing and displacement in the preparation of the probe are directly compared, and the optimal combinations of these two operations for mean photon number are identified.

Although homodyne detection is not the optimal (maximising the FI) POVM for squeezing estimation in the local/frequentist regime, our analysis provides efficient estimation strategies using only elementary quantum optics methods. In particular, these strategies rely only

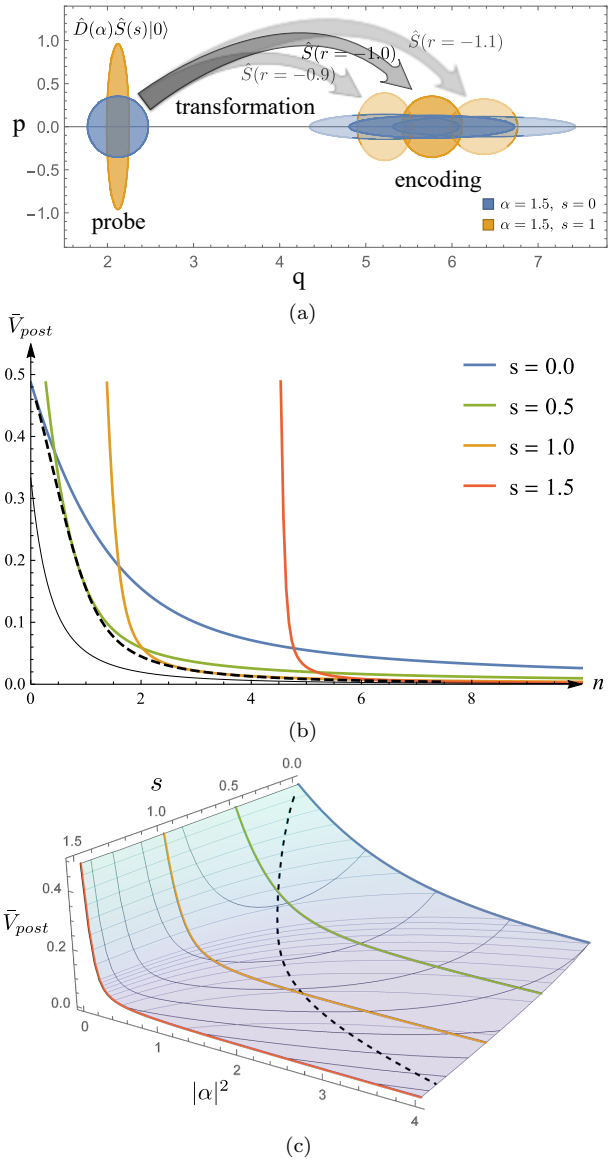

FIG. 8. Initial squeezing and displacement improve the estimation of the squeezing strength r. (a) shows this behaviour in phase space: two differently squeezed probe states (left side) are transformed with an unknown squeezing transformation. For slightly different squeezing strengths ($r = -0.9, -1.0, -1.1$) the unsqueezed probe state ($s = 0$, blue) overlaps for the different cases, thus making it hard to estimate the parameter exactly. The initially squeezed probe state ($s = 1$, orange) is still clearly distinguishable after the different transformations. (b) shows the average variance of the posterior $\bar{V}_{\text{post}}$ as a function of the average photon number n for the two states from (a) and for two more probes with $s = 0.5$ (green) and $s = 1.5$ (red). The dashed, black line shows $\bar{V}_{\text{post}}$ achieved with the optimal single-mode Gaussian states at fixed n. The solid, black line shows the lower bound given by the Van Trees inequality (see Sec. V.C). (c) shows $\bar{V}_{\text{post}}$ for different values of the squeezing s and displacement α of the probe state. The black curves represent lines of constant photon number $[n = |\alpha|^2 + \sinh^2(s)]$, whereas the dashed, black line minimises the average variance for fixed n. The four curves from (b) are shown in the same color-coding. The prior used in both (b) and (c) is a normal distribution with mean $r_0 = -0.5$ and variance $\sigma_0^2 = 1$.

on single-mode Gaussian states and homodyne detection, allowing a comparably straightforward experimental implementation.

VI. DISCUSSION & CONCLUSION

In this paper, we have aimed to provide a comprehensive investigation of Bayesian parameter estimation with single-mode Gaussian states and suitable Gaussian measurements. Notably, the Bayesian approach allows us to study regimes of uncertainty for the estimated parameter (e.g., flat priors, single measurements), where local estimation is not justified. Our focus has not been on finding optimal states and measurements maximising the quantum Fisher information. Instead, we have focused on discovering what can be achieved with practically easily realizable techniques: single-mode Gaussian states combined with heterodyne and homodyne detection. Besides the relevance for experimental implementations, this investigation of single-mode Gaussian states within the theory of Bayesian estimation also creates an important reference point for future explorations of more complicated probe states and measurements. Within this setting, we have investigated three paradigmatic cases of CV quantum metrology: the estimation of displacements, phase rotations, and single-mode squeezing strengths. For the Bayesian estimation of displacements, we provide a fully analytic treatment for Gaussian priors, and for arbitrary single-mode states combined with heterodyne or homodyne detection. For the estimation of a single phase-space coordinate we prove the optimality of the presented strategy. This optimal strategy entails investing all available energy into squeezing the probe state in the direction of the displacement and a homodyne measurement in the same direction.

For Bayesian phase estimation, many standard techniques from Bayesian parameter estimation have to be adapted to circular statistics. This makes it challenging to explore this scenario analytically, and we therefore focus on the case of flat priors (i.e., no initial information about the phase) as a polar opposite to the well-studied problem of local phase estimation. We provide closed expressions for the average variance achieved for coherent probe states and heterodyne detection. For all other scenarios we rely on numerical calculations, which show that homodyne detection generally outperforms heterodyne detection when restricting the phase to the interval $[0, \pi]$. In this case, it is best to invest all available energy into displacing the probe.

Finally we consider the estimation of an unknown squeezing strength. Almost all calculations here have to be done numerically. For this we make a series of well justified assumptions and restrict the large parameter space to a small subset, i.e., the displacement and squeezing of the probe state. Our analysis suggests that the best strategy in this case is to split the energy of the probe state amongst squeezing and displacement, and to perform homodyne measurements.

We envisage the results presented here as a first step in the exploration of Gaussian probe states and measurements in the framework of Bayesian parameter estimation. A number of interesting questions regarding optimality, as well as adaptive multi-round schemes come to mind. This could include the adaptive estimation of both coordinates of the complex displacement parameter with homodyne detection alternating in the measurement quadrature as well as adaptive schemes for phase estimation with more general prior functions. Also an extension to multi-mode Gaussian states [11] and the estimation of multiple parameters [73] seem fruitful directions for further investigations. Although these problems are thus left open for future research, the present work represents an important connection to the respective local estimation problems in that it provides practical strategies for drastically reducing the uncertainty about the estimated parameter. Once this has been achieved, one may employ suitable (e.g., asymptotically optimal) local estimation strategies.

ACKNOWLEDGMENTS

We are thankful to Michalis Skotiniotis for spending his time at the beach going through our work and for providing insightful comments. We acknowledge Thomas Busch and Stefan Ataman for useful comments. S.M. and N.F. acknowledge funding from the Austrian Science Fund (FWF): P 31339-N27. A. U. acknowledges support from IQOQI - Vienna for visiting, financial and computational support from OIST Graduate University, especially the computing resources of the Scientific Computing and Data Analysis section, and financial support from a Research Fellowship of JSPS for Young Scientists. E.A. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie IF (InDiQE - EU project 845486).

[1] Rafał Demkowicz-Dobrzański, Janek Kołodyński, and Mădălin Guţă, *The elusive Heisenberg limit in quantum-enhanced metrology,* Nat. Commun. **3**, 1063 (2012), arXiv:1201.3940.

[2] B. M. Escher, L. Davidovich, N. Zagury, and R. L. de Matos Filho, *Quantum Metrological Limits via a Variational Approach,* Phys. Rev. Lett. **109**, 190404 (2012), arXiv:1207.3307.

[3] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, *Noisy Metrology beyond the Standard Quantum Limit,* Phys. Rev. Lett. **111**, 120401 (2013), arXiv:1212.3286.

[4] Pavel Sekatski, Michalis Skotiniotis, and Wolfgang Dür, *Dynamical decoupling leads to improved scaling in noisy quantum metrology,* New J. Phys. **18**, 073034 (2016), arXiv:1512.07476.

[5] Pavel Sekatski, Michalis Skotiniotis, Janek Kołodyński, and Wolfgang Dür, *Quantum metrology with full and fast quantum control,* Quantum **1**, 27 (2017), arXiv:1603.08944.

[6] Roberto Gaiba and Matteo G. A. Paris, *Squeezed vacuum as a universal quantum probe,* Phys. Lett. A **373**, 934 (2009), arXiv:0802.1682.

[7] Olivier Pinel, Pu Jian, Nicolas Treps, Claude Fabre, and Daniel Braun, *Quantum parameter estimation using general single-mode Gaussian states,* Phys. Rev. A **88**, 040102(R) (2013), arXiv:1307.5318.

[8] Alex Monràs, *Phase space formalism for quantum estimation of Gaussian states,* arXiv:1303.3682 [quant-ph] (2013).

[9] Zhang Jiang, *Quantum Fisher information for states in exponential form,* Phys. Rev. A **89**, 032128 (2014), arXiv:1310.2687.

[10] Nicolai Friis, Michalis Skotiniotis, Ivette Fuentes, and Wolfgang Dür, *Heisenberg scaling in Gaussian quantum metrology,* Phys. Rev. A **92**, 022106 (2015), arXiv:1502.07654.

[11] Dominik Šafránek, Antony R. Lee, and Ivette Fuentes, *Quantum parameter estimation using multi-mode Gaussian states,* New J. Phys. **17**, 073016 (2015), arXiv:1502.07924.

[12] Dominik Šafránek and Ivette Fuentes, *Optimal probe states for the estimation of Gaussian unitary channels,* Phys. Rev. A **94**, 062313 (2016), arXiv:1603.05545.

[13] Luca Rigovacca, Alessandro Farace, Leonardo A. M. Souza, Antonella De Pasquale, Vittorio Giovannetti, and Gerardo Adesso, *Versatile Gaussian probes for squeezing estimation,* Phys. Rev. A **95**, 052331 (2017), arXiv:1703.05554.

[14] Dominik Šafránek, *Estimation of Gaussian quantum states,* J. Phys. A: Math. Theor. **52**, 035304 (2019), arXiv:1801.00299.

[15] Changhun Oh, Changhyoup Lee, Leonardo Banchi, Su-Yong Lee, Carsten Rockstuhl, and Hyunseok Jeong, *Optimal measurements for quantum fidelity between Gaussian states and its relevance to quantum metrology,* Phys. Rev. A **100**, 012323 (2019), arXiv:1901.02994.

[16] Jesús Rubio, Paul Knott, and Jacob Dunningham, *Non-asymptotic analysis of quantum metrology protocols beyond the Cramér-Rao bound,* J. Phys. Commun. **2**, 015027 (2018), arXiv:1707.05022.

[17] Jesús Rubio and Jacob Dunningham, *Quantum metrology in the presence of limited data,* New J. Phys. **21**, 043037 (2019), arXiv:1810.12857.

[18] Jesús Rubio and Jacob Dunningham, *Bayesian multiparameter quantum metrology with limited data,* Phys. Rev. A **101**, 032114 (2020), arXiv:1906.04123.

[19] Ángel Rivas and Alfredo Luis, *Sub-Heisenberg estimation of non-random phase shifts,* New J. Phys. **14**, 093052 (2012), arXiv:1105.6310.

[20] Dominic W. Berry, Michael J. W. Hall, Marcin Zwierz, and Howard M. Wiseman, *Optimal Heisenberg-style bounds for the average performance of arbitrary phase estimates,* Phys. Rev. A **86**, 053813 (2012), arXiv:1209.3547.

[21] Y R Zhang, G R Jin, J P Cao, W M Liu, and H Fan, *Unbounded quantum Fisher information in two-pathinterferometry with finite photon number,* Phys. A: Math. Theor. **46**, 035302 (2013), arXiv:1105.2990.

[22] Marcin Jarzyna and Rafał Demkowicz-Dobrzański, *True precision limits in quantum metrology,* New J. Phys. **17**, 013010 (2015), arXiv:1407.4805.

[23] Wojciech Górecki, Rafał Demkowicz-Dobrzański, Howard M. Wiseman, and Dominic W. Berry, *π-Corrected Heisenberg Limit,* Phys. Rev. Lett. **124**, 030501 (2020), arXiv:1907.05428.

[24] Matteo G. A. Paris, *Quantum estimation for quantum technology,* Int. J. Quantum Inf. **7**, 125 (2009), arXiv:0804.2981.

[25] M. Aspachs, J. Calsamiglia, R. Muñoz Tapia, and E. Bagan, *Phase estimation for thermal Gaussian states,* Phys. Rev. A **79**, 033834 (2009), arXiv:0811.3408.

[26] Géza Tóth and Iagoba Apellaniz, *Quantum metrology from a quantum information science perspective,* J. Phys. A: Math. Theor. **47**, 424006 (2014), arXiv:1405.4878.

[27] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, *Quantum limits in optical interferometry,* Prog. Optics **60**, 345 (2015), arXiv:1405.7703.

[28] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein, *Quantum metrology with nonclassical states of atomic ensembles,* Rev. Mod. Phys. **90**, 035005 (2018), arXiv:1609.01609.

[29] Jasminder S. Sidhu and Pieter Kok, *Geometric perspective on quantum parameter estimation,* AVS Quantum Sci. **2**, 014701 (2020), arXiv:1907.06628.

[30] Lukas J. Fiderer, Jonas Schuff, and Daniel Braun, *Neural-Network Heuristics for Adaptive Bayesian Quantum Estimation,* arXiv:2003.02183 [quant-ph] (2020).

[31] Esteban Martínez-Vargas, Carlos Pineda, François Leyvraz, and Pablo Barberis-Blostein, *Quantum estimation of unknown parameters,* Phys. Rev. A **95**, 012136 (2017), arXiv:1606.07899.

[32] Valeria Cimini, Marco G. Genoni, Ilaria Gianani, Nicolò Spagnolo, Fabio Sciarrino, and Marco Barbieri, *Diagnosing Imperfections in Quantum Sensors via Generalized Cramér-Rao Bounds,* Phys. Rev. Appl. **13**, 024048 (2020), arXiv:2001.01926.

[33] Mauro Valeri, Emanuele Polino, Davide Poderini, Ilaria Gianani, Giacomo Corrielli, Andrea Crespi, Roberto Osellame, Nicoló Spagnolo, and Fabio Sciarrino, *Experimental adaptive Bayesian estimation of multiple phases with limited data,* npj Quantum Inf. **6**, 92 (2020), arXiv:2002.01232.

[34] J. Klauder and B. Skagerstam, *Coherent States* (World Scientific Publishing, Singapore, 1985).

[35] Ulrik L. Andersen, Tobias Gehring, Christoph Marquardt, and Gerd Leuchs, *30 years of squeezed light generation,* Phys. Scr. **91**, 053001 (2016), arXiv:1511.03250.

[36] Marco G. Genoni, Stefano Olivares, and Matteo G.A. Paris, *Optical phase estimation in the presence of phase diffusion,* Phys. Rev. Lett. **106**, 153603 (2011), arXiv:1012.1123.

[37] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, *Quantum Metrology,* Phys. Rev. Lett. **96**, 010401 (2006), arXiv:quant-ph/0509179.

[38] Giulio D'Agostini, *Bayesian Reasoning in Data Analysis: A Critical Introduction* (World Scientific Publishing, Singapore, 2003).

[39] William M. Bolstad, *Understanding Computational*

Bayesian Statistics (John Wiley & Sons, New Jersey, USA, 2009).
[40] M. Antónia Amaral Turkman, Carlos Daniel Paulino, and Peter Müller, *Computational Bayesian Statistics: An Introduction* (Cambridge University Press, Cambridge, U.K., 2019).
[41] Nicolai Friis, Davide Orsucci, Michalis Skotiniotis, Pavel Sekatski, Vedran Dunjko, Hans J. Briegel, and Wolfgang Dür, *Flexible resources for quantum metrology,* New J. Phys. **19**, 063044 (2017), arXiv:1610.09999.
[42] Howard Raïffa and Robert Schlaifer, *Applied statistical decision theory*, 6th ed. (Harvard University, Boston, USA, 1961).
[43] B. Roy Frieden, *Science from Fisher Information: A Unification* (Cambridge University Press, Cambridge, U.K., 2004).
[44] Samuel L. Braunstein and Carlton M. Caves, *Statistical distance and the geometry of quantum states,* Phys. Rev. Lett. **72**, 3439 (1994).
[45] Harry L. Van Trees, *Detection, Estimation, and Modulation Theory: Detection, Estimation, and Linear Modulation Theory* (Wiley Online Library, 2001).
[46] Steven M. Kay, *Fundamentals of statistical signal processing* (Prentice Hall, New Jersey, USA, 1993).
[47] Richard D. Gill and Boris Y. Levit, *Applications of the van Trees Inequality: A Bayesian Cramér-Rao Bound,* Bernoulli **1**, 59 (1995), https://projecteuclid.org/euclid.bj/1186078362.
[48] Werner Vogel and Dirk-Gunnar Welsch, *Quantum Optics*, 3rd ed. (WILEY-VCH Verlag, Weinheim, Germany, 2006).
[49] Peter L. Knight Chistopher C. Gerry, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, U.K., 2005).
[50] Samuel L. Braunstein and Peter van Loock, *Quantum information with continuous variables,* Rev. Mod. Phys. **77**, 513 (2005), arXiv:quant-ph/0410100.
[51] Alessandro Ferraro, Stefano Olivares, and Matteo G. A. Paris, *Gaussian States in Quantum Information*, Napoli Series on Physics and Astrophysics (Bibliopolis, 2005) arXiv:quant-ph/0503237.
[52] Gerardo Adesso and Fabrizio Illuminati, *Entanglement in continuous variable systems: Recent advances and current perspectives,* J. Phys. A: Math. Theor. **40**, 7821 (2007), arXiv:quant-ph/0701221.
[53] Xiang Bin Wang, Tohya Hiroshima, Akihisa Tomita, and Masahito Hayashi, *Quantum information with Gaussian states,* Phys. Rep. **448**, 1 (2007), arXiv:0801.4604.
[54] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd, *Gaussian quantum information,* Rev. Mod. Phys. **84**, 621 (2012), arXiv:1110.3234.
[55] Gerardo Adesso, Sammy Ragy, and Antony R. Lee, *Continuous Variable Quantum Information: Gaussian States and Beyond,* Open Syst. Inf. Dyn. **21**, 1440001 (2014), arXiv:1401.4679.
[56] Arvind, B. Dutta, N. Mukunda, and R. Simon, *The real symplectic groups in quantum mechanics and optics,* Pramana **45**, 471 (1995).
[57] E. Wigner, *On the Quantum Correction For Thermodynamic Equilibrium,* Phys. Rev. **40**, 749 (1932).
[58] H. Yuen and J. Shapiro, *Optical communication with two-photon coherent states–Part I: Quantum-state propagation and quantum-noise,* IEEE Trans. Inf. Theory **24**, 657 (1978).
[59] J. Shapiro, H. Yuen, and A. Mata, *Optical communication with two-photon coherent states–Part II: Photoemissive detection and structured receiver performance,* IEEE Trans. Inf. Theory **25**, 179 (1979).
[60] H. Yuen and J. Shapiro, *Optical communication with two-photon coherent states–Part III: Quantum measurements realizable with photoemissive detectors,* IEEE Trans. Inf. Theory **26**, 78 (1980).
[61] Alexander S. Holevo, *The structure of general quantum Gaussian observable,* (2020), arXiv:2007.02340.
[62] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Massimiliano F. Sacchi, *Covariant quantum measurements that maximize the likelihood,* Phys. Rev. A **70**, 062105 (2004), arXiv:quant-ph/0403083.
[63] S. D. Personick, *Application of quantum estimation theory to analog communication over quantum channels,* IEEE Trans. Inf. Theory **17**, 240 (1971).
[64] A. A. Michelson and E. W. Morley, *On the relative motion of the Earth and the luminiferous ether,* Am. J. Sci. **34**, 333 (1887).
[65] Nathan Wiebe and Chris Granade, *Efficient Bayesian Phase Estimation,* Phys. Rev. Lett. **117**, 010503 (2016), arXiv:1508.00869.
[66] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, *Experimental Bayesian Quantum Phase Estimation on a Silicon Photonic Chip,* Phys. Rev. Lett. **118**, 100503 (2017), arXiv:1703.05169.
[67] F. Martínez-García, D. Vodola, and M. Müller, *Adaptive Bayesian phase estimation for quantum error correcting codes,* New J. Phys. **21**, 123027 (2019), arXiv:1904.06166.
[68] Changhun Oh, Changhyoup Lee, Carsten Rockstuhl, Hyunseok Jeong, Jaewan Kim, Hyunchul Nha, and Su-Yong Lee, *Optimal Gaussian measurements for phase estimation in single-mode Gaussian metrology,* npj Quantum Inf. **5**, 10 (2019), arXiv:1805.08495.
[69] Stefan Ataman, *Optimal Mach-Zehnder phase sensitivity with Gaussian states,* Phys. Rev. A **100**, 063821 (2019), arXiv:1912.04018.
[70] Alexander S. Holevo, *Covariant measurements and imprimitivity systems,* Lect. Notes Math. **1055**, 153 (1984).
[71] G. J. Milburn, Wen Yu Chen, and K. R. Jones, *Hyperbolic phase and squeeze-parameter estimation,* Phys. Rev. A **50**, 801 (1994).
[72] G. Chiribella, G. M. D'Ariano, and M. F. Sacchi, *Optimal estimation of squeezing,* Phys. Rev. A **73**, 062103 (2006), arXiv:quant-ph/0601103.
[73] Lahcen Bakmou, Mohammed Daoud, and Rachid ahl Iaamara, *Multiparameter quantum estimation theory in quantum Gaussian states,* J. Phys. A: Math. Theor. **53**, 385301 (2020), arXiv:2009.00762

APPENDIX

A.1. Phase estimation with heterodyne measurement

In this section, we provide additional details on the calculations for phase estimation using heterodyne detection discussed in Sec. IV.A of the main text.

A.1.I. Coherent probe states & heterodyne measurements

We begin with the estimator for coherent probe states $|\alpha\rangle$ with $\alpha \in \mathbb{R}$ and $\alpha > 0$. In this case, the posterior given outcome β, is given by

$$p(\theta|\beta) = \frac{p(\theta)\,p(\beta|\theta)}{p(\beta)} = \frac{e^{2\alpha|\beta|\cos(\theta-\phi_\beta)}}{2\pi\, I_0(2\alpha|\beta|)}. \tag{A.1}$$

For evaluating the estimator $\hat{\theta}(\beta) = \arg\langle e^{i\theta}\rangle_{p(\theta|\beta)}$, we then note that

$$\int_{-\pi}^{\pi} d\theta\, e^{2\alpha|\beta|\cos(\theta-\phi_\beta)}\,\sin(\theta-\phi_\beta) = 0, \tag{A.2}$$

which implies that

$$\begin{aligned}\int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\,\sin(\theta-\phi_\beta) &= \\ = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\left[\sin\theta\,\cos\phi_\beta - \cos\theta\,\sin\phi_\beta\right] &= 0.\end{aligned} \tag{A.3}$$

Consequently, we have

$$\tan\phi_\beta = \frac{\sin\phi_\beta}{\cos\phi_\beta} = \frac{\int d\theta\, p(\theta|\beta)\,\sin\theta}{\int d\theta\, p(\theta|\beta)\,\cos\theta} = \frac{\Im\big(\langle e^{i\theta}\rangle\big)}{\Re\big(\langle e^{i\theta}\rangle\big)}, \tag{A.4}$$

such that our estimator is simply the phase of the outcome β, i.e.,

$$\hat{\theta}(\beta) = \arg\langle e^{i\theta}\rangle_{p(\theta|\beta)} = \phi_\beta. \tag{A.5}$$

For the average variance of the posterior, we have to evaluate an integral over all values of $\beta \in \mathbb{C}$, which can easily be done in polar coordinates, i.e., $\beta = |\beta|e^{-i\phi_\beta}$ such that $\int d^2\beta = \int_0^\infty d|\beta|\,|\beta| \int_{-\pi}^{\pi} d\phi_\beta$. With this, we can insert from Eqs. (34) and (37), and calculate

$$\begin{aligned}\bar{V}_{\text{post}} &= \int d^2\beta\; p(\beta)\, V_{\text{post}}(\beta) \\ &= \tfrac{e^{-\alpha^2}}{2\pi\,\Gamma(2)} \int_0^\infty d|\beta| \int_{-\pi}^{\pi} d\phi_\beta\; |\beta|\, e^{-|\beta|^2}\, {}_0F_1(2;\alpha^2|\beta|^2) \\ &= \tfrac{e^{-\alpha^2}}{\Gamma(2)} \int_0^\infty d|\beta|\,|\beta|\, e^{-|\beta|^2}\, {}_0F_1(2;\alpha^2|\beta|^2) = \tfrac{1-e^{-|\alpha|^2}}{2\,|\alpha|^2},\end{aligned} \tag{A.6}$$

which yields the result as stated in Eq. (38).

A.1.II. Displaced squeezed probe states & heterodyne measurements

In this section, we provide additional details on the calculations in Sec. IV.A.2 of the main text. There, we consider Bayesian phase estimation using displaced squeezed states $\hat{D}(\alpha)\hat{S}(\xi)$, where $\alpha > 0$ and $\xi = re^{i\varphi}$ with $r \geq 0$ and $\varphi = \pi$, combined with heterodyne detection represented by a POVM $\left\{\frac{1}{\pi}|\beta\rangle\langle\beta|\right\}_{\beta\in\mathbb{C}}$ with elements that are proportional to projectors on coherent states $|\beta\rangle = \hat{D}(\beta)|0\rangle$. In this scenario, the likelihood for obtaining measurement outcome $\beta = |\beta|e^{-i\phi_\beta}$ given that the estimated phase has the value θ, given by Eq. (39) in the main text, can be rewritten as

$$\begin{aligned}p(\beta|\theta) = \frac{e^{-\alpha^2(1-\tanh r)-|\beta|^2}}{\pi\cosh r} &\exp\Big[2\alpha|\beta|\cos(\theta-\phi_\beta)\Big] \\ &\times \exp\Big[|\beta|^2\,\tanh r\,\cos\left[2(\theta-\phi_\beta)\right]\Big] \\ &\times \exp\Big[-2\alpha|\beta|\tanh r\,\cos(\theta-\phi_\beta)\Big].\end{aligned} \tag{A.7}$$

We can then use the Jacobi-Anger expansion in terms of the modified Bessel functions of the first kind, i.e.,

$$e^{x\cos\theta} = \sum_{n=-\infty}^{\infty} I_n(x) e^{in\theta}, \tag{A.8}$$

and write the unconditional probability $p(\beta)$ as

$$\begin{aligned}p(\beta) = \tfrac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\, p(\beta|\theta) = \sum_{\substack{n,m_1,m_2\\=-\infty}}^{\infty} \int_{-\pi}^{\pi} d\theta\, e^{i(n+2m_1+m_2)(\theta-\phi_\beta)} \\ \times \frac{e^{-\alpha^2(1-\tanh r)-|\beta|^2}}{2\pi^2\cosh r} e^{i(n+m_1+m_2)\pi}\, I_n(-2\alpha|\beta|) \\ \times I_{m_1}(-|\beta|^2\tanh r)\, I_{m_2}(2\alpha|\beta|\tanh r).\end{aligned} \tag{A.9}$$

We then make use of the identity

$$\int_{-\pi}^{\pi} d\theta\, e^{i(n+2m_1+m_2)(\theta-\phi_\beta)} = \begin{cases} 2\pi & \text{if } n = -2m_1 - m_2 \\ 0 & \text{otherwise}\end{cases}, \tag{A.10}$$

such that we obtain

$$\begin{aligned}p(\beta) = \frac{e^{-\alpha^2(1-\tanh r)-|\beta|^2}}{\pi\cosh r} \sum_{\substack{m_1,m_2\\=-\infty}}^{\infty} e^{-im_1\pi}\, I_{-2m_1-m_2}(-2\alpha|\beta|) \\ \times I_{m_1}(-|\beta|^2\tanh r)\, I_{m_2}(2\alpha|\beta|\tanh r).\end{aligned} \tag{A.11}$$

By setting $e^{-im_1\pi} = (-1)^{m_1}$, we thus obtain the expression for the unconditional probability $p(\beta)$ from Eq. (41) of the main text. Using Bayes' law, the posterior is obtained as $p(\theta|\beta) = p(\beta|\theta)/\big[2\pi\,p(\beta)\big]$.

To evaluate the estimator $\hat{\theta}(\beta) = \arg\langle e^{i\theta}\rangle_{p(\theta|\beta)}$, we pro-

ceed in a similar way as above. We first calculate

$$\langle e^{i\theta}\rangle = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\, e^{i\theta} = \frac{1}{2\pi K} \sum_{n_1,n_2,n_3=-\infty}^{\infty} e^{i(n_1+n_2+n_3)\pi} \times I_{n_1}(-2\alpha|\beta|)\, I_{n_2}(-|\beta|^2 \tanh r)\, I_{n_3}(2\alpha|\beta|\tanh r) \times \int_{-\pi}^{\pi} d\theta\, e^{i\theta} e^{i(n_1+2n_2+n_3)(\theta-\phi_\beta)}, \tag{A.12}$$

where

$$K := \sum_{m_1,m_2=-\infty}^{\infty} (-1)^{m_1} I_{-2m_1-m_2}(-2\alpha|\beta|) \times I_{m_1}(-|\beta|^2\tanh r) I_{m_2}(2\alpha|\beta|\tanh r). \tag{A.13}$$

Here, we can make use of a similar identity as in Eq. (A.10), i.e.,

$$\int_{-\pi}^{\pi} d\theta\, e^{i\theta} e^{i(n_1+2n_2+n_3)(\theta-\phi_\beta)} = \begin{cases} 2\pi\, e^{i\phi_\beta} & \text{if } n_1 = -2n_2 - n_3 - 1 \\ 0 & \text{otherwise,} \end{cases} \tag{A.14}$$

such that we obtain

$$\langle e^{i\theta}\rangle = \frac{e^{i\phi_\beta}}{K} \sum_{n_2,n_3=-\infty}^{\infty} (-1)^{-n_2-1} I_{-2n_2-n_3-1}(-2\alpha|\beta|) \times I_{n_2}(-|\beta|^2\tanh r) I_{n_3}(2\alpha|\beta|\tanh r). \tag{A.15}$$

Here, $K \geq 0$, since $K = K(\beta)$ is proportional to the probability distribution $p(\beta)$ and the proportionality factor is non-negative. The remaining sum on the right-hand side of Eq. (A.15) is strictly real-valued, which can be seen by noting that $I_n(x)$ is real when both the order n and argument x are real. However, the sum over modified Bessel functions may take positive and negative values.

If the sum is positive, the estimator corresponds to the phase of the outcome, $\hat{\theta}(\beta) = \arg\langle e^{i\theta}\rangle_{p(\theta|\beta)} = \phi_\beta$, whereas the estimate is shifted by π [i.e., $\hat{\theta}(\beta) = \phi_\beta + \pi$] if the sum is negative. As seen below (particularly, Eq. (A.17)), the distinction between these two cases does not affect the variance of the posterior, because the deviation function $\sin^2\big[\theta - \hat{\theta}(\beta)\big]$ is invariant under shift by π. For the variance of the posterior, we take the average of $\sin^2\big[\theta - \hat{\theta}(\beta)\big]$, and find

$$V_{\text{post}}(\beta) = \int_{-\pi}^{\pi} d\theta\, p(\theta|\beta)\, \sin^2(\theta - \hat{\theta}(\beta)) \tag{A.16}$$
$$= \frac{1}{2\pi K} \sum_{\substack{n_1,n_2,n_3 \\ =-\infty}}^{\infty} I_{n_1}(-2\alpha|\beta|) I_{n_2}(-|\beta|^2\tanh r) \times I_{n_3}(2\alpha|\beta|\tanh r) e^{i(n_1+n_2+n_3)\pi} \times \int_{-\pi}^{\pi} d\theta\, e^{i(n_1+2n_2+n_3)(\theta-\phi_\beta)}\, \sin^2[\theta - \hat{\theta}(\beta)].$$

We can again make use of an identify similar to Eq. (A.10), i.e.,

$$\int_{-\pi}^{\pi} d\theta\, e^{i(n_1+2n_2+n_3)(\theta-\phi_\beta)}\, \sin^2[\theta - \hat{\theta}(\beta)] = \begin{cases} \pi & \text{if } n_1 = -2n_2 - n_3 \\ -\frac{\pi}{2} & \text{if } n_1 = -2n_2 - n_3 \pm 2 \\ 0 & \text{otherwise} \end{cases}, \tag{A.17}$$

such that we obtain

$$V_{\text{post}}(\beta) = \frac{1}{2K} \sum_{n_1,n_2,n_3=-\infty}^{\infty} (-1)^{n_1+n_2+n_3} I_{n_1}(-2\alpha|\beta|) \times I_{n_2}(-|\beta|^2\tanh r) I_{n_3}(2\alpha|\beta|\tanh r) \times \big(\delta_{n_1,-2n_2-n_3+2}\tfrac{-1}{2} + \delta_{n_1,-2n_2-n_3} + \delta_{n_1,-2n_2-n_3-2}\tfrac{-1}{2}\big). \tag{A.18}$$

To obtain the average variance of the posterior, we switch to polar coordinates, $\beta = |\beta| e^{-i\phi_\beta}$, such that

$$\bar{V}_{\text{post}} = \int d^2\beta\, p(\beta)\, V_{\text{post}}(\beta) = \int_0^{\infty} d|\beta| \int_{-\pi}^{\pi} d\phi_\beta \frac{e^{-\alpha^2(1-\tanh r)}}{2\pi\cosh r} \times \sum_{\substack{n_2,n_3 \\ =-\infty}}^{\infty} |\beta|\, e^{-|\beta|^2} I_{n_2}(-|\beta|^2\tanh r) I_{n_3}(2\alpha|\beta|\tanh r) \times \tfrac{1}{2}(-1)^{n_2}\Big[2I_{-2n_2-n_3}(-2\alpha|\beta|) - I_{2-2n_2-n_3}(-2\alpha|\beta|) - I_{-2-2n_2-n_3}(-2\alpha|\beta|)\Big]$$
$$= \frac{e^{-\alpha^2(1-\tanh r)}}{\cosh r} \sum_{\substack{n_2,n_3 \\ =-\infty}}^{\infty} \int_0^{\infty} d|\beta|\, |\beta|\, e^{-|\beta|^2} I_{n_2}(-|\beta|^2\tanh r) \times I_{n_3}(2\alpha|\beta|\tanh r) \tfrac{1}{2}(-1)^{n_2}\Big[2I_{-2n_2-n_3}(-2\alpha|\beta|) - I_{2-2n_2-n_3}(-2\alpha|\beta|) - I_{-2-2n_2-n_3}(-2\alpha|\beta|)\Big], \tag{A.19}$$

which coincides with the expression in Eq. (43). We have not found an analytical expression for the above integral so far, but we have evaluated the integral numerically.

A.1.III. Coherent probe states & homodyne measurements

Here, we provide additional details on the calculations in Sec. IV.B.1 of the main text. There, we consider Bayesian phase estimation with coherent probe states $\hat{D}(\alpha)|0\rangle = |\alpha\rangle$, where $\alpha > 0$, combined with homodyne detection represented by a POVM $\{|q\rangle\langle q|\}_{\beta\in\mathbb{R}}$. In this scenario, the likelihood for measurement outcome q given the phase θ is provided by Eq. (45) in the main text, which can be rewritten as

$$p(q|\theta) = \frac{1}{\pi\sqrt{\pi}} e^{-q^2-\alpha^2} \int_0^{\pi} d\theta e^{2\sqrt{2}q\alpha\cos\theta} e^{-\alpha^2\cos(2\theta)}. \tag{A.20}$$

We express the Jacobi-Anger expansion Eq. (A.8) in a real representation as

$$e^{x\cos\theta} = I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)\cos(n\theta), \tag{A.21}$$

since $I_n(x) = I_{-n}(x)$. Noticing that the range of θ is $[0,\pi]$, the (unconditional) probability to obtain outcome q is given by

$$\begin{aligned} p(q) = \frac{1}{\pi}\int_0^{\pi} d\theta\, p(q|\theta) = \frac{e^{-q^2-\alpha^2}}{\pi\sqrt{\pi}}\int_0^{\pi} d\theta \Big[& I_0(2\sqrt{2}q\alpha)I_0(-\alpha^2) \\ & + 2\,I_0(-\alpha^2)\sum_{n=1}^{\infty} I_n(2\sqrt{2}q\alpha)\cos(n\theta) \\ & + 2\,I_0(2\sqrt{2}q\alpha)\sum_{m=1}^{\infty} I_m(-\alpha^2)\cos(2m\theta) \\ & + 4\sum_{m,n=1}^{\infty} I_m(-\alpha^2)I_n(2\sqrt{2}q\alpha)\cos(n\theta)\cos(2m\theta)\Big]. \end{aligned} \tag{A.22}$$

We then use the identities $\int_0^{\pi} d\theta\,\cos(n\,\theta) = 0\ \forall\, n \geq 1$ and

$$\int_0^{\pi} d\theta\,\cos(n\,\theta)\cos(2m\,\theta) = \begin{cases} \frac{\pi}{2} & \text{if } n = 2m \\ 0 & \text{otherwise} \end{cases}. \tag{A.23}$$

With this, we obtain

$$\begin{aligned} p(q) = \frac{1}{\pi}\int_0^{\pi} d\theta\, p(q|\theta) = \frac{e^{-q^2-\alpha^2}}{\sqrt{\pi}}\Big[& I_0(2\sqrt{2}q\alpha)I_0(-\alpha^2) \\ & + 2\sum_{m=1}^{\infty} I_{2m}(2\sqrt{2}q\alpha)I_m(-\alpha^2)\Big] = \frac{e^{-q^2-\alpha^2}}{\sqrt{\pi}}\,M, \end{aligned} \tag{A.24}$$

where

$$M := \sum_{m=-\infty}^{\infty} I_{2m}(2\sqrt{2}q\alpha)I_m(-\alpha^2). \tag{A.25}$$

The posterior $p(\theta|q)$ is then obtained as $p(q|\theta)/[\pi p(q)]$.

To determine the estimator $\hat{\theta}(q) = \arg\langle e^{i\theta}\rangle_{p(\theta|q)}$, we calculate $\langle e^{i\theta}\rangle_{p(\theta|q)}$, i.e.,

$$\begin{aligned} \langle e^{i\theta}\rangle = \int_0^{\pi} d\theta\, p(\theta|q)\, e^{i\theta} = \frac{1}{M}\frac{1}{\pi}\int_0^{\pi} d\theta \Big[& I_0(2\sqrt{2}q\alpha)I_0(-\alpha^2)e^{i\theta} \\ & + 2I_0(-\alpha^2)\sum_{n=1}^{\infty} I_n(2\sqrt{2}q\alpha)\cos(n\theta)e^{i\theta} \\ & + 2I_0(2\sqrt{2}q\alpha)\sum_{m=1}^{\infty} I_m(-\alpha^2)\cos(2m\theta)e^{i\theta} \\ & + 4\sum_{m,n=1}^{\infty} I_n(2\sqrt{2}q\alpha)I_m(-\alpha^2)\cos(n\theta)\cos(2m\theta)e^{i\theta}\Big]. \end{aligned} \tag{A.26}$$

We then use the identities $\int_0^{\pi} d\theta\, e^{i\theta} = 2i$,

$$\int_0^{\pi} d\theta\,\cos(n\theta)\, e^{i\theta} = \begin{cases} \pi/2 & \text{if } n = 1 \\ \frac{i(1+(-1)^n)}{1-n^2} & \text{if } n \geq 2 \\ 0 & \text{otherwise} \end{cases}, \tag{A.27}$$

and

$$\int_0^{\pi} d\theta\,\cos(n\theta)\cos(2m\theta)e^{i\theta} = \begin{cases} \frac{\pi}{4} & \text{if } n = 2m \pm 1 \\ \frac{i(1+(-1)^n)(1-4m^2-n^2)}{(n-2m-1)(n-2m+1)(n+2m+1)(n+2m-1)} & \text{otherwise} \end{cases}. \tag{A.28}$$

With this, we obtain

$$\begin{aligned} \langle e^{i\theta}\rangle = \frac{1}{\pi M}\Bigg\{ & 2iI_0(2\sqrt{2}q\alpha)\Big[I_0(-\alpha^2) + \sum_{m=1}^{\infty} I_m(-\alpha^2)\tfrac{2}{1-4m^2}\Big] \\ & + 2I_0(-\alpha^2)\Big[\tfrac{\pi}{2}I_1(2\sqrt{2}q\alpha) + \sum_{n=2}^{\infty} I_n(2\sqrt{2}q\alpha)\tfrac{i(1+(-1)^n)}{1-n^2}\Big] \\ & + 4\sum_{m=1}^{\infty} I_m(-\alpha^2)\Big[\tfrac{\pi}{4}I_{2m-1}(2\sqrt{2}q\alpha) + \tfrac{\pi}{4}I_{2m+1}(2\sqrt{2}q\alpha) \\ & + \sum_{\substack{n=1 \\ n\neq 2m\pm 1}}^{\infty} I_n(2\sqrt{2}q\alpha)\tfrac{i(1+(-1)^n)(1-4m^2-n^2)}{(n-2m-1)(n-2m+1)(n+2m+1)(n+2m-1)}\Big]\Bigg\} \\ = \frac{i}{\pi M}\Bigg\{ & \sum_{n=1}^{\infty}\frac{4}{1-4n^2}\Big[I_0(-\alpha^2)I_{2n}(2\sqrt{2}q\alpha) + I_n(-\alpha^2)I_0(2\sqrt{2}q\alpha)\Big] \\ & + 2I_0(2\sqrt{2}q\alpha)I_0(-\alpha^2) + 8\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} I_{2n}(2\sqrt{2}q\alpha)I_m(-\alpha^2) \\ & \times \frac{1-4m^2-4n^2}{(2n-2m-1)(2n-2m+1)(2n+2m+1)(2n+2m-1)}\Bigg\} \\ & + \frac{1}{M}\Bigg\{I_0(-\alpha^2)I_1(2\sqrt{2}q\alpha) + \sum_{n=1}^{\infty} I_n(-\alpha^2)\Big[I_{2n-1}(2\sqrt{2}q\alpha) \\ & + I_{2n+1}(2\sqrt{2}q\alpha)\Big]\Bigg\}. \end{aligned} \tag{A.29}$$

Finally, we can express the real and imaginary parts of $\langle e^{i\theta}\rangle$ as

$$\Re[\langle e^{i\theta}\rangle] = \frac{\sum_{n=-\infty}^{\infty} I_{2n+1}(2\sqrt{2}q\alpha)I_n(-\alpha^2)}{\sum_{m=-\infty}^{\infty} I_{2m}(2\sqrt{2}q\alpha)I_m(-\alpha^2)} \tag{A.30}$$

and

$$\begin{aligned} \Im[\langle e^{i\theta}\rangle] = \frac{2}{\pi}\frac{\sum_{m,n=-\infty}^{\infty} I_{2n}(2\sqrt{2}q\alpha)I_m(-\alpha^2)}{\sum_{k=-\infty}^{\infty} I_{2k}(2\sqrt{2}q\alpha)I_k(-\alpha^2)} \\ \times \frac{1-4m^2-4n^2}{(2n-2m-1)(2n-2m+1)(2n+2m+1)(2n+2m-1)}, \end{aligned} \tag{A.31}$$

respectively, where we have used the fact that functions $C_{n,m}$ invariant under the exchanges $n \to -n$ and $m \to -m$ satisfy

$$\sum_{n=1}^{\infty} C_{n,m} = \frac{1}{2}\left(\sum_{n=-\infty}^{\infty} C_{n,m} - C_{0,m}\right) \tag{A.32}$$

and

$$\sum_{m,n=1}^{\infty} C_{n,m} = \frac{1}{4}\left(\sum_{m,n=-\infty}^{\infty} C_{n,m} - \sum_{m=-\infty}^{\infty} C_{0,m} - \sum_{n=-\infty}^{\infty} C_{n,0} + C_{0,0}\right). \tag{A.33}$$

The estimator can then be calculated from Eqs. (A.30) and (A.31) via

$$\hat{\theta}(q) = \arctan\left(\frac{\Im[\langle e^{i\theta}\rangle]}{\Re[\langle e^{i\theta}\rangle]}\right). \tag{A.34}$$

A.2. Squeezing estimation using the vacuum state and homodyne detection

In this appendix, we provide additional details on the estimation of the squeezing strength using a vacuum probe state in combination with homodyne detection. We include this to illustrate that the theory of conjugate priors can applied also in more general cases, even if the calculations might become more involving.

The likelihood is given by Eq. (52),

$$p(q|\delta) = \frac{\exp(-\frac{q^2}{2\delta^2})}{\delta\sqrt{2\pi}}, \tag{A.35}$$

where we have defined $\delta := e^{-r}/\sqrt{2}$. For normal distributions with unknown standard deviation δ, the conjugate priors are gamma distributions

$$p(\delta) = \frac{b^a \delta^{a-1} e^{-b\delta}}{\Gamma(a)}, \tag{A.36}$$

$a,\ b > 0$. The mean and variance of such a distribution is given by $\mathrm{E}[p(\delta)] = a/b$ and $\mathrm{Var}[p(\delta)] = a/b^2$, respectively. If the prior is gamma distributed with parameters a and b, then the posterior after m measurements is gamma distributed as well with parameters $a + m/2$ and $b + \sum_i q_i^2/2$, where q_i is the measurement outcome in each round. The mean and variance of the posterior after m repeated measurements with outcomes $\mathbf{q} = (q_1, \ldots, q_m)$ then becomes

$$\mathrm{E}[p(\delta|\mathbf{q})] = \frac{2a+m}{2b+\sum_i q_i^2} \tag{A.37}$$

$$\mathrm{Var}[p(\delta|\mathbf{q})] = \frac{2(2a+m)}{(2b+\sum_i q_i^2)^2}, \tag{A.38}$$

From this point on the formulas become really cumbersome. Since homodyning is not a covariant measurement for the squeezing operator, the variance of our posterior distribution depends on the outcome. To calculate the average variance $\int dq p(q)\,\mathrm{Var}[p(\delta|q)]$, one first needs to calculate

$$\begin{aligned} p(q) &= \int d\delta p(\delta) p(q|\delta) \qquad (A.39)\\ &= \int_0^\infty d\delta \frac{\exp(-\frac{q^2}{2\delta^2})}{\delta\sqrt{2\pi}} \frac{b^a \delta^{a-1} e^{-b\delta}}{\Gamma(a)} \\ &= \frac{1}{\sqrt{\pi}\Gamma(a)}\Big[\frac{b}{\sqrt{2}}\Gamma(a-1)\,{}_pF_q(1; 1-\frac{a}{2}, \frac{3-a}{2}; -\frac{b^2q^2}{8}) \\ &\quad - \frac{b^{a+1}|q|^a}{2^{\frac{3+a}{2}}}\Gamma(-\frac{a}{2})\,{}_pF_q(1; \frac{3}{2}, 1+\frac{a}{2}; -\frac{b^2q^2}{8}) \\ &\quad + \frac{\pi b^a |q|^{a-1}}{2^{1+\frac{a}{2}}\Gamma(\frac{1+a}{2})}\sec(\frac{\pi a}{2})\,{}_pF_q(1; \frac{1}{2}, \frac{1+a}{2}; -\frac{b^2q^2}{8})\Big], \end{aligned}$$

where ${}_pF_q(.;.;.)$ is the generalized hypergeometric function (the subscripts p and q are part of the notation for this function and have nothing to do with the phase space coordinates). With this now we can calculate the average variance after one measurement $m = 1$

$$\begin{aligned} \bar{V}_{\mathrm{post}} &= \int dq p(q)\,\mathrm{Var}[p(\delta|q)] \qquad (A.40)\\ &= \frac{\sqrt{\pi}(2a+1)}{4\sqrt{b}(a-1)}\,{}_pF_q(\frac{1}{2}; -\frac{1}{2}, 1-\frac{a}{2}, \frac{3-a}{2}; \frac{b^3}{4}) \\ &\quad + \frac{2}{3}b^4(2a+1)\frac{\Gamma(1-a)}{\Gamma(5-a)}\,{}_pF_q(2; \frac{5}{2}, \frac{5-a}{2}, 3-\frac{a}{2}; \frac{b^3}{4}) \\ &\quad - \frac{\pi^2(2a+1)2^{-a}b^{\frac{3a}{2}-2}\csc(\pi a)}{\Gamma(\frac{a-2}{2})\Gamma^2(\frac{a+1}{2})}\,{}_pF_q(\frac{a}{2}; \frac{1}{2}, \frac{a-2}{2}, \frac{1+a}{2}; \frac{b^3}{4}) \\ &\quad + \frac{\sqrt{\pi}}{8}(2a+1)b^{\frac{3a-1}{2}}\sec(\frac{\pi a}{2})\frac{\Gamma(-\frac{a}{2})}{\Gamma(a-1)} \\ &\quad \times {}_pF_q(\frac{1+a}{2}; \frac{3}{2}, \frac{a-1}{2}, \frac{2+a}{2}; \frac{b^3}{4}) \end{aligned}$$

Although we were able to calculate an analytical solution, the result in itself is not interesting, but the techniques we have used might be insightful to the reader.

8 Metrology-assisted entanglement distribution in noisy quantum networks

The following research article is currently under peer review. The manuscript can be accessed online under arXiv:2110.15627.

8.1 Contribution

In this work, I contributed to the theoretical formulation of the investigated scenario. The numerical calculations for the presented example were carried out by me, as well as the analysis comparing the presented strategy to other distillation based protocols and the investigation of strategies going beyond dephasing noise. Further I wrote major parts of the manuscript.

Metrology-assisted entanglement distribution in noisy quantum networks

Simon Morelli,[1,2,*] David Sauerwein,[3,†] Michalis Skotiniotis,[4,‡] and Nicolai Friis[1,2,§]

[1]*Institute for Quantum Optics and Quantum Information — IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria*
[2]*Atominstitut, Technische Universität Wien, 1020 Vienna, Austria*
[3]*Amazon Web Services Europe, Zürich, Switzerland*
[4]*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*

We consider the distribution of high-dimensional entangled states to multiple parties via noisy channels and the subsequent probabilistic conversion of these states to desired target states using stochastic local operations and classical communication. We show that such state-conversion protocols can be enhanced by embedded channel-estimation routines at no additional cost in terms of the number of copies of the distributed states. The defining characteristic of our strategy is the use of those copies for which the conversion was unsuccessful for the estimation of the noise, thus allowing one to counteract its detrimental effect on the successfully converted copies. Although this idea generalizes to various more complex situations, we focus on the realistic scenario, where only finitely many copies are distributed and where the parties are not required to process multiple copies simultaneously. In particular, we investigate the performance of so-called one-successful-branch protocols, applied sequentially to single copies and an adaptive Bayesian estimation strategy. Finally, we compare our strategy to more general but less easily practically implementable strategies involving distillation and the use of quantum memories to process multiple copies simultaneously.

I. INTRODUCTION

Entanglement between two or more parties is an important ingredient for the development of many quantum technologies such as, e.g., fault-tolerant quantum computation [1–4], quantum simulation [5–7], and, in particular, communication via quantum networks [8–12]. A landmark goal of quantum technologies is the establishment of a 'quantum internet' [13–15]—a highly interconnected quantum network able to distribute and manipulate entangled quantum states via fibres and optical links. Recent efforts [16] go in the direction of understanding the resources, challenges and opportunities that come along with such an endeavour. In this context, entanglement is a crucial resource that allows spatially separated parties to overcome the restriction of local operations and classical communication (LOCC) thereby implementing classically impossible tasks, such as quantum teleportation [17–19]. However, the technological requirements for quantum processors and networks on a large scale are immense and will have to involve significant improvements in the efficient manipulation and control of quantum systems. A deeper comprehension of the possibilities and limitations for multipartite entanglement distribution and its manipulation via LOCC is hence necessary. The present work aims to contribute to this endeavour by exploring the synergies between stochastic LOCC (SLOCC) protocols—that aim to convert multiparty quantum states—and quantum parameter estimation protocols whose goal is to improve entanglement distribution by precisely identifying the noise suffered by such states when distributed through quantum networks.

On the theoretical side, a lot of progress has been made with regards to the conversion of multipartite entanglement via LOCC (cf. [20–26]), despite the fact that LOCC is notoriously difficult to handle mathematically [27], while practical developments towards efficient entanglement certification techniques (see, e.g., [28–33]) and real-world quantum networks are rapidly progressing (cf. [34]). Yet, basic theoretical questions about multipartite entanglement structures arising in networks have only begun to be explored [35–39]. In particular, apart from recent examples [40, 41], mostly transformations of pure states in the single-copy regime or the asymptotic limit of infinitely many copies have been considered.

Here, we consider a scenario that more closely resembles situations expected to be encountered in future real-world quantum networks: the distribution of multiple but finitely many copies of multipartite quantum states via imperfect or varying channels available only for a restricted period of time. There, one central node, the vendor, is assumed to be capable of implementing the entangling operations necessary to create a highly entangled quantum state that is subsequently distributed to spatially separated parties via imperfect but local quantum channels. In general, the states provided by such a vendor will not match the target state that the parties might wish to employ in their communication or computation protocol of choice. For multiple parties, there is no unique resource state that can be deterministically converted to any other state via LOCC; in fact, the so-called maximally entangled set (MES) of states from which any state can be reached in this way has infinitely many elements [22]. In addition, the vendor might only be able to create a limited variety of such states and may therefore be unable to offer the desired state. Indeed, this might even be in the interest of the users, as they may not want to divulge information on the specific form (and hence the intended use) of their desired target state. Since deterministic LOCC transformations between multi-party pure states (with fixed local dimensions) almost never exist [26, 42], the

* simon.morelli@oeaw.ac.at
† sauerwein.david@gmail.com; This work was done prior to joining AWS.
‡ michail.skoteiniotis@uab.cat
§ nicolai.friis@univie.ac.at

distributed state will have to be converted to the target state *probabilistically*, i.e., via SLOCC. Two states that can be converted into each other via SLOCC are said to belong to the same *entanglement class*.

Optimal SLOCC conversion protocols will generally require acting on many copies simultaneously. Beside the fact that little is known about the structure of such protocols yet, the implementation of complex multi-copy operations also represents a substantial technological challenge. Therefore we focus on sequential strategies that process a single copy at a time. Compared to general protocols the difficulty to implement a sequential strategy does not increase with the number of copies. To employ a finite number, k_s, of copies for multipartite quantum communication or computation, users must purchase, on average, a larger number $k = k_\mathrm{s}/p_\mathrm{s}$ of copies from the vendor to compensate for the nonzero failure probability $p_\mathrm{f} = 1 - p_\mathrm{s} > 0$ of the conversion protocol.

Here, we demonstrate that this apparent disadvantage has important redeeming qualities in the presence of noise. We show that, for certain 'benign' types of noise, those copies for which the state conversion was unsuccessful can nevertheless be useful for obtaining information about the noise and hence partially compensate its effects on the successfully converted copies. Our results thus establish that multipartite entanglement distribution protocols can be assisted by quantum metrology at no additional resource costs in terms of additionally distributed states, and thus feature built-in noise robustness 'for free'.

To reach this conclusion, we first consider a particular situation as described above, i.e., a family of specific noisy multipartite state conversion protocols and show how the use of parameter estimation tools leads to an improvement in the quality of the resulting final states. Based on this exemplary situation, we then discuss the applicability of this approach to more general entanglement distribution protocols and contrast it with alternative strategies based on distillation. Finally, we discuss the merits of metrology-assisted entanglement conversion protocols and argue for their default integration in future quantum networks.

II. NOISY ENTANGLEMENT DISTRIBUTION

We consider a situation where N parties require several copies of a highly entangled N-qudit state ρ_target for a specific task. They purchase a number of copies from a vendor, where the purchased state ρ_vendor differs from the target state ρ_target, but belongs to the same entanglement class and can thus be converted to the target state via SLOCC. The vendor distributes the state ρ_vendor to the N parties via local quantum channels represented by completely-positive, trace-preserving (CPTP) maps, $\{\Lambda_i\}_{i=1}^N$, so that the state shared by the N parties is given by

$$\rho_\mathrm{received} = \Lambda(\rho_\mathrm{vendor}) = \bigotimes_{i=1}^{N} \Lambda_i(\rho_\mathrm{vendor}). \qquad (1)$$

The parties now wish to transform the received state ρ_received into the target state ρ_target within the same entanglement class. This can be done either deterministically or probabilistically. We focus on transformations of fully entangled (i.e., with full local rank) pure states via a so-called one-successful-branch protocol (OSBP) [43, 44]. The OSBP is in fact the optimal SLOCC protocol for almost all N-qudit pure-state transformations [44]. However, the idea of using unsuccessful branches of SLOCC protocols for parameter estimation presented in this paper applies in equal measure to any other probabilistic transformations of pure and mixed states of arbitrary dimensions.

In an OSBP that transforms $|\,\Psi\,\rangle$ into $|\,\Phi\,\rangle$, each party k performs a single two-outcome measurement $\{M_\mathrm{s}^k,\, M_\mathrm{f}^k\}$ with $\sum_{i=\mathrm{s,f}}(M_i^k)^\dagger M_i^k = \mathbb{1}$. The successful outcome s leads to a state that can still be transformed to the final state, while the failure outcome f leads to a state that is no longer fully entangled and can therefore no longer be transformed into $|\,\Phi\,\rangle$. Hence, the final state is only reached if the successful measurement outcome is realised for all parties. That is, only one branch of the OSBP is successful, as suggested by its name, and realized with some probability p_s.

The transformation performed by the N parties to convert the received state into the target state depends on the channel used to distribute the state. For the general case when the channel is not sufficiently known, the task for the parties becomes more challenging. They have to estimate the channel, i.e., identify the channel Λ from a discrete or continuous family of possible candidates, before transforming the received state ρ_received. The goal of the N parties is now to obtain the largest number of copies ρ_corr that are ϵ-close in trace distance to the desired target state ρ_target, i.e.,

$$D(\rho_\mathrm{corr}, \rho_\mathrm{target}) = \frac{1}{2}\mathrm{Tr}\Big[\sqrt{(\rho_\mathrm{corr} - \rho_\mathrm{target})^2}\,\Big] \;\leq\; \epsilon. \qquad (2)$$

If an asymptotically large number of copies of the state is available, the parties can always sacrifice a finite number of such states to estimate the channel to any desired precision before commencing with the state transformation. Consequently, the incurred cost for estimation has no bearing on the final yield of the target state. On the other hand, if only a finite number of copies is available, then the above strategy can be quite wasteful, and protocols that estimate the channel whilst implementing the OSBP at the same time may provide better yields. This is due to the observation that the states resulting from the failure branches of the OSBP potentially carry information about the noisy channel and can thus be used to estimate the latter at no extra cost.

A particular class of channels for which simultaneous channel estimation and state conversion may be beneficial is the family of channels that admit a decomposition into Kraus operators that commute with the POVM corresponding to the OSBP. Such channels include dephasing and phase drift [45] noise which are typical in quantum communication using optical fibres [46]. Alternatively, such channels also describe a situation often encountered in satellite communication where the sender (vendor) and receivers (N parties) do not share a common phase reference during the transmission stage [47].

In this latter case the receivers know that the state they receive is local-unitarily (LU) equivalent to the original state of the vendor and thus is in the same entanglement class. However, the lack of knowledge of the relative phase between the sender's and receivers' phase reference makes channel estimation indispensable, especially if the task they have in mind requires a particular state ρ_{target}, and not any other local-unitarily equivalent state.

In what follows, we shall consider the above class of channels as a special case of more general parameter estimation problems and employ a Bayesian estimation strategy to estimate the parameter (or set of parameters) θ that characterizes the channel, illustrated in Fig. 1. Such a strategy involves one or more rounds of subsequent measurements on different copies of the received states. The knowledge of the parameter, θ, is updated in every round depending on the measurement outcome, m, according to Bayes' law, i.e.,

$$p(\theta|m) = \frac{p(m|\theta)\,p(\theta)}{p(m)}. \tag{3}$$

Based on the posterior distribution of θ one can then assign an estimator $\hat{\theta}(m)$ after each round and adapt the next measurement to this estimate. This makes Bayesian parameter estimation more flexible and better suited for a limited number of probes compared to frequentist estimation. For an increasing number of copies, also a frequentist strategy can be employed to estimate the channel from the failure branches of the OSBP. Here, however, we will focus on scenarios with a very restricted number of copies. In particular, we consider the exemplary situation of local dephasing noise, to provide quantitative support for our strategy of metrology-assisted entanglement distribution.

III. ENTANGLEMENT DISTRIBUTION WITH DEPHASING NOISE

Assume the state distributed by the vendor is a pure N-qudit state from the GHZ class of the form

$$|\Psi\rangle = \sum_{m=0}^{d-1} \sqrt{\lambda_m}\,|m\rangle^{\otimes N}, \tag{4}$$

where we assume without loss of generality that $\lambda_m \geq \lambda_{m+1}$. Let us further assume that the vendor distributes these states via N local dephasing channels, i.e., each qudit undergoes a unitary transformation parameterized by an unknown angle θ,

$$U(\theta) = \sum_{m=0}^{d-1} e^{\mathrm{i}m\theta}\,|m\rangle\langle m|. \tag{5}$$

The parties wish to obtain the maximally-entangled qudit state

$$|\mathrm{GHZ}_d\rangle = \sqrt{\tfrac{1}{d}} \sum_{m=0}^{d-1} |m\rangle^{\otimes N}. \tag{6}$$

To this end, they estimate the unknown phase θ of the dephasing noise acting on the received states and recover the target state via SLOCC-transformations from the received state.

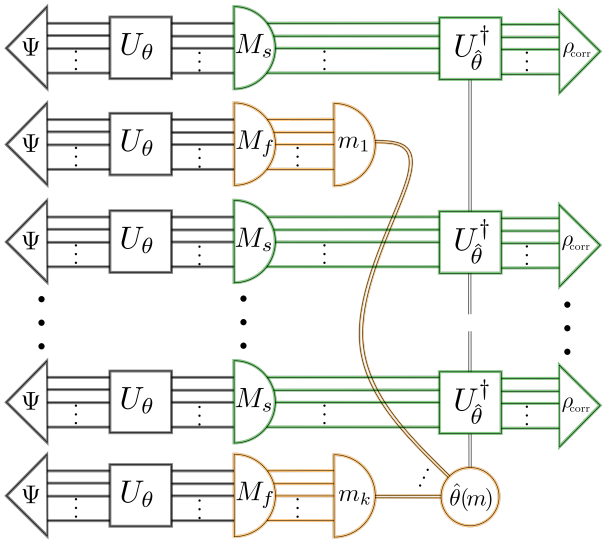

FIG. 1. Schematic of the metrology-assisted entanglement distribution protocol. Each copy of the distributed multi-party state $|\Psi\rangle$ is subject to a (local) noise channel (here represented by an unknown unitary U_θ) parameterized by θ, which is subsequently measured by an OSBP with measurement operators M_s and M_f, corresponding to successful (green) and failed (orange) state conversion, respectively. The copies in the failure branch can be used to perform further measurements with outcomes m_i to estimate the noise channel, and apply corresponding corrections (represented here by $U_{\hat{\theta}}^\dagger$) to obtain partially corrected states ρ_{corr}.

More specifically, let us assume the vendor produces the pure two-parameter qutrit state

$$|\Psi(\alpha,\beta)\rangle = \sin\alpha\cos\beta\,|0\rangle^{\otimes N} + \sin\alpha\sin\beta\,|1\rangle^{\otimes N} + \cos\alpha\,|2\rangle^{\otimes N}, \tag{7}$$

where $\tan^{-1}(\csc\beta) \leq \alpha \leq \pi/2$ and $0 \leq \beta \leq \pi/4$. This choice of angles ensures that $\sin\alpha\cos\beta \geq \sin\alpha\sin\beta \geq \cos\alpha \geq 0$. The parties wish to obtain the maximally entangled state of Eq. (6), for which the optimal OSBP comprises measurement operators [1]

$$M_\text{s} = \mathrm{diag}(\cot\alpha\sec\beta, \cot\alpha\csc\beta, 1), \tag{8}$$
$$M_\text{f} = \mathrm{diag}(\sqrt{1-\cot^2\alpha\sec^2\beta}, \sqrt{1-\cot^2\alpha\csc^2\beta}, 0)$$

for one party and the identity for the remaining parties. The conversion is successful with probability $p_\text{s} = 3\cos^2\alpha$, whereas with probability $p_\text{f} = 1 - 3\cos^2\alpha$ the conversion protocol fails and results in the state

$$|\Phi_\text{f}\rangle = a\,|0\rangle^{\otimes N} + \sqrt{1-a^2}\,|1\rangle^{\otimes N}, \tag{9}$$

[1] Although the optimal OSBP given here is not generally optimal, it is easy to check that this is indeed the case here. By splitting the state in Eq. (7) into one party versus the remaining $n-1$ parties and comparing the Schmidt coefficients, one immediately sees that the conversion rate for a single state is bounded by $3\cos^2\alpha$. Even if the OSBP is not optimal, the following ideas generalize to the optimal protocol.

where

$$a = \frac{\sqrt{\sin^2\alpha\cos^2\beta - \cos^2\alpha}}{\sqrt{1-3\cos^2\alpha}}. \tag{10}$$

Observing that $[M_x, U(\theta)] = 0, \forall\, x \in \{\mathrm{s}, \mathrm{f}\}$, it follows that the resulting states for the successful and failure branches of the OSBP are given by

$$|\,\mathrm{GHZ}_3(\theta)\,\rangle = \sqrt{\frac{1}{3}}\sum_{m=0}^{2} e^{\mathrm{i}mN\theta}\,|\,m\,\rangle^{\otimes N}, \tag{11a}$$

$$|\,\Phi_\mathrm{f}(\theta)\,\rangle = a\,|\,0\,\rangle^{\otimes N} + \sqrt{1-a^2}e^{\mathrm{i}N\theta}\,|\,1\,\rangle^{\otimes N}, \tag{11b}$$

respectively, with the corresponding probabilities invariant under the presence of dephasing noise.

In the case that the OSBP is successful, the parties obtain the state

$$\Lambda\big[|\,\mathrm{GHZ}_3\,\rangle\langle\,\mathrm{GHZ}_3\,|\big] = \int_0^{2\pi} p(\theta)U(\theta)^{\otimes N} |\,\mathrm{GHZ}_3\,\rangle\langle\,\mathrm{GHZ}_3\,|\,U(\theta)^{\dagger\otimes N}\mathrm{d}\theta, \tag{12}$$

where $p(\theta)$ is a probability distribution over $\theta \in [0, 2\pi)$ describing the believe/knowledge the parties have of[2] θ. Loosely speaking, the more peaked the distribution becomes, the closer the state gets to the target state. The goal of the N-parties now is to apply correction operations to the successfully converted copies to obtain the highest yield of N-partite highly entangled qutrit states ρ_corr, that are ϵ-close in trace distance to the ideal state $|\,\mathrm{GHZ}_3\,\rangle$.

In the case of a flat prior, $p(\theta) = (2\pi)^{-1}$, we can analytically calculate the average trace distance to the target state after the first measurement to be (see Appendix A for a detailed calculation)

$$\bar{D}(\rho_\mathrm{corr}, \rho_\mathrm{target}) = \tfrac{1}{6}\big(1 + \sqrt{9 + 8a^2(1-a^2) - 16a\sqrt{1-a^2}}\big). \tag{13}$$

The average distance after one measurement hence only depends on the parameter a. Since $a \in [0, 1]$, the maximal value $\frac{2}{3}$ is obtained for $a = 0$ or $a = 1$. This coincides with the distance based on the prior, i.e., ignoring measurement data. The minimal value of $\bar{D} = \frac{1}{6}(1+\sqrt{3})$ is obtained for $a = 1/\sqrt{2}$.

Given a particular measurement outcome the posterior distribution is updated using Bayes' law, [Eq. (3)]. In general the parties can also update their measurement directions at every step but this amount of generality quickly renders the problem analytically intractable. We thus resort to numerical analysis henceforth. The simplest strategy of fixed local measurements is treated in Appendix B.

In Fig. 2 we see how the trace distance $\bar{D}$ of the corrected state resulting from the successful branch to the target state reduces with the number of failure branches used for the estimation. After 20 measurement rounds the average trace distance drops to $\bar{D} \approx 0.038$ for $a = 1/\sqrt{2}$. This is a remarkable improvement considering that the failure states would have been discarded in the naive protocol.

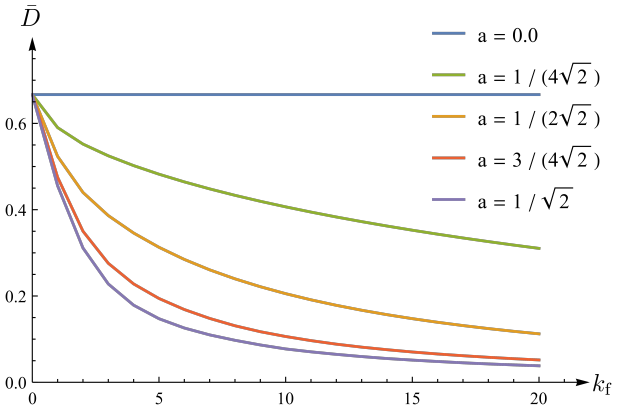

FIG. 2. Average distance to the target state. The curves show the average trace distance between the target state and the successfully converted copies for the exemplary metrology-assisted entanglement distribution protocol discussed in Sec. III for different values of a as a function of the number k_f of copies in the failure branch.

One could also think of more sophisticated strategies, where a swap operation is applied to the state by the vendor before it is distributed, changing the basis vectors $|\,1\,\rangle \leftrightarrow |\,2\,\rangle$. Since the state $|\,2\,\rangle$ reacts more sensitive to the transformation, i.e., picks up a phase of $e^{2\mathrm{i}N\theta}$, this can be advantageous for estimation. Note, however, that in this case one can draw conclusions about θ only up to a π/N phase shift, whereas the successful branch still depends on $N\theta$. Thus, this strategy can only be employed once the interval of θ is sufficiently narrowed down.

IV. COMPARISON WITH 'NAIVE' PROTOCOLS

In this section we compare the performance of our strategy to what we call 'naive' strategies. These are strategies that first use a number of copies for the estimation of the LU-channel and subsequently perform the transformation to the target state for all following copies. In contrast, our protocol *sequentially* processes each multipartite state. In addition, our protocol only uses local measurements both for the estimation and the state transformation. Although this drastically simplifies potential practical implementations, it is known that processing multiple copies simultaneously —via the use of quantum memories— can lead to a higher success rate for entanglement distillation [48, 49]. For a fair comparison we therefore first impose the same restrictions on the 'naive' strategy. Then we consider the case where each party can use quantum memories and is able to process multiple copies at the same time. In both scenarios we compare our previously obtained result for metrology-assisted entanglement distribution to asymptotic bounds valid for any distillation strategy.

Let us, assume we want to perform the task of the previous example, i.e., k copies of the state in Eq. (7) are distributed

[2] Actually the estimation only gives us information about $N\theta$, which allows us to draw conclusions about θ only up to a shift by $2\pi/N$. But this is not a problem, since the final state also only depends on $N\theta$.

via a local dephasing channel, described in Eq. (5), and the goal of the parties is to obtain the highest number of states ϵ-close to $|\,\mathrm{GHZ}_3\,\rangle$. They therefore use k_e copies of $|\,\Psi(\alpha,\beta)\,\rangle$ to estimate the phase θ and assign an estimator $\hat{\theta}(\mathbf{m})$ based on the measurement outcome $\mathbf{m} = (m_1, \ldots, m_{k_\mathrm{e}})$. The variance

$$V[\hat{\theta}(\mathbf{m})] = \sum_{\mathbf{m}} p(\mathbf{m}|\theta)(\hat{\theta}(\mathbf{m}) - \theta)^2 \tag{14}$$

of any *unbiased* estimator $\hat{\theta}(\mathbf{m})$ is bounded from below by the quantum Cramér-Rao bound, i.e., it is larger than the inverse quantum Fisher information of the state [50, 51]. For a single copy of the state in Eq. (7) the Fisher information is

$$\begin{aligned}\mathcal{I}\big[|\,\psi(\theta)\,\rangle\big] = \; & 4N^2[\sin^2\alpha\sin^2\beta + 4\cos^2\alpha \\ & - (\sin^2\alpha\sin^2\beta + 2\cos^2\alpha)^2],\end{aligned} \tag{15}$$

and the FI is additive for uncorrelated copies of the state. For repeated measurements the outcomes m_i are independent and identically distributed and the distribution of $\hat{\theta}$ over $\mathbb{R}$ is described by a Gaussian distribution with mean θ, which without loss of generality we set equal to 0, and variance $\sigma^2 = 1/(k_\mathrm{e}\,\mathcal{I}\big[|\,\psi(\theta)\,\rangle\big])$. As we are only interested in θ up to a phase of 2π, we represent the distribution of $\hat{\theta}$ as a wrapped Gaussian. We assume that the estimation is sufficiently precise such that the difference between the two is negligible, i.e., for values $\sigma < \pi/2$ a Gaussian distribution well-approximates a wrapped Gaussian (see the discussion in [52, A.II.5]).

With their knowledge of θ, the parties then assign the corrected state ρ_corr given by

$$\begin{aligned}\rho_\mathrm{corr} &= \int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{\hat{\theta}^2}{\sigma^2}}}{\sqrt{2\pi}\sigma}\frac{1}{3}\sum_{j,k=0}^{2}\mathrm{e}^{i(j-k)N\hat{\theta}}(|j\rangle\langle k|)^{\otimes N}d\hat{\theta} \\ &= \frac{1}{3}\sum_{j,k=0}^{2}\mathrm{e}^{-\frac{(j-k)^2N^2\sigma^2}{2}}(|j\rangle\langle k|)^{\otimes N},\end{aligned} \tag{16}$$

and the trace distance to the target state is

$$\begin{aligned}\bar{D}(\rho_\mathrm{corr},\rho_\mathrm{target}) = \tfrac{1}{6}\mathrm{e}^{-2N^2\sigma^2}\big(\mathrm{e}^{\frac{N^2\sigma^2}{2}}-1\big)\Big[1+\mathrm{e}^{\frac{N^2\sigma^2}{2}}+\mathrm{e}^{N^2\sigma^2} \\ + \mathrm{e}^{3N^2\sigma^2/2} + \sqrt{8\mathrm{e}^{3N^2\sigma^2} + (1+\mathrm{e}^{N^2\sigma^2/2})(1+\mathrm{e}^{N^2\sigma^2})^2}\Big].\end{aligned} \tag{17}$$

Inserting $\sigma^2 = 1/(k_\mathrm{e}\,\mathcal{I}\big[|\,\psi(\theta)\,\rangle\big])$ leaves us with an expression depending only on the number k_e of copies used for estimation but not on the number N of parties. Let us compare this to the numerical results of the previously employed adaptive strategy for the exemplary parameters $k = 100$, $\alpha = \arccos(2/\sqrt{15})$ and $\beta = \pi/4$. This leads to $p_\mathrm{s} = 4/5$, and so to $k_\mathrm{s} = 80$ and $k_\mathrm{f} = 20$ successful and failure branches, respectively, on average. After 20 rounds of estimation with the failure branch the adaptive strategy leads to a trace distance of $\epsilon \approx 0.038$, see Fig. 2. Using the FI of Eq. (15) one finds that achieving a comparable trace distance, any measurement performed needs at least $k_\mathrm{e} \geq 8$ copies of the received state. It should not surprise, that this number is lower than what is needed in the Bayesian approach above. After all, a different state, i.e. the received state instead of the failure branch, is used for the estimation. One can easily see that the FI is larger for the received state, as information about the parameter θ can only be lost in the OSBP. This means that the received state will generally be more advantageous for the estimation than the failure branch. In addition, we note that $k_\mathrm{e} \geq 8$ is only a lower bound, and it is not clear that a strategy exists that saturates this bound for finitely many copies.

Once the phase θ has been estimated sufficiently well, the states resulting from the successful branches can be transformed via SLOCC. We have already seen that any sequential strategy transforming the state in Eq. (7) to the state $|\,\mathrm{GHZ}_3\,\rangle$ is bounded by $p_\mathrm{s} \leq 3\cos^2\alpha$, compare footnote 1. In this example, this leads to obtaining at most 73.6 copies of the target state on average from the initial 100 copies of the distributed state. Protocols that act on all available states at once can perform better than that. For pure bipartite states the asymptotic distillation rate from $|\,\psi\,\rangle$ to $|\,\phi\,\rangle$ is bounded by the entropy ratio of the reduced states for any party i, i.e.,

$$\mathcal{D}(|\,\psi\,\rangle \to |\,\phi\,\rangle) = \frac{S(|\psi\rangle\langle\psi|_i)}{S(|\phi\rangle\langle\phi|_i)}, \tag{18}$$

and this bound is tight [53]. For multipartite states this immediately gives the, generally not tight, asymptotic bound

$$\mathcal{D}(|\,\psi\,\rangle \to |\,\phi\,\rangle) = \min_i\Big(\frac{S(|\psi\rangle\langle\psi|_i)}{S(|\phi\rangle\langle\phi|_i)}\Big). \tag{19}$$

In our case, however, the latter bound is tight [54, 55] and we can calculate

$$\begin{aligned}\mathcal{D}(|\,\Psi(\alpha,\beta)\,\rangle \to |\,\mathrm{GHZ}\,\rangle) = & -\cos^2\alpha\log_3(\cos^2\alpha) \\ & - \sin^2\alpha\cos^2\beta\log_3(\sin^2\alpha\cos^2\beta) \\ & - \sin^2\alpha\sin^2\beta\log_3(\sin^2\alpha\sin^2\beta).\end{aligned} \tag{20}$$

For our example this means there might exist a distillation protocol acting on all available copies at once able to successfully distill 91 copies of the desired GHZ-state, which clearly exceeds the 80 successful states obtained by our adaptive strategy. But these 91 copies, in turn, are only an upper bound, and it is not clear that a strategy attaining it for finitely many copies exists. Indeed, even if such a strategy exists, it would generally require acting on all remaining copies simultaneously. Even if the presented metrology-assisted protocol does not saturate these 'optimal' bounds, it nonetheless shows the potential of practically easily implementable adaptive strategies acting on single copies. In particular, the task of designing and implementing optimal strategies for estimation and especially distillation becomes extremely challenging for larger numbers of copies and may be out of reach for experimental realizations in the medium term. Therefore it is essential to find feasible strategies that perform well and whose implementation remains tractable with growing number of copies.

V. METROLOGY-ASSISTED PROTOCOLS BEYOND DEPHASING NOISE

A key assumption in our previous example was the commutativity of the measurement operators of the OSBP and the Kraus operators of the noisy channel. If this does not hold, as is the case in depolarizing, or lossy channels, the order in which the POVM measurements of the OSBP and those for the estimation protocol are carried out matter. If the channel does not commute with the POVM operators of the OSBP, this impacts the performance of the presented strategy. Nonetheless one could design strategies relying on the failure branch for the channel estimation. We now present a protocol that integrates the estimation of general unknown unitary transformations $\mathrm{SU}(d)$ into an OSBP. Possible strategies to estimate a general unitary transformation are given by [56, 57]. We focus on how an estimation strategy can incorporate the failure branch of an OSBP for a potentially more efficient estimation, without restricting ourselves to a specific strategy.

The naive way to perform this task would again be to first use a number of copies to estimate the unknown unitary sufficiently well, such that the received state can be corrected up to within trace distance $\bar{D} \leq \epsilon$ of the target. After that, some distillation procedure is applied to recover the desired state.

In our approach we combine both tasks in to one. The idea is to construct the POVM operators for the OSBP based on the knowledge of the unknown unitary at that specific stage of the protocol. We start with the POVM $\{\mathbb{1}, 0\}$, that will always result in the success branch. Then we check if the state is sufficiently close to the target state, based on our knowledge of the unitary. This will not be the case if the distribution $p(U)$ is sufficiently broad. If the state is not close enough to the target, we make a measurement and update our believe and the estimator. With increasing knowledge of the unitary transformation, also the POVM of the OSBP will change and converge towards the optimal POVM for the received state. During this process the copies in the successful branch close to the target (according to the knowledge at the time) are stored, while all the others are used for estimation. With an increasingly peaked distribution $p(U)$ also the ratio of stored copies to discarded copies (used for estimation) will increase. Given that enough copies were purchased, we can assume that the unitary transformation is known sufficiently well at the end of this protocol. For all stored copies it is checked if they are indeed close enough to the target according to the updated knowledge of U. If not, they are transformed into the target state via SLOCC. The exact estimation and updating procedure are left unspecified here on purpose because these will strongly depend on the exact state and noise model. In Algorithm 1 we consider the exemplary situation where the desired state is an N-partite GHZ-state of local dimension d and the distributed state is of the form of Eq. (4) with corresponding dimension and party number. But the idea works for general scenarios and Algorithm 1 can be adapted for broader classes of vendor and target states.

Once the first round is completed, all states that were saved, i.e., the states with a successful OSBP that were sufficiently close to the target state (with the information at the time), are checked to be sufficiently close with the final distribution. If not, one may try to transform them into the target via unitary transformations and an OSBP. The eventual failure branch copies can be used for estimation.

For the first rounds this protocol will do almost the same as the naive protocol. Since the distribution $p(U)$ is broad, the POVM will be close to $\{\mathbb{1}, 0\}$, resulting most likely in the success branch (not altering the state). But as the state will be far from the target, it will be used for estimation. After many rounds the distribution $p(U)$ will be very peaked, all copies in the successfull branch are now close to the target state and

Algorithm 1: Metrology-integrated OSBP

input : k copies of the state $|\psi\rangle$
output: $p_{(k)}(U)$, $\hat{U}_{(k)}$ and all saved copies $|\tilde{\psi}_s^{(i)}\rangle$
begin
  start with uniform distribution $p_{(0)}(U)$ over $SU(d)$
  set estimator $\hat{U}_{(0)} = \mathbb{1}$
  $i = 1$
  while $i \leq k$, *i.e., there are copies left* **do**
    send copy number i of the state $|\psi\rangle$
    obtain $U^{\otimes N}|\psi\rangle$
    correct to $|\tilde{\psi}^{(i)}\rangle = (\hat{U}_{(i-1)}^{\dagger} U)^{\otimes N}|\psi\rangle$
    construct the POVM for the OSBP:
      - define M_{s}^{V} as the positive matrix such that $M_{\mathrm{s}}^{V} V^{\otimes N} |\psi\rangle = |\mathrm{GHZ}_d\rangle$
      - $M_{\mathrm{s}}^{(i)} = \int_{SU(d)} p_{(i-1)}(V) {M_{\mathrm{s}}^{\hat{U}_{(i-1)}^{\dagger} V}} |\psi\rangle \, \mathrm{d}V$
      - $M_{\mathrm{f}}^{(i)}$ such that $(M_{\mathrm{s}}^{(i)})^{\dagger} M_{\mathrm{s}}^{(i)} + (M_{\mathrm{f}}^{(i)})^{\dagger} M_{\mathrm{f}}^{(i)} = \mathbb{1}$
    perform OSBP with $\{M_{\mathrm{s}}^{(i)}, M_{\mathrm{f}}^{(i)}\}$
    if *success* **then**
      get $|\tilde{\psi}_s^{(i)}\rangle = M_{\mathrm{s}}^{(i)}|\tilde{\psi}^{(i)}\rangle$
      if $\int_{SU(d)} p_{(i-1)}(U) |\langle \mathrm{GHZ}_d | M_{\mathrm{s}}^{(i)} (\hat{U}_{(i-1)}^{\dagger} U)^{\otimes N} |\psi\rangle|^2 \mathrm{d}U \geq 1 - \epsilon$
      then
        keep $|\tilde{\psi}_s^{(i)}\rangle$ and save corresponding $M_{\mathrm{s}}^{(i)}$ and $\hat{U}_{(i-1)}$
        leave probability distribution and estimator of the unitary unchanged
        $p_{(i)}(U) = p_{(i-1)}(U)$
        $\hat{U}_{(i)} = \hat{U}_{(i-1)}$
      else
        perform a measurement on $|\tilde{\psi}_s^{(i)}\rangle$ to estimate U
        update distribution $p_{(i)}(U)$ and estimator $\hat{U}_{(i)}$
      end
    else
      perform a measurement on $|\tilde{\psi}_f^{(i)}\rangle = M_{\mathrm{f}}^{(i)}|\tilde{\psi}^{(i)}\rangle$ to estimate U
      update distribution $p_{(i)}(U)$ and estimator $\hat{U}_{(i)}$
    end
    $i++$
  end
  return $p_{(k)}(U)$, $\hat{U}_{(k)}$ and all saved copies $|\tilde{\psi}_s^{(i)}\rangle$
end

only the failure branches are used for estimation, resulting in a protocol similar to the one presented above.

VI. CONCLUSION

Potential future implementations of large-scale quantum networks, such as the envisioned 'quantum internet' [13–15], can be expected to merge idealized protocols for quantum communication and information processing with real-world conditions arising from noise and technological constraints which will necessitate the development of pragmatic and flexible solutions. Here, we have considered a quantum communication protocol that will be relevant in this context: the distribution and conversion of high-dimensional entangled states among multiple parties using noisy channels. Procedures for quantum state-conversion employed under these circumstances will involve probabilistic transformations via SLOCC. As we have shown, such scenarios are naturally amenable to enhancements via embedded channel estimation routines.

While the ideas presented here apply quite generally, we have illustrated the improvements one may expect in such metrology-assisted entanglement distribution by focusing on a particular example: probabilistic conversion of finitely many copies of N-qutrit states distributed via a local dephasing channels to GHZ target states. For this example, we have shown that already practically easily implementable Bayesian estimation strategies based on local measurements of individual copies can lead to excellent performance comparable to optimal distillation procedures. The latter, however, might involve significantly more complex and technologically challenging joint operations on multiple copies.

We believe that metrology-assisted entanglement distribution has the potential to become a relevant factor in future quantum networks. We thus expect the study and optimization of such protocols for different target states and noisy models, as well as their comparison to relevant alternative protocols (e.g., based on distillation) to open up a broad range of interesting questions beyond this initial pilot study.

ACKNOWLEDGMENTS

We thank Cornelia Spee for feedback on the manuscript. S.M. acknowledges support from the Austrian Science Fund (FWF) through projects Y879-N27 (START) and P 31339-N27 (Stand-Alone). N.F. acknowledges support from the Austrian Science Fund (FWF) through the projects Y879-N27 (START) and P 31339-N27 (Stand-Alone), and the joint Czech-Austrian project MultiQUEST (I 3053-N27 and GF17-33780L). M.S. ackowledges support from Spanish MINECO reference FIS2016-80681-P (with the support of AEI/FEDER,EU), the Catalan Government for the project QuantumCAT 001-P-001644 (RIS3CAT comunitats) co-financed by the European Regional Development Fund (FEDER), and the Generalitat de Catalunya, project CIRIT 2017-SGR-1127.

[1] Andrew J. Scott, *Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions,* Phys. Rev. A **69**, 052330 (2004), arXiv:quant-ph/0310137.

[2] Tohya Hiroshima and Masahito Hayashi, *Entanglement and Quantum Error Correction,* in *Quantum Computation and Information*, edited by Hiroshi Imai and Masahito Hayashi (Springer, Berlin, Heidelberg, 2006) Chap. 5, pp. 111–132.

[3] Wolfgang Dür and Hans J. Briegel, *Entanglement purification and quantum error correction,* Rep. Prog. Phys. **70**, 1381 (2007), arXiv:0705.4165.

[4] Todd A. Brun and Min-Hsiu Hsieh, *Entanglement-assisted quantum error-correcting codes,* in *Quantum Error Correction*, edited by Daniel A. Lidar and Todd A. Brun (Cambridge University Press, Cambridge, U.K., 2013) Chap. 7, pp. 181–200, arXiv:1610.04013.

[5] Oliver Marty, Marcus Cramer, and Martin B. Plenio, *Practical Entanglement Estimation for Spin-System Quantum Simulators,* Phys. Rev. Lett. **116**, 105301 (2016), arXiv:1504.03572.

[6] Nicolai Friis, Oliver Marty, Christine Maier, Cornelius Hempel, Milan Holzäpfel, Petar Jurcevic, Martin B. Plenio, Marcus Huber, Christian Roos, Rainer Blatt, and Ben Lanyon, *Observation of Entangled States of a Fully Controlled 20-Qubit System,* Phys. Rev. X **8**, 021012 (2018), arXiv:1711.11092.

[7] Marcello Dalmonte, Benoît Vermersch, and Peter Zoller, *Quantum simulation and spectroscopy of entanglement Hamiltonians,* Nat. Phys. **14**, 827 (2018), arXiv:1707.04455.

[8] Michael Epping, Hermann Kampermann, Chiara Macchiavello, and Dagmar Bruß, *Multi-partite entanglement can speed up quantum key distribution in networks,* New J. Phys. **19**, 093012 (2017), arXiv:1612.05585.

[9] Stefan Bäuml and Koji Azuma, *Fundamental limitation on quantum broadcast networks,* Quantum Sci. Technol. **2**, 024004 (2017), arXiv:1609.03994.

[10] Matej Pivoluska, Marcus Huber, and Mehul Malik, *Layered quantum key distribution,* Phys. Rev. A **97**, 032312 (2018), arXiv:1709.00377.

[11] Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner, *Fully device-independent conference key agreement,* Phys. Rev. A **97**, 022307 (2018), arXiv:1708.00798.

[12] Matteo Pompili, Sophie L. N. Hermans, Simon Baier, Hans K. C. Beukers, Peter C. Humphreys, Raymond N. Schouten, Raymond F. L. Vermeulen, Marijn J. Tiggelman, Laura dos Santos Martins, Bas Dirkse, Stephanie Wehner, and Ronald Hanson, *Realization of a multinode quantum network of remote solid-state qubits,* Science **372**, 259 (2021), arXiv:2102.04471.

[13] H. Jeff Kimble, *The quantum internet,* Nature **453**, 1023 (2008), arXiv:0806.4195.

[14] Stephanie Wehner, David Elkouss, and Ronald Hanson, *Quantum internet: A vision for the road ahead,* Science **362**, eaam9288 (2018).

[15] Angela Sara Cacciapuoti, Marcello Caleffi, Francesco Tafuri, Francesco Saverio Cataliotti, Stefano Gherardini, and Giuseppe Bianchi, *Quantum Internet: Networking Challenges in Distributed Quantum Computing,* IEEE Network **34**, 137 (2020), arXiv:1810.08421.

[16] Wolfgang Dür, Raphael Lamprecht, and Stefan Heusler, *To-*

wards a quantum internet, Eur. J. Phys. **38**, 043001 (2017).

[17] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, *Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels,* Phys. Rev. Lett. **70**, 1895–1899 (1993).

[18] Dik Bouwmeester, Jian-Wei Pan, Klaus Mattle, Manfred Eibl, Harald Weinfurter, and Anton Zeilinger, *Experimental Quantum Teleportation,* Nature **390**, 575–579 (1997), arXiv:1901.11004.

[19] Daniel Llewellyn, Yunhong Ding, Imad I. Faruque, Stefano Paesani, Davide Bacco, Raffaele Santagati, Yan-Jun Qian, Yan Li, Yun-Feng Xiao, Marcus Huber, Mehul Malik, Gary F. Sinclair, Xiaoqi Zhou, Karsten Rottwitt, Jeremy L. O'Brien, John G. Rarity, Qihuang Gong, Leif K. Oxenlowe, Jianwei Wang, and Mark G. Thompson, *Chip-to-chip quantum teleportation and multi-photon entanglement in silicon,* Nat. Phys. **16**, 148 (2020), arXiv:1911.07839.

[20] Wolfgang Dür, Guifré Vidal, and Juan Ignacio Cirac, *Three qubits can be entangled in two inequivalent ways,* Phys. Rev. A **62**, 062314 (2000), arXiv:quant-ph/0005115.

[21] Antonio Acín, Dagmar Bruß, Maciej Lewenstein, and Anna Sanpera, *Classification of Mixed Three-Qubit States,* Phys. Rev. Lett. **87**, 040401 (2001), arXiv:quant-ph/0103025.

[22] Julio I. de Vicente, Cornelia Spee, and Barbara Kraus, *Maximally Entangled Set of Multipartite Quantum States,* Phys. Rev. Lett. **111**, 110502 (2013), arXiv:1305.7398.

[23] Katharina Schwaiger, David Sauerwein, Martí Cuquet, Julio I. de Vicente, and Barbara Kraus, *Operational Multipartite Entanglement Measures,* Phys. Rev. Lett. **115**, 150502 (2015), arXiv:1503.00615.

[24] Julio I. de Vicente, Cornelia Spee, David Sauerwein, and Barbara Kraus, *Entanglement manipulation of multipartite pure states with finite rounds of classical communication,* Phys. Rev. A **95**, 012323 (2017), arXiv:1607.05145.

[25] Cornelia Spee, Julio I. de Vicente, David Sauerwein, and Barbara Kraus, *Entangled Pure State Transformations via Local Operations Assisted by Finitely Many Rounds of Classical Communication,* Phys. Rev. Lett. **118**, 040503 (2017), arXiv:1606.04418.

[26] David Sauerwein, Nolan R. Wallach, Gilad Gour, and Barbara Kraus, *Transformations among Pure Multipartite Entangled States via Local Operations are Almost Never Possible,* Phys. Rev. X **8**, 031020 (2018), arXiv:1711.11056.

[27] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter, *Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask),* Commun. Math. Phys. **328**, 303 (2014), arXiv:1210.4583.

[28] Jessica Bavaresco, Natalia Herrera Valencia, Claude Klöckl, Matej Pivoluska, Paul Erker, Nicolai Friis, Mehul Malik, and Marcus Huber, *Measurements in two bases are sufficient for certifying high-dimensional entanglement,* Nat. Phys. **14**, 1032 (2018), arXiv:1709.07344.

[29] He Lu, Qi Zhao, Zheng-Da Li, Xu-Fei Yin, Xiao Yuan, Jui-Chen Hung, Luo-Kan Chen, Li Li, Nai-Le Liu, Cheng-Zhi Peng, Yeong-Cherng Liang, Xiongfeng Ma, Yu-Ao Chen, and Jian-Wei Pan, *Entanglement Structure: Entanglement Partitioning in Multipartite Systems and Its Experimental Detection Using Optimizable Witnesses,* Phys. Rev. X **8**, 021072 (2018), arXiv:1711.01784.

[30] Nicolai Friis, Giuseppe Vitagliano, Mehul Malik, and Marcus Huber, *Entanglement Certification From Theory to Experiment,* Nat. Rev. Phys. **1**, 72 (2019), arXiv:1906.10929.

[31] You Zhou, Qi Zhao, Xiao Yuan, and Xiongfeng Ma, *Detecting multipartite entanglement structure with minimal resources,* npj Quantum Inf. **5**, 83 (2019), arXiv:1904.05001.

[32] Natalia Herrera Valencia, Vatshal Srivastav, Matej Pivoluska, Marcus Huber, Nicolai Friis, Will McCutcheon, and Mehul Malik, *High-Dimensional Pixel Entanglement: Efficient Generation and Certification,* Quantum **4**, 376 (2020), arXiv:2004.04994.

[33] Gary J. Mooney, Gregory A. L. White, Charles D. Hill, and Lloyd C. L. Hollenberg, *Whole-Device Entanglement in a 65-Qubit Superconducting Quantum Computer,* Adv. Quantum Technol. , 2100061 (2021), arXiv:2102.11521.

[34] Raju Valivarthi, Samantha Davis, Cristian Pena, Si Xie, Nikolai Lauk, Lautaro Narvaez, Jason P. Allmaras, Andrew D. Beyer, Yewon Gim, Meraj Hussein, George Iskander, Hyunseong Linus Kim, Boris Korzh, Andrew Mueller, Mandy Rominsky, Matthew Shaw, Dawn Tang, Emma E. Wollman, Christoph Simon, Panagiotis Spentzouris, Neil Sinclair, Daniel Oblak, and Maria Spiropulu, *Teleportation Systems Towards a Quantum Internet,* PRX Quantum **1**, 020317 (2020), arXiv:2007.11157.

[35] Miguel Navascues, Elie Wolfe, Denis Rosset, and Alejandro Pozas-Kerstjens, *Genuine Network Multipartite Entanglement,* Phys. Rev. Lett. **125**, 240505 (2020), arXiv:2002.02773.

[36] Tristan Kraft, Sébastien Designolle, Christina Ritz, Nicolas Brunner, Otfried Gühne, and Marcus Huber, *Quantum entanglement in the triangle network,* Phys. Rev. A **103**, L060401 (2021), arXiv:2002.03970.

[37] Tristan Kraft, Cornelia Spee, Xiao-Dong Yu, and Otfried Gühne, *Characterizing quantum networks: Insights from coherence theory,* Phys. Rev. A **103**, 052405 (2021), arXiv:2006.06693.

[38] Johan Åberg, Ranieri Nery, Cristhiano Duarte, and Rafael Chaves, *Semidefinite tests for quantum network topologies,* Phys. Rev. Lett. **125**, 110505 (2020), arXiv:2002.05801.

[39] Cornelia Spee and Tristan Kraft, *Transformations in quantum networks via local operations assisted by finitely many rounds of classical communication,* arXiv:2105.01090 [quant-ph] (2021).

[40] Hayata Yamasaki, Simon Morelli, Markus Miethlinger, Jessica Bavaresco, Nicolai Friis, and Marcus Huber, *Activation of genuine multipartite entanglement: beyond the single-copy paradigm of entanglement characterisation,* arXiv:2106.01372 [quant-ph] (2021).

[41] Antoine Neven, David Gunn, Martin Hebenstreit, and Barbara Kraus, *Local Transformations of Multiple Multipartite States,* SciPost Phys. **11**, 42 (2021), arXiv:2007.06256.

[42] Gilad Gour, Barbara Kraus, and Nolan R. Wallach, *Almost all multipartite qubit quantum states have trivial stabilizer,* J. Math. Phys. **58**, 092204 (2017), arXiv:1609.01327.

[43] Antonio Acín, Enric Jané, Wolfgang Dür, and Guifre Vidal, *Optimal Distillation of a Greenberger-Horne-Zeilinger State,* Phys. Rev. Lett. **85**, 4811 (2000), arXiv:quant-ph/0007042.

[44] David Sauerwein, Katharina Schwaiger, and Barbara Kraus, *Discrete and differentiable entanglement transformations,* arXiv:1808.02819 [quant-ph] (2018).

[45] Marcin Jarzyna, Konrad Banaszek, and Rafał Demkowicz-Dobrzański, *Dephasing in coherent communication with weak signal states,* J. Phys. A: Math. Theor. **47**, 275302 (2014), arXiv:1307.6871.

[46] Keith H. Wanser, *Fundamental phase noise limit in optical fibres due to temperature fluctuations,* Electron. Lett. **28**, 53 (1992).

[47] Marco Fanizza, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Vittorio Giovannetti, *Squeezing-enhanced communication without a phase reference,* arXiv:2006.06522

[quant-ph] (2020).
[48] Nilanjana Datta and Felix Leditzky, *Second-Order Asymptotics for Source Coding, Dense Coding, and Pure-State Entanglement Conversions,* IEEE Trans. Inf. Theory **61**, 582 (2015), arXiv:1403.2543.
[49] Kun Fang, Xin Wang, Marco Tomamichel, and Runyao Duan, *Non-Asymptotic Entanglement Distillation,* IEEE Trans. Inf. Theory **65**, 6454 (2019), arXiv:1706.06221.
[50] Rafał Demkowicz-Dobrzański, Marcin Jarzyna, and Janek Kołodyński, *Quantum limits in optical interferometry,* Prog. Optics **60**, 345 (2015), arXiv:1405.7703.
[51] Samuel L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states,* Phys. Rev. Lett. **72**, 3439 (1994).
[52] Nicolai Friis, Davide Orsucci, Michalis Skotiniotis, Pavel Sekatski, Vedran Dunjko, Hans J. Briegel, and Wolfgang Dür, *Flexible resources for quantum metrology,* New J. Phys. **19**, 063044 (2017), arXiv:1610.09999.
[53] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher, *Concentrating partial entanglement by local operations,* Phys. Rev. A **53**, 2046 (1996), arXiv:quant-ph/9511030.
[54] John A. Smolin, Frank Verstraete, and Andreas Winter, *Entanglement of assistance and multipartite state distillation,* Physical Review A **72** (2005), arXiv:quant-ph/0505038.
[55] Alexander Streltsov, Clément Meignant, and Jens Eisert, *Rates of Multipartite Entanglement Transformations,* Physical Review Letters **125** (2020), arXiv:1709.09693.
[56] Giulio Chiribella, Giacomo Mauro DâĂŹAriano, Paolo Perinotti, and Massimiliano Federico Sacchi, *Efficient Use of Quantum Resources for the Transmission of a Reference Frame,* Phys. Rev. Lett. **93**, 180503 (2004), arXiv:quant-ph/0405095.
[57] Jonas Kahn, *Fast rate estimation of a unitary operation in SU(d),* Phys. Rev. A **75**, 022326 (2007), arXiv:quant-ph/0603115.

Appendix A: Details on single measurement

We consider an initial measurement where each of the N parties measures in the basis $\{|\pm\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i \pm |1\rangle_i)\}$ for $i = 1, 2, \ldots, N$ of the two-dimensional subspace relevant for the copies in the failure branch. Since the setup is symmetric with respect to the exchange of any of the N parties, only the parity of the overall measurement matters, i.e., if one obtains an even or odd number of "$-$" outcomes. To see this note that the probability to get an outcome "n" corresponding to a possible operator E_n is

$$\begin{aligned}
&\mathrm{Tr}(|\psi_\mathrm{f}(\theta)\rangle\langle\psi_\mathrm{f}(\theta)| E_n) = \langle\psi_\mathrm{f}(\theta)| E_n |\psi_\mathrm{f}(\theta)\rangle \\
&= a^2 \langle 0|^{\otimes N} E_n |0\rangle^{\otimes N} + (1-a^2)\langle 1|^{\otimes N} E_n |1\rangle^{\otimes N} \\
&\quad + a\sqrt{1-a^2}\, \mathrm{e}^{-i\theta N} \langle 1|^{\otimes N} E_n |0\rangle^{\otimes N} \\
&\quad + a\sqrt{1-a^2}\, \mathrm{e}^{i\theta N} \langle 0|^{\otimes N} E_n |1\rangle^{\otimes N},
\end{aligned} \tag{A1}$$

while E_n is one of 2^N products of projectors $|\pm\rangle\langle\pm|$, for instance $E_{+-+--} = |+\rangle\langle+|\otimes|-\rangle\langle-|\otimes|+\rangle\langle+|\otimes|-\rangle\langle-|\otimes|-\rangle\langle-|$. But only the parity of the number of "$-$" projector matters. Out of all 2^N combinations of $|+\rangle\langle+|$ and $|-\rangle\langle-|$ projectors for N qubits, half have an even number of "$-$" projectors, so we have

$$p(\text{even/odd}|\theta) = \tfrac{1}{2} \pm a\sqrt{1-a^2}\cos(N\theta) \tag{A2}$$

Given a prior $p(\theta)$, we then assign a probability

$$\begin{aligned}
p(\text{even/odd}) &= \int_0^{2\pi} p(\theta)\, p(\text{even/odd}|\theta) d\theta \\
&= \tfrac{1}{2} \pm a\sqrt{1-a^2} \int_0^{2\pi} p(\theta)\cos(N\theta) d\theta
\end{aligned} \tag{A3}$$

and for a flat prior $p(\theta) = \frac{1}{2\pi}$ this becomes $p(\text{even/odd}) = \frac{1}{2}$. We now use Bayes' law to obtain a posterior distribution. For a flat prior this becomes

$$p(\theta|\text{even/odd}) = \frac{1}{2\pi}[1 \pm 2a\sqrt{1-a^2}\cos(N\theta)]. \tag{A4}$$

Next, we nominate a suitable estimator for θ. As an estimator we can hence use

$$\hat{\theta}_N^{(m)} = \arg(\bar{z}_N^{(m)}), \tag{A5}$$

with

$$\bar{z}_N^{(m)} = \int_0^{2\pi} p(\theta|m)\, \mathrm{e}^{iN\theta}\, d\theta. \tag{A6}$$

For a flat prior we have

$$\bar{z}_N^{(\text{even/odd})} = a\sqrt{1-a^2}\mathrm{e}^{\frac{1\mp1}{2}i\pi}, \tag{A7}$$

and so

$$\hat{\theta}_N^{(\text{even})} = 0 \qquad \hat{\theta}_N^{(\text{odd})} = \pi \tag{A8}$$

We are then interested in evaluating the trace distance between the updated and corrected success-branch state

$$\rho_\text{corr}^{(m)} = \int_0^{2\pi} p(\theta|m)\, |\psi_\mathrm{s}(\theta - \tfrac{\hat{\theta}_N^{(m)}}{N})\rangle\langle\psi_\mathrm{s}(\theta - \tfrac{\hat{\theta}_N^{(m)}}{N})|, \tag{A9}$$

with

$$|\psi_\mathrm{s}(\theta)\rangle = \frac{1}{\sqrt{3}} \sum_{n=0}^{2} \mathrm{e}^{inN\theta} |n\rangle^{\otimes N} \tag{A10}$$

and the desired target state $|\psi_\mathrm{s}(0)\rangle$. The trace distance is

$$D(\rho_\text{corr}^{(m)}, \rho_\text{target}) = \frac{1}{2}\mathrm{Tr}(\sqrt{(\rho_\text{corr}^{(m)} - \rho_\text{target})^2}, \tag{A11}$$

and we can calculate

$$\rho_\text{corr}^{(m)} - \rho_\text{target} = \frac{1}{3}\sum_{j\neq k} c_{jk}^{(m)} (|j\rangle\langle k|)^{\otimes N}, \tag{A12}$$

with

$$c_{jk}^{(m)} = \int_0^{2\pi} p(\theta|m)\, \mathrm{e}^{i(N\theta-\hat{\theta}_N^{(m)})(j-k)} d\theta - 1. \tag{A13}$$

For real $c_{jk}^{(m)}$ this becomes

$$D(\rho_{\text{corr}}^{(m)}, \rho_{\text{target}}) = \frac{1}{6}(|c_{02}| + \sqrt{8c_{01}^2 + c_{02}^2}). \tag{A14}$$

For the flat prior we find

$$c_{0,1}^{(\text{even/odd})} = c_{1,2}^{(\text{even/odd})} = a\sqrt{1-a^2} - 1, \tag{A15a}$$

$$c_{0,2}^{(\text{even/odd})} = -1, \tag{A15b}$$

and with this the trace distance becomes

$$D(\rho_{\text{corr}}^{(m)}, \rho_{\text{target}}) = \tfrac{1}{6}\left(1+\sqrt{9+8a^2(1-a^2)-16a\sqrt{1-a^2}}\right), \tag{A16}$$

and since $p(\text{even/odd}) = 1/2$ this already matches the average trace distance $\bar{D}(\rho_{\text{corr}}, \rho_{\text{target}}) = \sum_m p(m) D(\rho_{\text{corr}}^{(m)}, \rho_{\text{target}})$.

Appendix B: Multiple copies without updating

Now let us see how measurements on more than one copy, in particular on k copies, can improve this value by considering a simple scenario where the measurement directions are not updated. This means, each copy is locally measured in the basis $\{|\pm\rangle_i\}$ by every party, resulting in k individual even/odd measurement outcomes. Since there is no updating in-between, the particular order of these outcomes is irrelevant and only the overall number of even/odd outcomes matter. We can thus view the entire procedure as one $(k+1)$-outcome measurement with outcomes $m = 0, 1, 2, \ldots, k$, i.e., the number of "even" outcomes. We hence assign the probabilities

$$p(m|\theta) = p(\text{even}|\theta)^m p(\text{odd}|\theta)^{k-m} \binom{k}{m}. \tag{B1}$$

The corresponding unconditional probabilities are

$$p(m) = \frac{1}{2^k}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\int_0^{2\pi} p(\theta)\, \cos^{i+j}(N\theta) d\theta. \tag{B2}$$

For a flat prior we calculate

$$p(m) = \frac{1}{2^k}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\frac{(-1)^i+(-1)^j}{2^{i+j+1}}\binom{i+j}{\frac{i+j}{2}}. \tag{B3}$$

The posterior given outcome m is then

$$p(\theta|m) = \frac{1}{2^k p(m)}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\, p(\theta)\, \cos^{i+j}(N\theta). \tag{B4}$$

For the estimator we have to calculate

$$\bar{z}_{N,k}^{(m)} = \int_0^{2\pi} p(\theta|m) \mathrm{e}^{iN\theta} d\theta. \tag{B5}$$

For the flat prior the estimator becomes the argument of

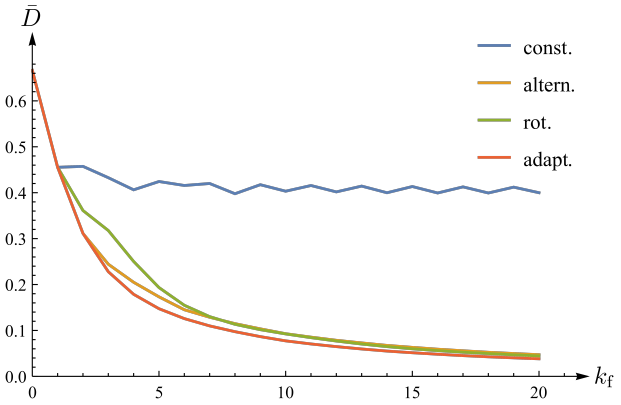

FIG. 3. Average distance to target state for different ways of adapting the measurement direction. The curves show the average trace distance between the target state and the successfully converted copies for different measurement-updating strategies for the exemplary metrology-assisted entanglement distribution protocol discussed in Sec. III for fixed value of $a = \frac{1}{\sqrt{2}}$ as a function of the number k_f of copies in the failure branch. The blue curve represents the strategy with constant measurement directions discussed in Appendix B, for the yellow curve measurements are alternated between σ_x and σ_y, for green the measurement directions are rotated by a constant angle $\pi/8$ after each measurement, and the red curve corresponds to measurement updates using the maximum likelihood estimator from Eq. (A8).

$$\bar{z}_{N,k}^{(m)} = \frac{1}{2^k p(m)}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\frac{(-1)^i+(-1)^j}{2^{i+j+1}}\binom{i+j}{\frac{i+j}{2}}. \tag{B6}$$

We can distinguish two cases, as long as $m < k/2$ we have $\bar{z}_{N,k}^{(m)} < 0$ and therefore $\arg(\bar{z}_{N,k}^{(m)}) = \pi = \hat{\theta}_{N,m}^{(m)}$. For $m \geq k/2$ we have $\bar{z}_{N,k}^{(m)} \geq 0$ and therefore $\arg(\bar{z}_{N,k}^{(m)}) = 0 = \hat{\theta}_{N,m}^{(m)}$. What remains is calculating the distance to our target state, and the result is illustrated in Fig. 3.

9 Activation of genuine multipartite entanglement: Beyond the single-copy paradigm of entanglement characterisation

The following research article is currently under peer review. The manuscript can be accessed online under arXiv:2110.15627.

9.1 Contribution

I participated from the beginning in this project, addressing relevant problems and contributing to the search of appropriate examples. Major parts of the calculations were carried out by me, such as the analytical decomposition in Appendix A and the identification of a suitable witness for the example in Appendix C. I also wrote substantial parts of the appendix.

Activation of genuine multipartite entanglement: Beyond the single-copy paradigm of entanglement characterisation

Hayata Yamasaki[1,2], Simon Morelli[1,2], Markus Miethlinger[1], Jessica Bavaresco[1,2], Nicolai Friis[1,2], and Marcus Huber[2,1]

[1]Institute for Quantum Optics and Quantum Information — IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

[2]Atominstitut, Technische Universität Wien, Stadionallee 2, 1020 Vienna, Austria

Entanglement shared among multiple parties presents complex challenges for the characterisation of different types of entanglement. One of the most fundamental insights is the fact that some mixed states can feature entanglement across every possible cut of a multipartite system yet can be produced via a mixture of states separable with respect to different partitions. To distinguish states that genuinely cannot be produced from mixing such partition-separable states, the term *genuine multipartite entanglement* was coined. All these considerations originate in a paradigm where only a single copy of the state is distributed and locally acted upon. In contrast, advances in quantum technologies prompt the question of how this picture changes when multiple copies of the same state become locally accessible. Here we show that multiple copies unlock genuine multipartite entanglement from partially separable states, i.e., mixtures of the partition-separable states, even from undistillable ensembles, and even more than two copies can be required to observe this effect. With these findings, we characterise the notion of genuine multipartite entanglement in the paradigm of multiple copies and conjecture a strict hierarchy of activatable states and an asymptotic collapse of the hierarchy.

1 Introduction

Entanglement shared among multiple parties is acknowledged as one of the fundamental resources driving the second quantum revolution [1], for instance, as a basis of quantum network proposals [2–5], as a key resource for improved quantum sensing [6] and quantum error correction [7] or as generic ingredient in quantum algorithms [8] and measurement-based quantum computation [9, 10]. Yet, its detection and characterisation are complicated by several factors: among them, the computational hardness of deciding whether any given system even exhibits any entanglement at all [11] as well as the fact that the usual paradigm of local operations and classical communication (LOCC) lead to infinitely many types of entanglement [12–18] already for single copies of multipartite states. Significant effort has thus been devoted to devising practical means of entanglement certification from limited experimental data [19, 20].

One of the principal challenges for the characterisation of multipartite entanglement lies in distinguishing between *partial separability* and its counterpart, *genuine multipartite entanglement* (GME)[1]. Here, a multipartite state is called *partially separable* if it can be decomposed as a mixture of *partition-separable* states, i.e., of states separable with respect to some (potentially different) partitions of the parties into two or more groups, whereas any state that cannot be decomposed in this way has GME (see Fig. 1 and Table 1). One may further classify partially separable states as k-separable states according to the maximal number k of tensor factors that all terms in the partially separable decomposition can be factorised into. If a state admits a decomposition where each term is composed of at least two tensor factors ($k = 2$), the state is called *biseparable*. Thus, every partially separable state is k-separable for some $k \geq 2$, and hence (at least) biseparable. This distinction arises naturally when considering the resources required to create a specific state: any biseparable state can be produced via LOCC in setups where all parties share classical randomness and subsets of parties share entangled states. One of the counter-intuitive features of partially separable states is the possibility for bipartite entanglement across every possible bipartition[2]. Consequently, the notion of bipartite entanglement across partitions is insufficient to capture the notion of

Hayata Yamasaki: hayata.yamasaki@oeaw.ac.at
Simon Morelli: simon.morelli@oeaw.ac.at
Nicolai Friis: nicolai.friis@univie.ac.at
Marcus Huber: marcus.huber@univie.ac.at

[1]Note that the term was also coined for multipartite pure states with exclusively non-vanishing n-tangle in Ref. [13].

[2]An explicit example of a $k/2$-separable (and thus biseparable) k-qubit state (for even k) with the bipartite entanglement between all neighbouring qubits in a linear arrangement can be found in [21, footnote 30].

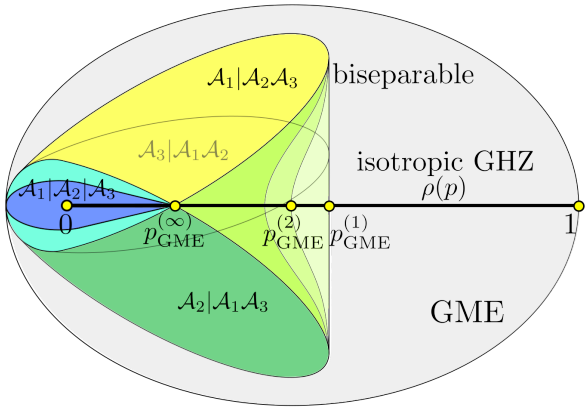

Figure 1: **GME and (partial) separability for three qubits**. All three-qubit states separable with respect to one of the three bipartitions, $\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3$ (yellow), $\mathcal{A}_2|\mathcal{A}_1\mathcal{A}_3$ (darker green), and $\mathcal{A}_3|\mathcal{A}_1\mathcal{A}_2$ (background), form convex sets, whose intersection (turquoise) contains (but is not limited to) all fully separable states $\mathcal{A}_1|\mathcal{A}_2|\mathcal{A}_3$ (dark blue). The convex hull of these partition-separable states contains all partially separable (the same as biseparable for tripartite systems) states. All states that are not biseparable are GME. States with k-copy activatable GME are contained in the set of biseparable but not partition-separable states and are conjectured to form the lighter green areas, with those states for which GME is activatable for higher values of k farther away from the border between GME and biseparability. The horizontal line represents the family of isotropic GHZ states $\rho(p)$, containing the maximally mixed state ($p = 0$) and the GHZ state ($p = 1$). The values $p_{\mathrm{GME}}^{(k)}$ indicate k-copy GME activation thresholds, which we discuss in the following.

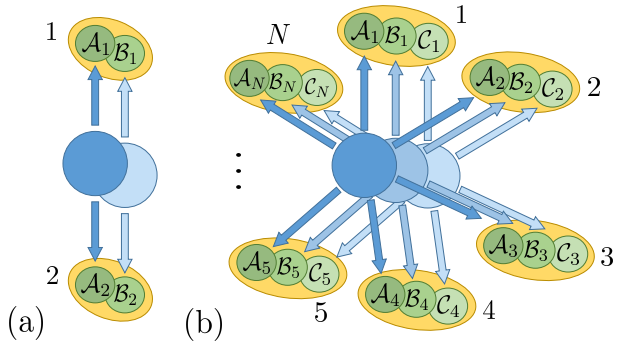

Figure 2: **Activation of GME from biseparable states**. (a) Separable bipartite states remain separable, no matter how many copies are shared, e.g., if $\rho_{\mathcal{A}_1\mathcal{A}_2}$ and $\rho_{\mathcal{B}_1\mathcal{B}_2}$ are separable with respect to the bipartitions $\mathcal{A}_1|\mathcal{A}_2$ and $\mathcal{B}_1|\mathcal{B}_2$, then so is $\rho_{\mathcal{A}_1\mathcal{A}_2} \otimes \rho_{\mathcal{B}_1\mathcal{B}_2}$. (b) In contrast, the joint state of multiple copies of biseparable states, e.g., $\rho_{\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_N}$, $\rho_{\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_N}$, and $\rho_{\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_N}$, can be GME with respect to the partition $\mathcal{A}_1\mathcal{B}_1\mathcal{C}_1|\mathcal{A}_2\mathcal{B}_2\mathcal{C}_2|\ldots|\mathcal{A}_N\mathcal{B}_N\mathcal{C}_N$.

partial separability, and conventional methods, such as positive maps [22, 23], cannot be straightforwardly applied to reveal GME (with new concepts for positive maps derived for that purpose in [24, 25]), which results in additional challenges compared to the — relatively — simpler scenario of detecting bipartite or partition entanglement (e.g., as in [26]).

An assumption inherent in the definitions above is that all parties locally act only on a single copy of the distributed state. However, in many experiments where quantum states are distributed among (potentially distant) parties, multiple independent but identically prepared copies of states are (or at least, can be) shared. For instance, exceptionally high visibilities of photonic states can only be achieved if each detection event stems from almost identical quantum states [27, 28]. Adding noise to the channel then produces the situation we focus on in this article: multiple copies of noisy quantum states produced in a laboratory [29, 30]. Even limited access to quantum memories or signal delays then allows one to act on multiple copies of the distributed states, which is a recurring theme also in research on quantum networks [31–33]. Characterising properties of GME in multi-copy scenarios is thus not only of fundamental theoretical interest but also crucial for practical applications that require GME to be distributed, such as conference key agreement [34].

However, we demonstrate here that, unlike the distinction between separable and entangled states, the distinction between biseparability and GME is not maintained in the transition from one to many copies; i.e., partial separability is not a tensor-stable concept. As we show, for N parties $1, \ldots, N$, there exist multipartite quantum states $\rho_{\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_N}$ that are biseparable, but which can be *activated* in the sense that sharing two copies results in a GME state, i.e., such that the joint state $\rho_{\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_N} \otimes \rho_{\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_N}$ of two identical copies (labelled $\mathcal{A}$ and $\mathcal{B}$, respectively) is not biseparable with respect to the partition $\mathcal{A}_1\mathcal{B}_1|\mathcal{A}_2\mathcal{B}_2|\ldots|\mathcal{A}_N\mathcal{B}_N$. (See Fig. 2.) That such activation of GME is in principle possible had previously only been noted in [35], where it was observed that two copies of a particular four-qubit state that is itself almost fully separable can become GME. Here, we systematically investigate this phenomenon of *multi-copy GME activation*. As the first main result, we show that the property of biseparability is not tensor stable in general by identifying a family of N-qubit isotropic Greenberger-Horne-Zeilinger (GHZ) states with two-copy activatable GME for all N. We further demonstrate the existence of biseparable states within this family for which two copies are not enough to activate GME, but three copies are. Moreover, we show that the bound for partition-separability coincides with the asymptotic (in terms of the number of copies) GME-activation bound for isotropic GHZ states.

Multi-copy GME activation is particularly remarkable — and may appear surprising at first — because it is in stark contrast to bipartite entanglement: Two copies of states separable with respect to a fixed partition always remain partition-separable and can

Table 1: Summary of terminology on GME in this paper.

Term	Meaning
k-separable	convex combination of pure states, each of which is a product of at least k projectors
biseparable	synonymous with 2-separable
partially separable	k-separable for some $k > 1$
partition-separable	separable for a specific partition of the multipartite Hilbert space, i.e., a convex combination of projectors, each of which is a product with respect to the same partition into subsystems
multipartite entangled	entangled across all bipartitions
genuine multipartite entangled	non-biseparable

never become GME. However, from the perspective of entanglement distillation — the concentration of entanglement from many weakly entangled (copies of) states to few strongly entangled ones — such an activation seems more natural. After all, if one party shares bipartite maximally entangled states with each other party, these could be used to establish any GME state among all N parties via standard teleportation, thus distributing GME by sharing only two-party entangled states. Nevertheless, such a procedure would require at least $N-1$ copies of these bipartite entangled states (in addition to a local copy of the GME state to be distributed), and already the example from [35] suggests that one does not have to go through first distilling bipartite entangled pairs, followed by teleportation, but two copies can naturally feature GME already. While we have seen that the phenomenon of GME activation is more than just distillation, one may still be tempted to think that distillable entanglement is required for GME activation. It is known that there exist bound entangled states — entangled states that do not admit distillation of entanglement no matter how many copies are provided. In particular, all entangled states with positive partial transpose (PPT) across a given cut are undistillable since any number of copies is also PPT. One might thus suspect that GME activation should not be possible for biseparable states that are PPT across every cut and hence have no distillable entanglement (even if multiple parties are allowed to collaborate). As another main result, we show that this is not the case by constructing a biseparable state that is PPT with respect to every cut, yet two copies of the state are indeed GME. Together, our results thus support the following conjectures:

(i) There exists a hierarchy of states with k-copy activatable GME, i.e., for all $k \geq 2$ there exists a biseparable but not partition-separable state ρ such that $\rho^{\otimes k-1}$ is biseparable, but $\rho^{\otimes k}$ is GME.

(ii) GME may be activated for any biseparable but not partition-separable state (light green areas in Fig. 1) of any number of parties as $k \to \infty$.

In the following, we first provide the formal definitions for biseparability and GME in Sec. 2 before turning to the family of N-qubit isotropic GHZ states in Sec. 3. For all biseparable states in this family, we provide upper bounds on the minimal number of copies required to activate GME in Sec. 4. In Sec. 5, we then consider the case of three qubits ($N = 3$), for which we can show that the bound on three-copy GME activation is tight in the sense that we identify all states in the family for which one requires at least three copies to activate GME, while two copies remain biseparable, and can also show that GME can indeed be activated for any biseparable but not partition-separable state in this family. Moreover, in Sec. 6, we construct an explicit example for two-copy GME activation from biseparable states with no distillable bipartite entanglement. Finally, we discuss the implications of our results and open questions in Sec. 8.

2 Definitions of biseparability & GME

We summarise the formal definitions of biseparability and GME in this paper. (See also Table 1 for the summary of the definitions here.) Formally, a pure quantum state of an N-partite system with Hilbert space $\mathcal{H}^{(N)} = \bigotimes_{i=1}^{N} \mathcal{H}_i$ is separable with respect to a k-partition $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k\}$, with $\mathcal{A}_i \subset \{1, 2, 3, \ldots, N\}$ and $\bigcup_{i=1}^{k} \mathcal{A}_i = \{1, 2, 3, \ldots, N\}$ such that $\bigotimes_{i=1}^{k} \mathcal{H}_{\mathcal{A}_i} = \mathcal{H}^{(N)}$, if it can be written as

$$|\Phi^{(k)}\rangle = \bigotimes_{i=1}^{k} |\phi_{\mathcal{A}_i}\rangle\,, \quad |\phi_{\mathcal{A}_i}\rangle \in \mathcal{H}_{\mathcal{A}_i}\,. \tag{1}$$

When generalising to density matrices, it is common not to specify all possible partitions, but to use the notion of *k-separability* instead: A density operator is called *k-separable* if it can be decomposed as a convex sum of pure states that are all separable with respect to *some* k-partition, i.e., if it is of the form (see, e.g., the review [20])

$$\rho^{(k)} = \sum_i p_i \, |\Phi_i^{(k)}\rangle\langle\Phi_i^{(k)}| \, . \tag{2}$$

Note that the lack of tensor stability of partial separability shown in the following also implies that the related concept of k-producibility [36, 37] is not tensor stable. Crucially, each $|\Phi_i^{(k)}\rangle$ may be k-separable with respect to a different k-partition. Consequently, k-separability does not imply separability of $\rho^{(k)}$ with respect to a specific partition, except when $\rho^{(k)}$ is a pure state or when $k = N$. In the latter case the state is called *fully separable*. To make this distinction more explicit, we refer to all (at least) biseparable states that are actually separable with respect to some bipartition as *partition-separable*. At the other end of this separability spectrum one encounters *biseparable states* ($k = 2$), while all states that are not at least biseparable (formally, $k = 1$) are called *genuinely N-partite entangled*. We will here use the term GME for the case $k = 1$. The operational reason for this definition of GME is easily explained: any biseparable state of the form of Eq. (2) can be created by N parties purely by sharing partition-separable states of the form of Eq. (1) and some classical randomness. In addition, this conveniently results in a convex notion of biseparability (as illustrated for the example in Fig. 1) amenable to entanglement witness techniques, which inherently rely on convexity.

3 GME of isotropic GHZ states

To overcome the difficulty in analysing GME, the crucial technique here is to use states in X-form, i.e., those with nonzero entries of density operators only on the main diagonal and main anti-diagonal with respect to the computational basis. Let us now consider a family of mixed N-qubit states, *isotropic GHZ states*, given by

$$\rho(p) = p \, |\mathrm{GHZ}_N\rangle\langle\mathrm{GHZ}_N| + (1-p) \, \tfrac{1}{2^N}\mathbb{1}_{2^N} \, , \tag{3}$$

obtained as convex combination of the N-qubit maximally mixed state $\frac{1}{2^N}\mathbb{1}_{2^N}$ and a pure N-qubit GHZ state

$$|\mathrm{GHZ}_N\rangle = \tfrac{1}{\sqrt{2}}\big(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}\big). \tag{4}$$

with real mixing parameter $p \in [-1/(2^N-1), 1]$. Since states in this family are in X-form with respect to the N-qubit computational basis, we can straightforwardly calculate the *genuine multipartite* (GM) *concurrence*, an entanglement measure for a multipartite state defined in terms of a polynomial of elements of its density matrix [38, 39]. For any N-qubit density operator ρ_X in X-form, i.e.,

$$\rho_X = \begin{pmatrix} \tilde{a} & \tilde{z}\,\tilde{d} \\ \tilde{d}\,\tilde{z}^\dagger & \tilde{d}\,\tilde{b}\,\tilde{d} \end{pmatrix}, \tag{5}$$

where $\tilde{a} = \mathrm{diag}\{a_1, \ldots, a_n\}$, $\tilde{b} = \mathrm{diag}\{b_1, \ldots, b_n\}$, and $\tilde{z} = \mathrm{diag}\{z_1, \ldots, z_n\}$ are diagonal $n \times n$ matrices with $n = 2^{N-1}$, $a_i, b_i \in \mathbb{R}$ and $z_i \in \mathbb{C}$ for all $i = 1, 2, \ldots, n$, and $\tilde{d} = \mathrm{antidiag}\{1, 1, \ldots, 1\}$ is antidiagonal, the GM concurrence is given by

$$C_{\mathrm{GM}}(\rho_X) = 2\max\Big\{0, \max_i\{|z_i| - \sum_{j\neq i}^{n}\sqrt{a_j b_j}\}\Big\}, \tag{6}$$

and provides a necessary and sufficient condition for GME whenever $C_{\mathrm{GM}} > 0$. In the case of the state $\rho(p)$ from Eq. (3), we have $a_i = b_i = \frac{1-p}{2^N} + \delta_{i1}\frac{p}{2}$ and $z_i = \delta_{i1}\frac{p}{2}$, such that

$$C_{\mathrm{GM}}\big[\rho(p)\big] = \max\{0, |p| - (1-p)(1-2^{1-N})\}. \tag{7}$$

Thus, $\rho(p)$ is GME if and only if

$$p > p^{(1)}_{\mathrm{GME}}(N) \coloneqq \frac{2^{N-1}-1}{2^N-1}\,, \tag{8}$$

i.e., if and only if p surpasses the single-copy threshold $p^{(1)}_{\mathrm{GME}}$. Conversely, we can be certain that $\rho(p)$ is not GME for $p \leq (2^{N-1}-1)/(2^N-1)$, and hence at least biseparable.

4 Multi-copy GME criterion

Our first goal is then to check if two copies of $\rho(p)$ are GME. Since the GM concurrence is an entanglement monotone, $C_{\mathrm{GM}}\big[\rho(p)^{\otimes k}\big]$ is monotonically non-decreasing as k increases [39]; that is, if we have $C_{\mathrm{GM}}\big[\rho(p)\big] = 0$ for $\rho(p)$ in X-form, it holds that $C_{\mathrm{GM}}\big[\rho(p)^{\otimes 2}\big] \geq 0$ in general. However, using $C_{\mathrm{GM}}\big[\rho(p)^{\otimes 2}\big] > 0$ as a necessary and sufficient criterion for GME is not an option in this case, since $\rho(p)^{\otimes 2}$ may not be of X-form even if a single copy is, and we therefore generally cannot directly calculate $C_{\mathrm{GM}}\big[\rho(p)^{\otimes 2}\big]$. The crucial idea here is to make use of the fact that stochastic LOCC (SLOCC) can never create GME from a biseparable state.

To construct a sufficient GME criterion, we therefore use a map $\mathcal{E}_\circ$ implementable via SLOCC [40], which, for any two density operators ρ and σ acting on $\mathcal{H}$, maps the state $\rho \otimes \sigma$ acting on $\mathcal{H}^{\otimes 2}$ to

$$\mathcal{E}_\circ[\rho \otimes \sigma] = \frac{\rho \circ \sigma}{\mathrm{Tr}(\rho \circ \sigma)} \quad \text{on } \mathcal{H}, \tag{9}$$

where the right-hand side is a density operator acting on $\mathcal{H}$, and "$\circ$" denotes the Hadamard product (or Schur product), i.e., the component-wise multiplication of the two matrices. What is useful for us here is that the Hadamard product of two X-form matrices results in an X-form matrix. Consequently, we can directly calculate the GM concurrence for the state resulting from applying the '*Hadamard-product map*' $\mathcal{E}_\circ$ to two copies of an originally biseparable state. If the GM concurrence of $\mathcal{E}_\circ[\rho(p)^{\otimes 2}]$ is nonzero, we can conclude that two copies of $\rho(p)$ are GME, even if a single copy is not. To decide whether $\mathcal{E}_\circ[\rho(p)^{\otimes 2}]$ is GME or not, i.e., whether the GM concurrence is nonzero or not, we can ignore the normalization and just consider $\rho(p) \circ \rho(p) = \rho(p)^{\circ 2}$. Moreover, in the maximization over the index i in Eq. (6), the maximum is obtained for $i = 1$. We can thus conclude that $\rho(p)^{\otimes 2}$ is GME if

$$|z_1^2| - \sum_{j\neq 1}^{n} \sqrt{a_j^2 b_j^2} \;=\; \tfrac{p^2}{4} - (2^{N-1}-1)\big(\tfrac{1-p}{2^N}\big)^2 \;>\; 0, \tag{10}$$

which translates to the condition $p/(1-p) > \sqrt{2^{N-1}-1}/2^{N-1}$, and in turn can be reformulated to the condition

$$p > p^{(2)}_{\mathrm{GME}}(N) \coloneqq \frac{\sqrt{2^{N-1}-1}}{2^{N-1}+\sqrt{2^{N-1}-1}}. \tag{11}$$

As we see, we have $p^{(1)}_{\mathrm{GME}} > p^{(2)}_{\mathrm{GME}}$ for all $N \geq 3$, confirming that *there exist biseparable states* with values $p < p^{(1)}_{\mathrm{GME}}$ *for which two copies are GME*, i.e., such that $p > p^{(2)}_{\mathrm{GME}}$.

Moreover, we can now concatenate multiple uses of the SLOCC map $\mathcal{E}_\circ$. For instance, we can identify the threshold value $p^{(3)}_{\mathrm{GME}}$ of p at which the state $\mathcal{E}_\circ\big[\rho(p)\otimes \mathcal{E}_\circ[\rho(p)^{\otimes 2}]\big]$ resulting from 2 applications of $\mathcal{E}_\circ$ to a total of 3 copies of $\rho(p)$ is GME, or, more generally, the corresponding threshold value $p^{(k)}_{\mathrm{GME}}$ for which k copies result in a GME state after applying the map $\mathcal{E}_\circ$ a total of $k-1$ times. From Eq. (10) it is easy to see that these threshold values are obtained as

$$p^{(k)}_{\mathrm{GME}}(N) \coloneqq \frac{\sqrt[k]{2^{N-1}-1}}{2^{N-1}+\sqrt[k]{2^{N-1}-1}}. \tag{12}$$

5 Hierarchy of k-copy activatable states

The threshold values $p^{(k)}_{\mathrm{GME}}$ provide upper bounds on the minimal number of copies required to activate GME: a value p satisfying $p^{(k)}_{\mathrm{GME}} < p < p^{(k-1)}_{\mathrm{GME}}$ for $k \geq 2$ implies that k copies are enough to activate GME. But since the map $\mathcal{E}_\circ$ (does not create and) may reduce GME, it does not imply that k copies are actually needed; up to this point, there is a possibility that two copies are all it takes.

However, at least for the case of three qubits ($N = 3$) and three copies ($k = 3$), we find that this is not the case. That is, for all isotropic three-qubit GHZ states with $p \leq p^{(2)}_{\mathrm{GME}}(N=3) = \sqrt{3}/(4+\sqrt{3})$, we find that two copies are still biseparable, and thus at least three copies are required to activate GME. The explicit biseparable decomposition of two copies of the states in this range is presented in Appendix A. Although it does not constitute conclusive proof, this result nevertheless supports our first conjecture, repeated here for convenience:

Conjecture (i): There exists a hierarchy of states with k-copy activatable GME, i.e., for all $k \geq 2$ there exists a biseparable but not partition-separable state ρ such that $\rho^{\otimes k-1}$ is biseparable, but $\rho^{\otimes k}$ is GME.

The conjectured existence of a hierarchy of biseparable states with k-copy activatable GME means that states become less and less 'valuable' as the number of copies k required to obtain GME increases. At the same time, it is also clear that all partition-separable states cannot be used to activate GME because separability with respect to any fixed partition is tensor stable. But it is not clear where exactly the boundary between activatable and non-activatable states really lies (see Fig. 1).

To shed light on this question, let us again examine the isotropic GHZ states from Eq. (3) with regards to partition-separability with respect to the bipartition separating the first qubit $\mathcal{A}_1$ from the remaining $N-1$ qubits (collected in $\tilde{\mathcal{A}}_2$), i.e., $\mathcal{A}_1|\tilde{\mathcal{A}}_2$. Using this partition, we can write

$$\begin{aligned}\rho(p) &= p\,|\Phi^+\rangle\langle\Phi^+|_{\mathcal{A}_1\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2^N}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2^N}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp} \\ &= \tfrac{1+p}{2}\,\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2}\,\tfrac{1}{2^{N-1}}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp},\end{aligned} \tag{13}$$

where $|\Phi^+\rangle_{\mathcal{A}_1\tilde{\mathcal{A}}_2} = \frac{1}{\sqrt{2}}\big(|0\rangle_{\mathcal{A}_1}|\tilde{0}\rangle_{\tilde{\mathcal{A}}_2} + |1\rangle_{\mathcal{A}_1}|\tilde{1}\rangle_{\tilde{\mathcal{A}}_2}\big)$ with $|\tilde{i}\rangle_{\tilde{\mathcal{A}}_2} = \bigotimes_{j=2}^{N}|i\rangle_{\mathcal{A}_j}$ for $i = 0, 1$, $\mathbb{1}_{\tilde{\mathcal{A}}_2} = \sum_{i=0,1}|\tilde{i}\rangle\langle\tilde{i}|$ and $\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp} = \mathbb{1}_{2^{N-1}} - \mathbb{1}_{\tilde{\mathcal{A}}_2}$. From this decomposition, it becomes clear that the state can be written as a convex sum of a two-qubit state $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ (where the second qubit lives on the two-dimensional subspace of $\tilde{\mathcal{A}}_2$, spanned by the states $|\tilde{i}\rangle$ for $i = 0, 1$) and diagonal terms proportional to $\mathbb{1}_{\mathcal{A}_1} \otimes \mathbb{1}_{\tilde{\mathcal{A}}_2^\perp}$ with support in a subspace $\tilde{\mathcal{A}}_2^\perp$ orthogonal to $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$. The latter diagonal terms trivially have a separable decomposition with respect to the bipartition $\mathcal{A}_1|\tilde{\mathcal{A}}_2$. For the two-qubit state $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$, the PPT criterion offers a necessary and sufficient separability criterion, and one easily finds that the partial transpose of $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ is non-negative if $p \leq p_{\mathrm{crit}} \coloneqq 1/(1+2^{N-1})$ (see Appendix B). Further taking into account its qubit exchange symmetry, we thus find that $\rho(p)$ is partition-separable with respect to any bipartition for $p \leq p_{\mathrm{crit}}$. At the same time, we find that $\lim_{k\to\infty} p^{(k)}_{\mathrm{GME}}(N) = p_{\mathrm{crit}}$, which implies that any isotropic GHZ state with $p > p_{\mathrm{crit}}$ features k-copy

activatable GME, at least asymptotically as $k \to \infty$, and is thus also not partition-separable. This leads us to our second conjecture, also repeated here for convenience:

Conjecture (ii): GME may be activated for any biseparable but not partition-separable state of any number of parties as $k \to \infty$.

Conjecture (ii) holds for isotropic GHZ states. But does it hold in general?

6 GME activation from PPT entangled states

A situation where one might imagine Conjecture (ii) to fail is the situation of biseparable (but not partition-separable) states with PPT entanglement across every bipartition, as discussed in Sec. 1. For isotropic GHZ states, however, the PPT criterion across every cut coincides exactly with the threshold p_{crit} for biseparability (and GME activation), as one can confirm by calculating the eigenvalues of the partial transpose of $\rho(p)$ (see Appendix B). We thus turn to a different family of states, for which this is not the case.

Specifically, as we show in detail in Appendix C, we construct a family of biseparable three-party states

$$\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3} = \sum_{\substack{i,j,k=1 \\ i\neq j\neq k\neq i}}^{3} p_i \; \rho_{\mathcal{A}_i} \otimes \rho^{\text{PPT}}_{\mathcal{A}_j\mathcal{A}_k} \tag{14}$$

where the $\rho^{\text{PPT}}_{\mathcal{A}_j\mathcal{A}_k}$ are (different) two-qutrit states with PPT entanglement across the respective cuts $\mathcal{A}_j|\mathcal{A}_k$ for $j \neq k \in \{1,2,3\}$ and $\sum_i p_i = 1$. Via LOCC, three copies (labelled $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, respectively) of this state $\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$ can be converted to what we call *PPT-triangle states* of the form

$$\rho^{\text{PPT}}_{\mathcal{A}_2\mathcal{A}_3} \otimes \rho^{\text{PPT}}_{\mathcal{B}_1\mathcal{B}_3} \otimes \rho^{\text{PPT}}_{\mathcal{C}_1\mathcal{C}_2}. \tag{15}$$

Using a GME witness based on the lifted Choi map (cf. [24, 25]), we show that there exists a parameter range where these PPT-triangle states are GME. Therefore, it is proved that GME activation is possible even from biseparable states only with PPT entanglement across every bipartition.

7 GME activation and shared randomness

Provided that our conjectures are true, incoherent mixing (access to shared randomness) can lead to situations where the number of copies needed for GME activation is reduced. In the extreme case, and this is true even based only on the results already proven here (and thus independently of whether or not the conjectures turn out to be true or not), the probabilistic combination of partition-separable states (without activatable GME) can results in a state — a biseparable isotropic GHZ state — which has activatable GME. Although this may at first glance appear to be at odds with the usual understanding of bipartite entanglement, which cannot arise from forming convex combinations of separable states, we believe this can be understood rather intuitively if we view incoherent mixing as a special case of a more general scenario in which one may have any amount of information on the states that are shared between different observers. As an example, consider the following situation:

Three parties, labelled, 1, 2 and 3, share two identical (as in, the system and its subsystems have the same Hilbert space dimensions and are represented by the same physical degrees of freedom) tripartite quantum systems, labelled $\mathcal{A}$ and $\mathcal{B}$, in the states $\rho_{\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3}$ and $\rho_{\mathcal{B}_1\mathcal{B}_2|\mathcal{B}_3}$, respectively, where we assume that $\rho_{\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3}$ is separable with respect to the bipartition $\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3$ and $\rho_{\mathcal{B}_1\mathcal{B}_2|\mathcal{B}_3}$ is separable with respect to the bipartition $\mathcal{B}_1\mathcal{B}_2|\mathcal{B}_3$. Clearly, both of these systems and states individually are biseparable, but if the parties have full information about which system is which, e.g., the first system is A and the second system is B, then the joint state $\rho_{\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3} \otimes \rho_{\mathcal{B}_1\mathcal{B}_2|\mathcal{B}_3}$ can be GME with respect to the partition $\mathcal{A}_1\mathcal{B}_1|\mathcal{A}_2\mathcal{B}_2|\mathcal{A}_3\mathcal{B}_3$. In this sense, two biseparable systems can yield one GME system. Now, let us suppose that the parties do not have full information which system is in which state. For simplicity, let us assume that either system may be in either state with the same probability $\frac{1}{2}$. Then the state of either of the systems is described by the convex mixture $\rho_{\text{mix}} = \frac{1}{2}\rho_{A_1|A_2A_3} + \frac{1}{2}\rho_{B_1B_2|B_3}$, where we have kept the labels A and B, but they now refer to the same subsystems, i.e., $A_i = B_i$ for all i. The state ρ_{mix} may not be partition separable anymore, but is certainly still biseparable. In particular, it may have activatable GME, even though neither $\rho_{A_1|A_2A_3}$ nor $\rho_{B_1B_2|B_3}$ do. For the sake of the argument let us assume that the latter is indeed the case and that GME is activated for 2 copies in this case, such that $\rho^{\otimes 2}_{\text{mix}}$ is GME. That means, if one has access to both systems, A and B, even without knowing which system is in which state, one would end up with GME. However, the additional randomness with respect to the case where one knows exactly which state which system is in results in an increased entropy of $\rho^{\otimes 2}_{\text{mix}}$ with respect to $\rho_{\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3} \otimes \rho_{\mathcal{B}_1\mathcal{B}_2|\mathcal{B}_3}$, and thus represents a disadvantage with respect to the latter case.

In general, it is therefore not problematic that the conjectures, if true, would imply that incoherent mixtures of k-activatable states may result in k'-activatable states with $k' < k$. Instead, this can be considered as a sign that scenarios with multiple copies of multipartite quantum states give rise to fea-

tures that are not captured by convex structures on the level of the single-copy state space.

8 Conclusion and outlook

Our results show that a modern theory of entanglement in multipartite systems, which includes the potential to locally process multiple copies of distributed quantum states, exhibits a rich structure that goes beyond the convex structure of partially separable states on single copies. While we conjecture that asymptotically, an even simpler description might be possible, i.e., separability in multipartite systems collapses to a simple bipartite concept of separability, we show that two copies are certainly not sufficient for reaching this simple limit, thus leaving the practical certification with finite copies a problem to be studied.

Indeed, our results show that GME is a resource with a complex relationship to bipartite entanglement in the context of local operations and shared randomness (cf. [41]). An array of important open questions arises from our results, which can thus be considered to establish an entirely new direction of research: first and foremost, this includes the quest for conclusive evidence for or against our conjectures. Besides determining whether these conjectures are ultimately correct or not, it will be of high interest to determine which properties (of the biseparable decompositions) of given states permit or prevent GME activation with a certain number of copies. Another open question is the minimal local dimension necessary for GME activation from biseparable states with PPT entanglement across every cut. Furthermore, from a practical point of view, it will be desirable to develop a theory of k-copy multipartite entanglement witnesses that are non-linear expressions of density matrices and allow for a more fine-grained characterisation of multipartite entanglement in networks with local memories. Finally, although separable states and shared classical randomness are free under LOCC, i.e., under a conventional choice of free operations in the resource theory of bipartite entanglement, our results suggest that convex combinations of different partition-separable states with shared classical randomness can be used as a resource for GME activation in multi-copy scenarios; that is, it may not be straightforward to study GME activation within the usual resource-theoretical framework under LOCC. In view of this situation, it would be interesting for future research to establish a new framework for understanding such a complicated aspect of multipartite entanglement as GME activation by, e.g., considering non-convex quantum resource theories where classical randomness can be used as a resource [42, 43].

Acknowledgments

We acknowledge support from the Austrian Science Fund (FWF) through the START project Y879-N27, the project P 31339-N27, and the Zukunftskolleg ZK03. H.Y. was supported by JSPS Overseas Research Fellowships and JST, PRESTO Grant Number JPMJPR201A, Japan.

References

[1] Jonathan P. Dowling and Gerard J. Milburn, *Quantum technology: the second quantum revolution,* Phil. Trans. R. Soc. A **361**, 1655 (2003), arXiv:quant-ph/0206091.

[2] Michael Epping, Hermann Kampermann, Chiara Macchiavello, and Dagmar Bruß, *Multi-partite entanglement can speed up quantum key distribution in networks,* New J. Phys. **19**, 093012 (2017), arXiv:1612.05585.

[3] Matej Pivoluska, Marcus Huber, and Mehul Malik, *Layered quantum key distribution,* Phys. Rev. A **97**, 032312 (2018), arXiv:1709.00377.

[4] Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner, *Fully device-independent conference key agreement,* Phys. Rev. A **97**, 022307 (2018), arXiv:1708.00798.

[5] Stefan Bäuml and Koji Azuma, *Fundamental limitation on quantum broadcast networks,* Quantum Sci. Technol. **2**, 024004 (2017), arXiv:1609.03994.

[6] Géza Tóth, *Multipartite entanglement and high-precision metrology,* Phys. Rev. A **85**, 022322 (2012), arXiv:1006.4368.

[7] Andrew J. Scott, *Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions,* Phys. Rev. A **69**, 052330 (2004), arXiv:quant-ph/0310137.

[8] Dagmar Bruß and Chiara Macchiavello, *Multipartite entanglement in quantum algorithms,* Phys. Rev. A **83**, 052313 (2011), arXiv:1007.4179.

[9] Robert Raussendorf and Hans J. Briegel, *A One-Way Quantum Computer,* Phys. Rev. Lett. **86**, 5188 (2001), arXiv:quant-ph/0010033.

[10] Hans J. Briegel and Robert Raussendorf, *Persistent Entanglement in Arrays of Interacting Particles,* Phys. Rev. Lett. **86**, 910 (2001), arXiv:quant-ph/0004051.

[11] Leonid Gurvits, *Classical complexity and quantum entanglement,* J. Comput. Syst. Sci. **69**, 448 (2004), Special Issue on STOC 2003, arXiv:quant-ph/0303055.

[12] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, *Four qubits can be entangled in nine different ways,* Phys. Rev. A **65**, 052112 (2002), arXiv:quant-ph/0109033.

[13] Andreas Osterloh and Jens Siewert, *Constructing n-qubit entanglement monotones from antilinear operators,* Phys. Rev. A **72**, 012337 (2005), arXiv:quant-ph/0410102.

[14] Julio I. de Vicente, Cornelia Spee, and Barbara Kraus, *Maximally Entangled Set of Multipartite Quantum States,* Phys. Rev. Lett. **111**, 110502 (2013), arXiv:1305.7398.

[15] Katharina Schwaiger, David Sauerwein, Martí Cuquet, Julio I. de Vicente, and Barbara Kraus, *Operational Multipartite Entanglement Measures,* Phys. Rev. Lett. **115**, 150502 (2015), arXiv:1503.00615.

[16] Julio I. de Vicente, Cornelia Spee, David Sauerwein, and Barbara Kraus, *Entanglement manipulation of multipartite pure states with finite rounds of classical communication,* Phys. Rev. A **95**, 012323 (2017), arXiv:1607.05145.

[17] C. Spee, J. I. de Vicente, D. Sauerwein, and B. Kraus, *Entangled Pure State Transformations via Local Operations Assisted by Finitely Many Rounds of Classical Communication,* Phys. Rev. Lett. **118**, 040503 (2017), arXiv:1606.04418.

[18] David Sauerwein, Nolan R. Wallach, Gilad Gour, and Barbara Kraus, *Transformations among Pure Multipartite Entangled States via Local Operations are Almost Never Possible,* Phys. Rev. X **8**, 031020 (2018), arXiv:1711.11056.

[19] Géza Tóth and Otfried Gühne, *Entanglement detection in the stabilizer formalism,* Phys. Rev. A **72**, 022340 (2005), arXiv:quant-ph/0501020.

[20] Nicolai Friis, Giuseppe Vitagliano, Mehul Malik, and Marcus Huber, *Entanglement Certification From Theory to Experiment,* Nat. Rev. Phys. **1**, 72 (2019), arXiv:1906.10929.

[21] Nicolai Friis, Oliver Marty, Christine Maier, Cornelius Hempel, Milan Holzäpfel, Petar Jurcevic, Martin B. Plenio, Marcus Huber, Christian Roos, Rainer Blatt, and Ben Lanyon, *Observation of Entangled States of a Fully Controlled 20-Qubit System,* Phys. Rev. X **8**, 021012 (2018), arXiv:1711.11092.

[22] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki, *Separability of mixed states: necessary and sufficient conditions,* Phys. Lett. A **223**, 25 (1996), arXiv:quant-ph/9605038.

[23] Asher Peres, *Separability Criterion for Density Matrices,* Phys. Rev. Lett. **77**, 1413 (1996), arXiv:quant-ph/9604005.

[24] Marcus Huber and Ritabrata Sengupta, *Witnessing Genuine Multipartite Entanglement with Positive Maps,* Phys. Rev. Lett. **113**, 100501 (2014), arXiv:1404.7449.

[25] Fabien Clivaz, Marcus Huber, Ludovico Lami, and Gláucia Murta, *Genuine-multipartite entanglement criteria based on positive maps,* J. Math. Phys. **58**, 082201 (2017), arXiv:1609.08126.

[26] Andrea Rodriguez-Blanco, Alejandro Bermudez, Markus Müller, and Farid Shahandeh, *Efficient and Robust Certification of Genuine Multipartite Entanglement in Noisy Quantum Error Correction Circuits,* PRX Quantum **2**, 020304 (2021), arXiv:2010.02941.

[27] Siddarth Koduru Joshi, Djeylan Aktas, Sören Wengerowsky, Martin Lončarić, Sebastian Philipp Neumann, Bo Liu, Thomas Scheidl, Guillermo Currás Lorenzo, Željko Samec, Laurent Kling, Alex Qiu, Mohsen Razavi, Mario Stipčević, John G. Rarity, and Rupert Ursin, *A trusted node–free eight-user metropolitan quantum communication network,* Sci. Adv. **6** (2020), arXiv:1907.08229.

[28] Sören Wengerowsky, Siddarth Koduru Joshi, Fabian Steinlechner, Hannes Hübel, and Rupert Ursin, *An entanglement-based wavelength-multiplexed quantum communication network,* Nature **564**, 225 (2018), arXiv:1801.06194.

[29] Sebastian Ecker, Frédéric Bouchard, Lukas Bulla, Florian Brandt, Oskar Kohout, Fabian Steinlechner, Robert Fickler, Mehul Malik, Yelena Guryanova, Rupert Ursin, and Marcus Huber, *Overcoming Noise in Entanglement Distribution,* Phys. Rev. X **9**, 041042 (2019), arXiv:1904.01552.

[30] Xiao-Min Hu, Wen-Bo Xing, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, Paul Erker, and Marcus Huber, *Efficient generation of high-dimensional entanglement through multipath down-conversion,* Phys. Rev. Lett. **125**, 090503 (2020), arXiv:2004.09964.

[31] Hayata Yamasaki, Alexander Pirker, Mio Murao, Wolfgang Dür, and Barbara Kraus, *Multipartite entanglement outperforming bipartite entanglement under limited quantum system sizes,* Phys. Rev. A **98**, 052313 (2018), arXiv:1808.00005.

[32] Miguel Navascues, Elie Wolfe, Denis Rosset, and Alejandro Pozas-Kerstjens, *Genuine Network Multipartite Entanglement,* Phys. Rev. Lett. **125**, 240505 (2020), arXiv:2002.02773.

[33] Tristan Kraft, Sébastien Designolle, Christina Ritz, Nicolas Brunner, Otfried Gühne, and Marcus Huber, *Quantum entanglement in the triangle network,* Phys. Rev. A **103**, L060401 (2021), arXiv:2002.03970.

[34] Gláucia Murta, Federico Grasselli, Hermann Kampermann, and Dagmar Bruß, *Quantum conference key agreement: A review,* Adv. Quantum Technol. **3**, 2000025 (2020), arXiv:2003.10186.

[35] Marcus Huber and Martin Plesch, *Purification of genuine multipartite entanglement,* Phys. Rev. A **83**, 062321 (2011), arXiv:1103.4294.

[36] Otfried Gühne and Géza Tóth, *Entanglement detection,* Phys. Rep. **474**, 1 (2009), arXiv:0811.2803.

[37] Szilárd Szalay, *k-stretchability of entanglement, and the duality of k-separability and*

k-producibility, Quantum **3**, 204 (2019), arXiv:1906.10798.

[38] Seyed Mohammad Hashemi Rafsanjani, Marcus Huber, Curtis J. Broadbent, and Joseph H. Eberly, *Genuinely multipartite concurrence of N-qubit X matrices,* Phys. Rev. A **86**, 062303 (2012), arXiv:1208.2706.

[39] Zhi-Hao Ma, Zhi-Hua Chen, Jing-Ling Chen, Christoph Spengler, Andreas Gabriel, and Marcus Huber, *Measure of genuine multipartite entanglement with computable lower bounds,* Phys. Rev. A **83**, 062325 (2011), arXiv:1101.2001.

[40] Ludovico Lami and Marcus Huber, *Bipartite depolarizing channels,* J. Math. Phys. **57**, 092201 (2016), arXiv:1603.02158.

[41] David Schmid, Denis Rosset, and Francesco Buscemi, *The type-independent resource theory of local operations and shared randomness,* Quantum **4**, 262 (2020), arXiv:1909.04065.

[42] Kohdai Kuroiwa and Hayata Yamasaki, *General Quantum Resource Theories: Distillation, Formation and Consistent Resource Measures,* Quantum **4**, 355 (2020), arXiv:2002.02458.

[43] Kohdai Kuroiwa and Hayata Yamasaki, *Asymptotically consistent measures of general quantum resources: Discord, non-Markovianity, and non-Gaussianity,* Phys. Rev. A **104**, L020401 (2021), arXiv:2103.05665.

Appendix

The appendices are organised as follows. In Appendix A, we analyse which values of the parameter allow for a biseparable decomposition of two copies of the three-qubit isotropic Greenberger-Horne-Zeilinger (GHZ) states. In Appendix B, we study the positive-partial-transpose (PPT) criterion for isotropic GHZ states. In Appendix C, we show that multi-copy activation of genuine multipartite entanglement (GME) is possible from PPT bound entanglement.

A Biseparable decomposition of two-copy three-qubit isotropic GHZ states

In this appendix we analyse which values of the parameter p allow for a biseparable decomposition of two copies of the three-qubit isotropic GHZ states. To be more precise, we look for a biseparable decomposition with respect to the partition $\mathcal{A}_1\mathcal{B}_1|\mathcal{A}_2\mathcal{B}_2|\mathcal{A}_3\mathcal{B}_3$ of the state $\rho_3(p)^{\otimes 2}$, where

$$\rho_3(p) = p\ |\mathrm{GHZ}_3\rangle\langle\mathrm{GHZ}_3| + (1-p)\,\tfrac{1}{2^3}\mathbb{1}_{2^3} \tag{16}$$

is the three-qubit isotropic GHZ state defined in the main text.

To construct a biseparable decomposition, we first construct separable states for two or four qubits. We then map these states to different six-qubit states in such a way that all resulting six-qubit states are separable with respect to one of the bipartitions

$$\mathcal{A}_1\mathcal{B}_1|\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3\mathcal{B}_3, \mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2|\mathcal{A}_3\mathcal{B}_3, \mathcal{A}_2\mathcal{B}_2|\mathcal{A}_1\mathcal{B}_1\mathcal{A}_3\mathcal{B}_3. \tag{17}$$

For convenience of notation, we henceforth reorder the subsystems to $\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3\mathcal{B}_3$. We then group together these different states to define biseparable states for the whole six-qubit system. This allows us to rewrite the state $\rho_3(p)^{\otimes 2}$ as a convex sum of these biseparable states and a diagonal matrix. Finally, we find conditions for which this diagonal matrix has only non-negative entries, i.e., is positive semi-definite and thus itself a state.

Let us begin by defining the separable two-qubit state

$$\gamma = \frac{1}{4}(|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}| + |rl\rangle\langle rl| + |lr\rangle\langle lr|), \tag{18}$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, $|r\rangle = (|0\rangle - i\,|1\rangle)/\sqrt{2}$ and $|l\rangle = (|0\rangle + i\,|1\rangle)/\sqrt{2}$. We partition the six-qubit space $\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3\mathcal{B}_3$ into two subsystems $\mathcal{C}$ and $\mathcal{D}$ in such a way that the bipartition $\mathcal{C}|\mathcal{D}$ coincides with one of the three bipartitions in (17). We then define a map E from a two-qubit state space to the six-qubit space $\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3$ as the unique linear map such that $|00\rangle \to |ii'\rangle$, $|01\rangle \to |ij'\rangle$, $|10\rangle \to |ji'\rangle$ and $|11\rangle \to |jj'\rangle$, where $|i\rangle$ and $|j\rangle$ are orthogonal states of subsystem $\mathcal{C}$ and $|i'\rangle$ and $|j'\rangle$ are orthogonal states of subsystem $\mathcal{D}$. Applying this map to the two-qubit separable state γ above, we have a six-qubit state $E(\gamma)$ that is separable across the cut $\mathcal{C}|\mathcal{D}$ by construction. In the following we will consider only such embeddings E that map $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ onto four of the standard-basis states of the six-qubit space. For example consider the partition $\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2|\mathcal{A}_3\mathcal{B}_3$ and the embedding E that maps $|00\rangle \to |000000\rangle$, $|01\rangle \to |000001\rangle$, $|10\rangle \to |010100\rangle$ and $|11\rangle \to |010101\rangle$. The embedded state then reads

$$\begin{aligned} E(\gamma) = \frac{1}{4}(&|000000\rangle\langle 000000| + |000000\rangle\langle 010101| + |010101\rangle\langle 000000| + |010101\rangle\langle 010101| \\ &+ |000001\rangle\langle 000001| + |010100\rangle\langle 010100|). \end{aligned} \tag{19}$$

For every index m running from 1 to 64, we let $|m\rangle = |i_1 i_2 i_3 i_4 i_5 i_6\rangle$ denote a standard-basis state of $\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3\mathcal{B}_3$ such that

$$m = 32i_1 + 16i_2 + 8i_3 + 4i_4 + 2i_5 + i_6 + 1, \tag{20}$$

that is, m is the decimal representation of the number represented by the bit string $i_1 i_2 i_3 i_4 i_5 i_6$. Let E_{m_1,m_2,m_3,m_4} be the linear map from a two-qubit space to the previously considered six-qubit space such that

$$|00\rangle \mapsto |m_1\rangle\,, \tag{21a}$$
$$|01\rangle \mapsto |m_2\rangle\,, \tag{21b}$$
$$|10\rangle \mapsto |m_3\rangle\,, \tag{21c}$$
$$|11\rangle \mapsto |m_4\rangle\,. \tag{21d}$$

We then define

$$\gamma(m_1, m_2, m_3, m_4) = E_{m_1,m_2,m_3,m_4}(\gamma). \tag{22}$$

For example, the state (19) is denoted by $\gamma(1, 2, 21, 22) = E_{1,2,21,22}(\gamma)$. Note that not all combinations m_1, m_2, m_3, m_4 define a two-qubit subspace across the bipartitions in (17), and among all subspaces, we are only interested in those pertaining to different parties. With this notation, we introduce the following states

$$\begin{aligned}\Gamma_1 = \frac{1}{24}[&\gamma(2,10,36,44) + \gamma(2,12,34,44) + \gamma(33,37,50,54) + \gamma(3,7,20,24) + \gamma(3,8,19,24)\\ &+\gamma(5,7,45,47) + \gamma(5,15,37,47) + \gamma(9,10,29,30) + \gamma(9,14,25,30) + \gamma(18,20,58,60)\\ &+\gamma(18,28,50,60) + \gamma(41,45,58,62) + \gamma(41,46,57,62) + \gamma(21,29,55,63) + \gamma(21,31,53,63)\\ &+\gamma(35,36,55,56) + \gamma(35,40,51,56) + \gamma(6,8,46,48) + \gamma(6,14,40,48) + \gamma(11,12,31,31)\\ &+\gamma(11,15,28,32) + \gamma(17,19,57,59) + \gamma(17,25,51,59) + \gamma(33,34,53,54)],\end{aligned} \tag{23}$$

$$\begin{aligned}\Gamma_2 = \frac{1}{12}[&\gamma(1,2,21,22) + \gamma(1,5,18,22) + \gamma(1,6,17,22) + \gamma(1,3,41,43) + \gamma(1,9,35,43)\\ &+\gamma(1,11,33,43) + \gamma(22,24,62,64) + \gamma(22,30,56,64) + \gamma(22,32,54,64) + \gamma(43,44,63,64)\\ &+\gamma(43,47,60,64) + \gamma(43,48,59,64)].\end{aligned} \tag{24}$$

With the same notation as before we define the four-qubit separable state

$$\begin{aligned}\sigma = \frac{1}{16}(&|++++\rangle\langle++++| + |+-+-\rangle\langle+-+-| + |-+-+\rangle\langle-+-+| + |----\rangle\langle----|\\ &+|+r+l\rangle\langle+r+l| + |+l+r\rangle\langle+l+r| + |-r-l\rangle\langle-r-l| + |-l-r\rangle\langle-l-r|\\ &+|r+l+\rangle\langle r+l+| + |r-l-\rangle\langle r-l-| + |l+r+\rangle\langle l+r+| + |l-r-\rangle\langle l-r-|\\ &+|rrll\rangle\langle rrll| + |rllr\rangle\langle rllr| + |lrrl\rangle\langle lrrl| + |llrr\rangle\langle llrr|),\end{aligned} \tag{25}$$

shared between three parties. It can be split in three different ways: $\sigma_{\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{A}_3}$, $\sigma_{\mathcal{A}_1\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3}$ and $\sigma_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3\mathcal{B}_3}$. Next, we define the biseparable six-qubit state

$$\Sigma = \frac{1}{3}(U_1\sigma_{\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{A}_3}U_1^\dagger + U_2\sigma_{\mathcal{A}_1\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3}U_2^\dagger + U_3\sigma_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3\mathcal{B}_3}U_3^\dagger) \tag{26}$$

where U_k are isometries of the form $U_1\,|ij\rangle_{\mathcal{A}_2\mathcal{A}_3} = |iijj\rangle_{\mathcal{A}_2\mathcal{B}_2\mathcal{A}_3\mathcal{B}_3}$, $U_2\,|ij\rangle_{\mathcal{A}_1\mathcal{A}_3} = |iijj\rangle_{\mathcal{A}_1\mathcal{B}_1\mathcal{A}_3\mathcal{B}_3}$ and $U_3\,|ij\rangle_{\mathcal{A}_1\mathcal{A}_2} = |iijj\rangle_{\mathcal{A}_1\mathcal{B}_1\mathcal{A}_2\mathcal{B}_2}$.

With this we can finally rewrite the two copies of the original state as

$$\rho(p)^{\otimes 2} = (1-2p)^2\rho_{\text{diag}} + p(3-7p)\Gamma_1 + p(1-p)\Gamma_2 + 4p^2\Sigma, \tag{27}$$

where ρ_{diag} is a normalized diagonal matrix. With m defined as (20), the matrix $64(1-2p)^2\rho_{\text{diag}}$ has the following entries:

$\underline{m}$	$\underline{\rho_{\text{diag}}(m,m)}$
$1, 22, 43, 64:$	$(1-p)^2,$
$2, 3, 5, 6, 9, 11, 17, 18, 21, 24, 30, 32, 33,$	
$35, 41, 44, 47, 48, 54, 56, 59, 60, 62, 63:$	$1 - 10/3p + 7/3p^2,$
$4, 13, 16, 23, 26, 27, 38, 39, 42, 49, 52, 61:$	$1 - 2p - 13/3p^2,$
$7, 8, 10, 12, 14, 15, 19, 20, 25, 28, 29, 31,$	
$34, 36, 37, 40, 45, 46, 50, 51, 53, 55, 57, 58:$	$1 - 6p + 31/3p^2.$

The terms $(1-p)^2$ and $1-6p+31/3p^2$ are positive for all values of p. The term $1-10/3p+7/3p^2$ is non-negative for $p \leq 3/7$ and $p \geq 1$ and finally the term $1-2p-13/3p^2$ is non-negative for $(3-4\sqrt{3})/13 \leq p \leq (3+4\sqrt{3})/13$. With this we have found a biseparable decomposition for all values $-1/7 \leq p \leq (3+4\sqrt{3})/13$. From the bound shown in the main text

$$p > p^{(2)}_{\text{GME}}(N) \coloneqq \frac{\sqrt{2^{N-1}-1}}{2^{N-1}+\sqrt{2^{N-1}-1}}, \tag{28}$$

we know that all values above this bound are already GME.

B PPT criterion for isotropic GHZ states

The isotropic GHZ states defined in the main text can be rewritten as

$$\begin{aligned}\rho(p) &= p\,|\Phi^+\rangle\langle\Phi^+|_{\mathcal{A}_1\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2^N}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2^N}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp} \\ &= \tfrac{1+p}{2}\,\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2} + \tfrac{1-p}{2}\,\tfrac{1}{2^{N-1}}\,\mathbb{1}_{\mathcal{A}_1}\otimes\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp},\end{aligned} \tag{29}$$

where $|\Phi^+\rangle_{\mathcal{A}_1\tilde{\mathcal{A}}_2} = \frac{1}{\sqrt{2}}\big(|0\rangle_{\mathcal{A}_1}|\tilde{0}\rangle_{\tilde{\mathcal{A}}_2} + |1\rangle_{\mathcal{A}_1}|\tilde{1}\rangle_{\tilde{\mathcal{A}}_2}\big)$ with $|\tilde{i}\rangle_{\tilde{\mathcal{A}}_2} = \bigotimes_{j=2}^{N}|i\rangle_{\mathcal{A}_j}$ for $i = 0, 1$, $\mathbb{1}_{\tilde{\mathcal{A}}_2} = \sum_{i=0,1}|\tilde{i}\rangle\langle\tilde{i}|$ and $\mathbb{1}_{\tilde{\mathcal{A}}_2^\perp} = \mathbb{1}_{2^{N-1}} - \mathbb{1}_{\tilde{\mathcal{A}}_2}$. We are now interested in checking for which values of p the partial transpose of the two-qubit state $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ is positive semi-definite. Since the normalisation is irrelevant for this calculation, we can instead consider the partial transpose of the unnormalised operator $\frac{1+p}{2}\,\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ whose partial transpose is given by

$$\big(\tfrac{1+p}{2}\,\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}\big)^{T_{\tilde{\mathcal{A}}_2}} = \begin{pmatrix} \frac{p}{2}+\frac{1-p}{2^N} & 0 & 0 & 0 \\ 0 & \frac{1-p}{2^N} & \frac{p}{2} & 0 \\ 0 & \frac{p}{2} & \frac{1-p}{2^N} & 0 \\ 0 & 0 & 0 & \frac{p}{2}+\frac{1-p}{2^N} \end{pmatrix}. \tag{30}$$

The only potentially negative eigenvalue of this matrix is $(1-p)/2^N - p/2$ and we hence find that $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ is positive semi-definite for $p \leq p_{\text{crit}} \coloneqq 1/(1+2^{N-1})$. Since $\tilde{\rho}_{\mathcal{A}_1\tilde{\mathcal{A}}_2}$ is a two-qubit state, the PPT criterion is necessary and sufficient for separability, and the state $\rho(p)$ hence has a separable decomposition with respect to the bipartition $\mathcal{A}_1|\mathcal{A}_2\ldots\mathcal{A}_N$ for $p \leq p_{\text{crit}}$.

Since $\rho(p)$ is invariant under exchanges of any qubits, this separability threshold applies for any bipartition of separating any one qubit from the remaining $N-1$ qubits. Moreover, it is easy to see that the arguments presented above hold also for any bipartition into M and $N-M$ qubits by choosing suitable single-qubit subspaces in both the M-qubit and $(M-N)$-qubit Hilbert spaces.

We also note that the threshold value p_{crit} for partition-separability trivially coincides with the PPT threshold for any chosen bipartition of $\rho(p)$ because the only non-diagonal 2×2-block of the partial transpose is always of the form of the right-hand side of Eq. (30). In particular, this implies that all states $\rho(p)$ are non-PPT (NPT) entangled across any bipartition for $p > p_{\text{crit}}$ and separable below this value. Consequently, there are no PPT entangled isotropic GHZ states.

C PPT-triangle states and GME activation

To investigate whether multi-copy GME activation is possible from bound entanglement, we first consider a biseparable three-party state with no distillable bipartite entanglement across any bipartition; i.e., the state is positive under partial transposition across all cuts. Since the set of PPT states is convex, we may construct such a state as a convex combination of terms where one party is uncorrelated with the others, while the remaining two parties share a PPT entangled state, i.e.,

$$\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3} = p_1\rho_{\mathcal{A}_1}\otimes\rho^{\mathrm{PPT}}_{\mathcal{A}_2\mathcal{A}_3} + p_2\rho_{\mathcal{A}_2}\otimes\rho^{\mathrm{PPT}}_{\mathcal{A}_1\mathcal{A}_3} + p_3\rho_{\mathcal{A}_3}\otimes\rho^{\mathrm{PPT}}_{\mathcal{A}_1\mathcal{A}_2}, \tag{31}$$

where $\sum_i p_i = 1$, $p_i \geq 0$ and $\rho^{\mathrm{PPT}}_{\mathcal{A}_i\mathcal{A}_j}$ for $i,j\in\{1,2,3\}$ are PPT entangled states. Here we note that the existence of such a decomposition guarantees biseparability, but it does not a priori rule out that such a state may be partition-separable (or even fully separable). If $\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$ is separable with respect to one or several of the bipartitions $\mathcal{A}_1|\mathcal{A}_2\mathcal{A}_3$, $\mathcal{A}_1\mathcal{A}_2|\mathcal{A}_3$ and $\mathcal{A}_2|\mathcal{A}_1\mathcal{A}_3$, then GME activation is not possible for any number of copies. However, as we show here, for certain choices of the $\rho^{\mathrm{PPT}}_{\mathcal{A}_i\mathcal{A}_j}$ and $\rho_{\mathcal{A}_k}$, three copies of $\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$ are GME, which thus also shows that the single-copy states in question are not partition-separable (or fully separable).

To continue, let us consider the particular situation where each of the three parties $\mathcal{A}_i$ for $i=1,2,3$ consists of three subsystems $\mathcal{A}_i^{(j)}$ for $j=1,2,3$. In this situation, a particular example for a state of the form of Eq. (31) is given by

$$\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3} = \sum_{\substack{i=1,2,3\\ i\neq j\neq k\neq i\\ j<k}} p_i\ \rho_{\mathcal{A}_i^{(i)}}\otimes\rho^{\mathrm{PPT}}_{\mathcal{A}_j^{(i)}\mathcal{A}_k^{(i)}}\otimes\bigotimes_{\substack{m,n=1\\ n\neq i}}^{3}|0\rangle\langle 0|_{\mathcal{A}_m^{(n)}}\,, \tag{32}$$

where the states $\rho^{\mathrm{PPT}}_{\mathcal{A}_j^{(i)}\mathcal{A}_k^{(i)}}$ are PPT entangled states that will be specified later. Now, suppose that three copies, $\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$, $\rho_{\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3}$, and $\rho_{\mathcal{C}_1\mathcal{C}_2\mathcal{C}_3}$, are shared. By projecting the subsystems $\mathcal{A}_1^{(1)}$ of the first copy, $\mathcal{B}_2^{(2)}$ of the second copy, and $\mathcal{C}_3^{(3)}$ of the third copy into the subspaces orthogonal to the states $|0\rangle_{\mathcal{A}_1^{(1)}}$, $|0\rangle_{\mathcal{B}_2^{(2)}}$, and $|0\rangle_{\mathcal{C}_3^{(3)}}$, respectively, the three parties can (deterministically) prepare the states $\rho^{\mathrm{PPT}}_{\mathcal{A}_2^{(1)}\mathcal{A}_3^{(1)}}$, $\rho^{\mathrm{PPT}}_{\mathcal{B}_1^{(2)}\mathcal{B}_3^{(2)}}$, and $\rho^{\mathrm{PPT}}_{\mathcal{C}_1^{(3)}\mathcal{C}_2^{(3)}}$. All other subsystems can be discarded. Consequently, three copies of $\rho_{\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$ allow the parties to establish a state of the form

$$\rho^{\Delta\mathrm{PPT}}_{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3} \coloneqq \rho^{\mathrm{PPT}}_{\mathcal{A}_2\mathcal{A}_3}\otimes\rho^{\mathrm{PPT}}_{\mathcal{B}_1\mathcal{B}_3}\otimes\rho^{\mathrm{PPT}}_{\mathcal{C}_1\mathcal{C}_2} \tag{33}$$

via LOCC. For ease of notation we have dropped the superscripts identifying the particular subsystems, e.g., using the label $\mathcal{A}_i$ instead of $\mathcal{A}_i^{(j)}$. We call a state in this form a *PPT-triangle* state, where the parties 1, 2, and 3 have access to systems $\mathcal{B}_1\mathcal{C}_1$, $\mathcal{A}_2\mathcal{C}_2$, and $\mathcal{A}_3\mathcal{B}_3$, respectively. We further note that every such PPT-triangle state can be created via LOCC from three copies of a biseparable state of the form of Eq. (31).

Therefore, we reach the following claim: if there is a GME state that is PPT-triangle, then multi-copy GME activation is achievable for (some) biseparable states that are PPT across every cut. Consequently, the problem reduces to proving the existence of a PPT-triangle state that exhibits GME. To find such a state, we construct a one-parameter family of two-qutrit states given by

$$\begin{aligned}\rho^{\mathrm{PPT}}_{\mathcal{X}\mathcal{Y}}(p) \coloneqq \tfrac{1}{\mathcal{N}_p}\big[&(|00\rangle+|11\rangle+|22\rangle)(\langle 00|+\langle 11|+\langle 22|) + p(|01\rangle\langle 01|+|12\rangle\langle 12|+|20\rangle\langle 20|)\\ &+\tfrac{1}{p}(|02\rangle\langle 02|+|10\rangle\langle 10|+|21\rangle\langle 21|)\big],\end{aligned} \tag{34}$$

for all $p>0$, where $\mathcal{X}$ and $\mathcal{Y}$ labels the first and second qutrit, respectively, $\mathcal{N}_p = 3(1+p+\frac{1}{p}) > 0$ is a normalization constant. The partial transpose of $\rho^{\mathrm{PPT}}_{\mathcal{X}\mathcal{Y}}(p)$ has eigenvalues $\lambda_1 = 0$, $\lambda_2 = \mathcal{N}_p > 0$, and $\lambda_3 = \mathcal{N}_p(p+\frac{1}{p}) > 0$, each thrice degenerate, and $\rho^{\mathrm{PPT}}_{\mathcal{X}\mathcal{Y}}(p)$ is hence PPT. We can then choose the PPT states in Eq. (33) from this family of two-qutrit states, such that

$$\rho^{\Delta\mathrm{PPT}}_{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3}(x,y,z) \coloneqq \rho^{\mathrm{PPT}}_{\mathcal{A}_2\mathcal{A}_3}(x)\otimes\rho^{\mathrm{PPT}}_{\mathcal{B}_1\mathcal{B}_3}(y)\otimes\rho^{\mathrm{PPT}}_{\mathcal{C}_1\mathcal{C}_2}(z). \tag{35}$$

To show that the state is GME with respect to the partition $\mathcal{O}_1|\mathcal{O}_2|\mathcal{O}_3$ it suffices to detect GME between subspaces $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$ of $\mathcal{O}_1$, $\mathcal{O}_2$, and $\mathcal{O}_3$, respectively. Specifically, we consider the single-qutrit subspaces $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$ spanned by $\{|ii\rangle_{\mathcal{B}_1\mathcal{C}_1}\}_{i=0,1,2}$, $\{|jj\rangle_{\mathcal{A}_2\mathcal{C}_2}\}_{j=0,1,2}$, and $\{|kk\rangle_{\mathcal{A}_3\mathcal{B}_3}\}_{k=0,1,2}$, respectively, and thus the projection of $\rho^{\Delta\mathrm{PPT}}_{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3}(x,y,z)$ onto the three-qutrit subspace spanned by $\{|ii\rangle_{\mathcal{B}_1\mathcal{C}_1}\otimes|jj\rangle_{\mathcal{A}_2\mathcal{C}_2}\otimes|kk\rangle_{\mathcal{A}_3\mathcal{B}_3}\}_{i,j,k=0,1,2}$. We let $\rho^{\Delta\mathrm{PPT}}_{\mathcal{D}_1\mathcal{D}_2\mathcal{D}_3}(x,y,z)$ denote the resulting state.

To this state, we apply a three-party GME witness W_3 (see [24, example 2]) based on the lifted Choi-map witnesses from [25] of the form

$$\begin{aligned} W_3 = \; & |000\rangle\langle 000| + |001\rangle\langle 001| + |011\rangle\langle 011| + |020\rangle\langle 020| + |101\rangle\langle 101| + |111\rangle\langle 111| + |112\rangle\langle 112| \\ & + |122\rangle\langle 122| + |200\rangle\langle 200| + |212\rangle\langle 212| + |220\rangle\langle 220| + |222\rangle\langle 222| - |000\rangle\langle 111| - |000\rangle\langle 222| \\ & - |111\rangle\langle 222| - |111\rangle\langle 000| - |222\rangle\langle 000| - |222\rangle\langle 111|. \end{aligned} \tag{36}$$

Applying it to our state yields the expression

$$\mathrm{Tr}[W_3\, \rho^{\Delta\mathrm{PPT}}_{\mathcal{D}_1\mathcal{D}_2\mathcal{D}_3}(x,y,z)] = \tfrac{3}{\mathcal{N}_x\mathcal{N}_y\mathcal{N}_z}(xy + \tfrac{z}{x} + yz - 1). \tag{37}$$

We see that for certain values of x, y and z the expected value of the witness can be negative, e.g., for states of the form $\rho^{\Delta\mathrm{PPT}}_{\mathcal{D}_1\mathcal{D}_2\mathcal{D}_3}(1,y,y)$ with $0 < y < \sqrt{2} - 1$, thus detecting GME in this range.

Finally, an observation that we can make about the PPT-triangle states in Eq. (35) is that the third tensor factor $\rho^{\mathrm{PPT}}_{\mathcal{C}_1\mathcal{C}_2}(z)$ is not even necessary to obtain GME. Indeed, the state

$$\rho^{\wedge\mathrm{PPT}}_{\mathcal{A}_2\mathcal{A}_3\mathcal{B}_1\mathcal{B}_3}(x,y) = \rho^{\mathrm{PPT}}_{\mathcal{A}_2\mathcal{A}_3}(x) \otimes \rho^{\mathrm{PPT}}_{\mathcal{B}_1\mathcal{B}_3}(y) \tag{38}$$

is GME for certain values of x and y. To show this it again suffices detecting GME on a subspace. Consider the projection onto the three-qutrit subspace spanned by $\{|i\rangle_{\mathcal{B}_1} \otimes |j\rangle_{\mathcal{A}_2} \otimes |kk\rangle_{\mathcal{A}_3\mathcal{B}_3}\}_{i,j,k=0,1,2}$ and denote the resulting state by $\rho^{\wedge\mathrm{PPT}}_{\mathcal{D}_1\mathcal{D}_2\mathcal{D}_3}(x,y)$. With the same witness W_3 as before we obtain

$$\mathrm{Tr}[W_3\, \rho^{\wedge\mathrm{PPT}}_{\mathcal{D}_1\mathcal{D}_2\mathcal{D}_3}(x,y)] = \tfrac{3}{\mathcal{N}_x\mathcal{N}_y}\,(x + y + xy - 1). \tag{39}$$

For instance, for $x = y < \sqrt{2} - 1$, this expression becomes negative, thus detecting GME. We can thus conclude that PPT entanglement across two out of the three cuts and two copies of the original state are already enough for GME activation.

10 Entanglement detection with imprecise measurements

The following research article is currently under peer review. The manuscript can be accessed online under arXiv:2202.13131.

10.1 Contribution

In this project I participated from the start, influencing the direction the project developed. I substantially contributed in the calculation of high-dimensional entanglement criteria. The bounds on witnesses in Appendix D were derived by me. Finally I helped in the exploration of various other directions, looking for analytical solutions were finally we had to resolve to numerical methods.

Entanglement detection with imprecise measurements

Simon Morelli,[1,2,*] Hayata Yamasaki,[1,2,*] Marcus Huber,[1,2] and Armin Tavakoli[1,2]

[1]*Institute for Quantum Optics and Quantum Information – IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria*
[2]*Atominstitut, Technische Universität Wien, Stadionallee 2, 1020 Vienna, Austria*

We investigate entanglement detection when the local measurements only nearly correspond to those intended. This corresponds to a scenario in which measurement devices are not perfectly controlled, but nevertheless operate with bounded inaccuracy. We formalise this through an operational notion of inaccuracy that can be estimated directly in the lab. To demonstrate the relevance of this approach, we show that small magnitudes of inaccuracy can significantly compromise several renowned entanglement witnesses. For two arbitrary-dimensional systems, we show how to compute tight corrections to a family of standard entanglement witnesses due to any given level of measurement inaccuracy. We also develop semidefinite programming methods to bound correlations in these scenarios.

Introduction.— Deciding whether an initially unknown state is entangled is one of the central challenges of quantum information science [1–3]. The most common approach is the method of entanglement witnesses, in which one hypothesises that the state is close to a known target and then finds suitable local measurements that can reveal its entanglement [4–6]. In principle, this allows for the detection of every entangled state. However, it crucially requires the experimenter to flawlessly perform the stipulated quantum measurements. This is an idealisation to which one may only aspire: even for the simplest system of two qubits, small alignment errors can cause false positives [7]. In contrast, by adopting a device-independent approach, any concerns about the modelling of the measurement devices can be dispelled. This entails viewing them as quantum black boxes and detecting entanglement through the violation of a Bell inequality [8, 9]. However, Bell experiments are practically demanding [10]. Also, many entangled states either cannot, or are not known to, violate any Bell inequality [11, 12]. In addition, for the common purpose of verifying that a non-malicious entanglement source operates as intended, a device-independent approach is to break a butterfly on a wheel. In the interest of a compromise, entanglement detection has also been investigated in steering scenarios, in which some devices are assumed to be perfectly controlled and others are quantum black boxes [13]. Nevertheless, such asymmetry is often not present in non-malicious scenarios, and the approach still suffers from drawbacks similar to both the device-independent case, albeit it milder, and the standard, fully controlled, scenario. A much less explored compromise route is to only assume knowledge of the Hilbert space dimension [14, 15]. This essentially adopts the view that the experimenter has no control over the relevant degrees of freedom. Such ideas have also been used to strengthen steering-based entanglement detection [16].

Here, we introduce an approach to entanglement detection that neither assumes flawless control of the measurements nor views them as mostly uncontrolled operations. The main idea is that an experimenter can quantitatively estimate the accuracy of their measurement devices and then base entanglement detection on this benchmark. Such knowledge naturally requires a fixed Hilbert space dimension: the experimenter knows the degrees of freedom on which they operate. To quantify the inaccuracy between the intended target measurement and the lab measurement, we use a simple fidelity-based notion that can handily be measured experimentally.

In what follows, we first establish the relevance of small inaccuracies by showcasing that the conclusions of renowned entanglement witnesses can be substantially compromised. We show that the magnitude of detrimental influence associated to a small inaccuracy does not have to decrease for higher-dimensional systems. This is important because higher-dimensional entangled systems are increasingly interesting for experiments [17–21] but typically cannot be controlled as precisely as qubits. Secondly, we develop entanglement criteria that explicitly take the degree of inaccuracy into account. For two-qubit scenarios, we provide this based on the simplest entanglement witness and the Clauser-Horne-Shimony-Holt (CHSH) quantity. For a pair of systems of any given local dimension, we show that such criteria can be analytically established as corrections to a simple family of standard entanglement witnesses. Finally, we present semidefinite programming (SDP) relaxations for bounding the set of quantum correlations under measurement inaccuracies. We use this both to estimate the potentially constructive influence of measurement inaccuracy on entanglement-based correlations and to systematically place upper bounds for separable states on linear witnesses.

Framework.— We consider sources of bipartite states $\rho = \rho_{\mathrm{AB}}$ of local dimension d. The subsystems are measured individually with settings x and y respectively, producing outcomes $a, b \in \{1, \ldots, o\}$. The experimenter's aim is to measure the first (second) system using a set of projective measurements $\{\tilde{A}_{a|x}\}$ ($\{\tilde{B}_{b|y}\}$). These are called target measurements. However, the measurements actually performed in the lab do not precisely correspond to the targeted measurements, but instead to positive operator-valued measures (POVMs) $\{A_{a|x}\}$ ($\{B_{b|y}\}$). These are called lab measurements and do not need to be projective. The correlations in the experiment are given by the Born-rule

$$p(a,b|x,y) = \mathrm{tr}\left[A_{a|x} \otimes B_{b|y}\rho\right]. \tag{1}$$

We quantify the correspondence between each of the target measurements and the associated lab measurements through

* S.M. and H.Y. contributed equally to this manuscript.

their average fidelity,

$$\mathcal{F}_x^{\mathrm{A}} \equiv \frac{1}{d}\sum_{a=1}^{o} \mathrm{tr}\left[A_{a|x}\tilde{A}_{a|x}\right], \quad \mathcal{F}_y^{\mathrm{B}} \equiv \frac{1}{d}\sum_{b=1}^{o} \mathrm{tr}\left[B_{b|y}\tilde{B}_{b|y}\right]. \tag{2}$$

The fidelity respects $\mathcal{F} \in [0,1]$ with $\mathcal{F} = 1$ if and only if the lab measurement is identical to the target measurement. Importantly, the fidelity admits a simple operational interpretation: it is the average probability of obtaining outcome a (b) when the lab measurement is applied to each of the orthonormal states spanning the eigenspace of the a'th (b'th) target projector. Thus, the fidelities $\{\mathcal{F}_x^{\mathrm{A}}, \mathcal{F}_y^{\mathrm{B}}\}$ can be directly determined by probing the lab measurements with single qudits from a well-calibrated, auxiliary, source. This requires no entanglement and can routinely be achieved, see e.g. Ref. [22]. It motivates the assumption of a bounded inaccuracy, i.e. a lower bound on each of the fidelities,

$$\mathcal{F}_x^{\mathrm{A}} \geq 1 - \varepsilon_x^{\mathrm{A}}, \qquad \mathcal{F}_y^{\mathrm{B}} \geq 1 - \varepsilon_y^{\mathrm{B}}, \tag{3}$$

where the parameter $\varepsilon \in [0,1]$ is the inaccuracy of the considered lab measurement. In the extreme case of $\varepsilon = 0$, the lab measurement is identical to the target measurement and our scenario reduces to a standard entanglement witness. In the other extreme, $\varepsilon = 1$, only the Hilbert space dimension of the measurement is known. Away from these extremes, one encounters the more realistic scenario, in which the experimenter knows the degrees of freedom, but is only able to control them up to a limited accuracy.

The simplest tests of entanglement use the minimal number of outcomes ($o = 2$). In such scenarios the fidelity constrains (3) can be simplified into

$$\mathrm{tr}\left(A_x\tilde{A}_x\right) \geq d\left(1 - 2\varepsilon_x^{\mathrm{A}}\right), \quad \mathrm{tr}\left(B_y\tilde{B}_y\right) \geq d\left(1 - 2\varepsilon_y^{\mathrm{B}}\right) \tag{4}$$

where we have defined observables $A_x \equiv A_{1|x} - A_{2|x}$ and $B_y \equiv B_{1|y} - B_{2|y}$. The observables can be arbitrary Hermitian operators whose extremal eigenvalue is bounded by unity, i.e. $\|A_x\|_\infty \leq 1$ and $\|B_y\|_\infty \leq 1$.

Notice that the proposed framework immediately extends also to multipartite scenarios.

Impact of inaccuracies in entanglement witnessing.— A crucial motivating question for our approach is whether, and to what extent, small inaccuracies in the measurement devices ($\varepsilon \ll 1$) impact the analysis of a conventional entanglement witness. We discuss this matter based on several well-known witnesses.

Firstly, consider the simplest entanglement witness for two qubits, involving two pairs local Pauli observables: $\mathcal{W} = \langle\sigma_X \otimes \sigma_X\rangle + \langle\sigma_Z \otimes \sigma_Z\rangle$. For separable states we have $\mathcal{W} \leq \mathcal{W}_{\mathrm{sep}} = 1$ and for entangled states $\mathcal{W} \leq \mathcal{W}_{\mathrm{ent}} = 2$. Consider now that the lab observables $\{A_1, A_2\}$ and $\{B_1, B_2\}$ only nearly correspond (4) to the target observables $\{\sigma_X, \sigma_Z\}$. Since $\mathcal{W}_{\mathrm{ent}} = 2$ is algebraically maximal, it remains unchanged, but such is not the case for the separable bound $\mathcal{W}_{\mathrm{sep}}$. Thanks to the simplicity of $\mathcal{W}$, we can precisely evaluate $\mathcal{W}_{\mathrm{sep}}$ in the prevalent scenario when all measurement devices are equally inaccurate, i.e. $\varepsilon_x^{\mathrm{A}} = \varepsilon_y^{\mathrm{B}} = \varepsilon$. For a product state, we have $\mathcal{W} = \langle A_1\rangle\langle B_1\rangle + \langle A_2\rangle\langle B_2\rangle \leq \sqrt{\langle A_1\rangle^2 + \langle A_2\rangle^2}\sqrt{\langle B_1\rangle^2 + \langle B_2\rangle^2}$. Since the target measurements are identical on both sites and the factors are independent, they are optimally chosen equal. Then, it is easily shown that the optimal choice of Bloch vectors corresponds to aligning A_1 and A_2 (B_1 and B_2) to the extent allowed by ε. This leads to the following tight condition for entanglement detection (see Appendix A)

$$\mathcal{W}_{\mathrm{sep}}(\varepsilon) = 1 + 4\left(1 - 2\varepsilon\right)\sqrt{\varepsilon\left(1-\varepsilon\right)}, \tag{5}$$

when $\varepsilon \leq \frac{1}{2} - \frac{1}{2\sqrt{2}}$ and $\mathcal{W}_{\mathrm{sep}} = 2$ otherwise. Importantly, the derivative diverges at $\varepsilon \to 0^+$. Hence, a small ε induces a large perturbation in the ideal ($\varepsilon = 0$) separable bound. In the vicinity of $\varepsilon = 0$, it scales as $\mathcal{W}_{\mathrm{sep}} \sim 1 + 4\sqrt{\varepsilon}$. For example, $\varepsilon = 0.5\%$ leads to $\mathcal{W}_{\mathrm{sep}}(\varepsilon) \approx 1.28$, which eliminates over a quarter of the range in which standard entanglement detection is possible, indicating the relevance of false positives.

Secondly, consider the CHSH quantity for entanglement detection, namely $\mathcal{W} = \langle\sigma_X \otimes (\sigma_X + \sigma_Z)\rangle + \langle\sigma_Z \otimes (\sigma_X - \sigma_Z)\rangle$. Here, we have targeted observables optimal for violating the CHSH Bell inequality [23]. One has $\mathcal{W}_{\mathrm{sep}} = \sqrt{2}$ and $\mathcal{W}_{\mathrm{ent}} = 2\sqrt{2}$. In contrast to the previous example, the fact that all correlations from d-dimensional separable states constitute a subset of all correlations based on local hidden variables implies that entanglement can be detected for any value of ε. However, as we show in Appendix B through an explicit separable model that we conjecture to be optimal, this fact does not qualitatively improve the robustness of idealised ($\varepsilon = 0$) entanglement detection to small inaccuracies. We obtain

$$\mathcal{W}_{\mathrm{sep}} = 4\left(1-2\varepsilon\right)\sqrt{\varepsilon(1-\varepsilon)} + \sqrt{2 - 16\varepsilon\left(1-\varepsilon\right)\left(1-2\varepsilon\right)^2}, \tag{6}$$

when $\varepsilon \leq \frac{1}{2} - \frac{1}{2\sqrt{2}}$ and $\mathcal{W}_{\mathrm{sep}} = 2$ otherwise. For small ε, we find $\mathcal{W}_{\mathrm{sep}} \sim \sqrt{2} + 4\sqrt{\varepsilon}$. An inaccuracy of $\varepsilon = 0.5\%$ ensures $\mathcal{W}_{\mathrm{sep}} \gtrsim 1.67$, which eliminates nearly a fifth of the range in which standard entanglement detection is possible.

Interestingly, it is *a priori* not clear how small ε should impact standard entanglement witnessing as d increases. On the one hand, the impact ought to increase due to the increasing number of orthogonal directions in Hilbert space. On the other hand, it ought to decrease due to the growing distances in Hilbert space. For instance, the ε required to transform the computational basis into its Fourier transform scales as $\varepsilon = \frac{\sqrt{d}-1}{\sqrt{d}}$, which rapidly approaches unity. To investigate the trade-off between these two effects, we consider the d-dimensional generalisation of the simplest entanglement witness. Both subsystems are subject to the same pair of target measurements, namely the computational basis $\{|e_i\rangle\}_{i=1}^d$ and its Fourier transform $\{|f_i\rangle\}_{i=1}^d$, where $|f_i\rangle = \Omega|e_i\rangle$ with $\Omega_{jk} = \frac{1}{\sqrt{d}}e^{\frac{2\pi i}{d}jk}$. The witness is $\mathcal{W}^{(d)} = \sum_{i=1}^d \langle e_i, e_i|\rho|e_i, e_i\rangle + \langle f_i, f_i|\rho|f_i, f_i\rangle$. One has $\mathcal{W}_{\mathrm{sep}}^{(d)} = 1 + \frac{1}{d}$ and $\mathcal{W}_{\mathrm{ent}}^{(d)} = 2$ [24]. Allowing for measurement inaccuracy, we use an alternating convex search algorithm to numerically optimise over the lab measurements and

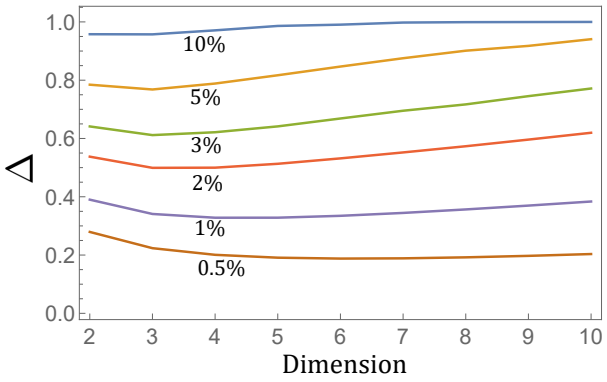

FIG. 1. Numerically obtained lower bounds on the relative magnitude of the entangled-to-separable gap, Δ, for entanglement witnessing based on two conjugate bases at different degrees of measurement inaccuracy $\varepsilon \in \{0.5\%, 1\%, 2\%, 3\%, 5\%, 10\%\}$.

shared separable states to obtain lower bounds on $\mathcal{W}_{\text{sep}}^{(d)}(\varepsilon)$. See Appendix C for details about the method. In order to compare the impact of measurement inaccuracy for different dimensions, we consider the following ratio between the entangled-to-separable gap in the inaccurate and ideal case, $\Delta \equiv \frac{\mathcal{W}_{\text{ent}}^{(d)}(0)-\mathcal{W}_{\text{sep}}^{(d)}(\varepsilon)}{\mathcal{W}_{\text{ent}}^{(d)}(0)-\mathcal{W}_{\text{sep}}^{(d)}(0)} = \frac{d}{d-1}\left[2-\mathcal{W}_{\text{sep}}^{(d)}(\varepsilon)\right]$. Notice that the numerator features $\mathcal{W}_{\text{ent}}^{(d)}(0)$ instead of $\mathcal{W}_{\text{ent}}^{(d)}(\varepsilon)$ because ε is not in itself a resource for the experimenter. The results of the numerics are illustrated in Figure 1 for some different choices of ε. We observe that Δ is not monotonic in d, but instead features a minimum, that shifts downwards in d as ε increases. Beyond this minimum point, the impact of measurement inaccuracies grows as the dimension becomes large.

Finally, for multipartite qubit states, it is natural to expect that the detrimental influence of small ε grows with the number of qubits under consideration. The reason is that measurement inaccuracies can accumulate separately in the different subsystems. This intuition is confirmed by the models of Ref. [7], in which small alignment errors are used to spoof, with increasing magnitude, the standard fidelity-based witness of genuine multipartite entanglement for Greenberger-Horne-Zeilinger states [25]. This further confirms the need of considering measurement inaccuracies.

High-dimensional entanglement criterion.— In view of the the relevance of small measurement inaccuracies, it is natural to formulate entanglement criteria that take them explicitly into account beyond the simplest, two-qubit, scenario. Consider a pair of d-dimensional systems and $n \in \{1, \ldots, d^2-1\}$ measurements. For system A, the observables ideally correspond to (subsets of) a generalised Bloch basis $\{\lambda_i\}_{i=1}^n$ and for system B, the ideal observables are the complex conjugates $\{\bar{\lambda}_i\}_{i=1}^n$. Here, λ_i is d-dimensional, traceless and satisfies $\text{tr}\left(\lambda_i \lambda_j^\dagger\right) = d\delta_{ij}$ [26]. Defining $\rho = \frac{1}{d}\left(\mathbb{1} + \sum_{i=1}^{d^2-1} \mu_i \lambda_i\right)$, one has $\|\vec{\mu}\|^2 \leq d-1$. A simple standard entanglement witness, based on a total of n measurements, is then given by

$$\mathcal{W}^{(d)} = \sum_{i=1}^{n} \langle \lambda_i \otimes \bar{\lambda}_i \rangle. \tag{7}$$

Using Hölder's inequality, one finds that separable states obey $\mathcal{W}_{\text{sep}}^{(d)} = d-1$. When the choice of Bloch basis is fixed, entangled states can achieve at most $\mathcal{W}_{\text{ent}}^{(d)} = \nu_{\max}\left[\sum_{i=1}^n \lambda_i \otimes \bar{\lambda}_i\right]$, by choosing the state as the eigenvector corresponding to the largest eigenvalue ($\nu_{\max}$). When the choice of Bloch basis is not fixed, a general upper bound for entanged states is $\mathcal{W}_{\text{ent}}^{(d)} \leq \min\left\{\sqrt{n\,(d^2-1)}, n(d-1)\right\}$, as shown in Appendix D. Note that $n(d-1)$ only is relevant when $d = 2$. Notice also that the maximally entangled state $|\phi_d^+\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle$ achieves $\mathcal{W}^{(d)} = n$ regardless of the choice of Bloch basis.

Consider now that the lab observables only nearly correspond to $\{\lambda_i\}$ and $\{\bar{\lambda}_i\}$ respectively. We write them as $A_i = q\lambda_i + \sqrt{1-q^2}\lambda_i^\perp$ and $B_i = q\bar{\lambda}_i + \sqrt{1-q^2}\bar{\lambda}_i^\perp$, where $q \in [-1,1]$ is related to the inaccuracy through $q = 1-2\varepsilon$ and $\lambda_i^\perp$ and $\bar{\lambda}_i^\perp$ are observables orthogonal to λ_i and $\bar{\lambda}_i$, respectively, on the generalised Bloch sphere. In Appendix D, we prove that the witness $\mathcal{W}^{(d)} = \sum_{i=1}^n \langle A_i \otimes B_i \rangle$ for separable states obeys

$$\mathcal{W}_{\text{sep}}^{(d)}(\varepsilon) \leq (d-1)\left(q + \sqrt{n-1}\sqrt{1-q^2}\right)^2, \tag{8}$$

when $q \geq \frac{1}{\sqrt{n}}$ and otherwise $\mathcal{W}_{\text{sep}}^{(d)}(\varepsilon) \leq n\,(d-1)$, which is algebraically maximal. As is intuitive, the window for detecting entanglement shrinks as ε increases.

We investigate the tightness of the bound. To this end, choose the state as $|\phi^\dagger\rangle \otimes |\phi^T\rangle$, where the local Bloch vector is $\mu_i = \frac{\sqrt{d-1}}{\sqrt{n}}$ and where $\lambda_i \to \lambda_i^\dagger$ ($\lambda_i \to \lambda_i^T$) for $|\phi^\dagger\rangle$ ($|\phi^T\rangle$). Choose the observables as $A_i = q\lambda_i + \sum_{j\neq i}\frac{\sqrt{1-q^2}}{\sqrt{n-1}}\lambda_j$ and $B_i = q\bar{\lambda}_i + \sum_{j\neq i}\frac{\sqrt{1-q^2}}{\sqrt{n-1}}\bar{\lambda}_j$. This returns the separable bound (8). However, we need to check that the Bloch vector $\vec{\mu}$ corresponds to a valid state. Curiously, for the most powerful case, namely $n = d^2-1$, tightness would be implied by a positive answer to the long-standing open question of whether there exists a Weyl-Heisenberg covariant symmetric informationally complete (SIC) POVM in dimension d. To see the connection, simply choose the Bloch basis as the non-Hermitian Weyl-Heisenberg basis $\{X^u Z^v\}$ for $u, v \in \{0, \ldots, d-1\}$ and $u+v > 0$, where $X = \sum_{k=0}^{d-1}|k+1\rangle\langle k|$ and $Z = \sum_{k=0}^{d-1} e^{\frac{2\pi i k}{d}}|k\rangle\langle k|$. It follows immediately that $|\langle\phi|X^u Z^v|\phi\rangle| = \frac{1}{\sqrt{d+1}}$, which defines a SIC-POVM. Since these SIC-POVMs are conjectured to exist in all dimensions [27], and are known to exist up to well above the first hundred dimensions [28, 29], our bound is plausibly tight for any d.

SDP methods.— We develop a hierarchy of SDP relaxations to bound the largest possible value of any linear witness, $\mathcal{W} = \sum_{a,b,x,y} c_{abxy} p(a,b|x,y)$, for some real coefficients c_{abxy}. The method applies both for correlations originating from entangled states and from separable states, under

any given degree of measurement inaccuracy and arbitrary target measurements. Thus, we systematically establish upper bounds $\mathcal{W}^{\uparrow}_{\text{ent}}(\varepsilon) \geq \mathcal{W}_{\text{ent}}(\varepsilon)$ and $\mathcal{W}^{\uparrow}_{\text{sep}}(\varepsilon) \geq \mathcal{W}_{\text{sep}}(\varepsilon)$. This has a three-fold motivation. Firstly, $\mathcal{W}_{\text{ent}}$ will generally depend on ε; cases with $\mathcal{W}^{(d)} > \mathcal{W}^{(d)}_{\text{ent}}(0)$ can be observed when the inaccuracies accumulate in a constructive way (e.g. a favourable systematic error in the local reference frames). It is relevant to bound such occurances. Secondly, knowledge of $\mathcal{W}^{\uparrow}_{\text{ent}}(\varepsilon)$ allows an experimenter to give lower bounds on the inaccuracy of the measurement devices. Thirdly, and most importantly, this enables a general and systematic construction of entanglement witnesses of the form $\mathcal{W} \leq \mathcal{W}^{\uparrow}_{\text{sep}}(\varepsilon)$.

We discuss the main features of the method for computing $\mathcal{W}^{\uparrow}_{\text{ent}}(\varepsilon)$ and then see how it can be extended to also compute $\mathcal{W}^{\uparrow}_{\text{sep}}(\varepsilon)$. To this end, as is standard, the SDP relaxation method is based on the positivity of a moment matrix. This matrix consists of traces of monomials (in the spirit of e.g. [30]) which are composed of products of the state, the lab measurements and the target measurements (see Appendix E for specifics). Moments corresponding to products of the first two can be used to build a generic linear witness $\mathcal{W}$ via Eq. (1). Moments corresponding to products of the final two can be used to build the constraints on the fidelities $\mathcal{F}^{\text{A}}_x$ and $\mathcal{F}^{\text{B}}_y$. Our construction draws inspiration from two established ideas. Firstly, one can capture the constraints of d-dimensional Hilbert space, on the level of the moment matrix, by numerically sampling states and measurements [31]. Secondly, in scenarios without entanglement, constraints capturing the fidelity of a quantum state with a target can be incorporated into the moment matrix [32]. We adapt the latter to entanglement-based scenarios and measurement fidelities as needed for Eq. (3). Details are given in Appendix E. We have applied this method, at low relaxation level, in several different case studies in low dimensions and frequently found that the obtained upper bounds coincide with those obtained from interior point optimisation routines. We note that the computational requirements for this tool can be much reduced since sampling-based symmetrisation methods of Ref. [33] can straightforwardly be incorporated.

To extend this method for the computation of $\mathcal{W}^{\uparrow}_{\text{sep}}(\varepsilon)$, we must incorporate constraints on the set of quantum states. Since the set of separable states is generally difficult to characterise (see e.g. [34]), we instead adopt an approach in which we use the ideal entanglement witness condition, $\mathcal{W} \leq \mathcal{W}_{\text{sep}}(0)$, which we may realistically assume to posses, in place of the set of separable states. Then, since the probabilities associated to performing the target measurements on the state explicitly appear in our moment matrix, we can introduce it as an additional linear constraint in our SDP. Hence, the optimisation is effectively a relaxation of the subset of entangled states for which the original entanglement witness holds. In fact, since the set of separable states is characterised by infinitely many linear entanglement witnesses, one can in this way continue to introduce linear standard witnesses to constrain the effective state space in the SDP and thus further improve the accuracy of the bound $\mathcal{W}^{\uparrow}_{\text{sep}}(\varepsilon)$. In Appendix E we exemplify the use of this method, in its basic version, using only a single witness constraint $\mathcal{W} \leq \mathcal{W}_{\text{sep}}(0)$ on the state space, and show that it returns non-trivial, albeit not tight, bounds for two simple entanglement witnesses for relevant values of ε.

Discussion.— We have introduced and investigated entanglement detection when the measurements only nearly correspond to those intended to be performed in the laboratory. We have shown the relevance of the concept, presented explicit entanglement witnesses that take measurement inaccuracy into account, and finally shown how SDP methods can be applied to these types of problems. These results are a step towards a theoretical framework for detecting entanglement based on devices that quantitatively benchmarked in an operationally meaningful and experimentally accessible manner.

Our work leaves several natural open problems. If given an arbitrary standard entanglement witness, how can we compute corrections due to the introduction of measurement inaccuracies? Our SDP method is a first step towards addressing this problem but better methods are necessary both in terms of computational cost and in terms of the accuracy of the separable bound. Moreover, for a given d, what is the smallest number of auxiliary global measurement settings needed to eliminate the diverging derivative for optimal standard entanglement witnesses under small measurement inaccuracy? In addition, can one extend our entanglement witnesses to witnesses of genuine higher-dimensional entanglement, e.g. by detecting the Schmidt number? Also, in this first work, we have focused on bipartite entanglement. It would be interesting to identify useful entanglement witnesses for multipartite states at bounded measurement inaccuracy. Finally, the framework proposed here for entanglement detection draws inspiration from ideas proposed in semi-device-independent quantum communications. Given that several frameworks for semi-device-independence recently have been proposed [32, 35–38], there may be other similarly inspired avenues for entanglement detection based on quantitative benchmarks.

ACKNOWLEDGMENTS

The authors thank Mateus Araújo for discussions. This project was supported by the Wenner-Gren Foundations, the Austrian Science Fund (FWF) through the projects Y879-N27 (START) and P 31339-N27 (Stand-Alone), JSPS Overseas Research Fellowships, and JST PRESTO Grant Number JPMJPR201A.

[1] O. Gühne and G. Tóth, Entanglement detection, Physics Reports **474**, 1 (2009).

[2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[3] N. Friis, G. Vitagliano, M. Malik, and M. Huber, Entanglement certification from theory to experiment, Nature Reviews Physics **1**, 72 (2019).

[4] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A **223**, 1 (1996).

[5] B. M. Terhal, Bell inequalities and the separability criterion, Physics Letters A **271**, 319 (2000).

[6] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Optimization of entanglement witnesses, Phys. Rev. A **62**, 052310 (2000).

[7] D. Rosset, R. Ferretti-Schöbitz, J.-D. Bancal, N. Gisin, and Y.-C. Liang, Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses, Phys. Rev. A **86**, 062325 (2012).

[8] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Device-independent witnesses of genuine multipartite entanglement, Phys. Rev. Lett. **106**, 250404 (2011).

[9] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Device-independent entanglement quantification and related applications, Phys. Rev. Lett. **111**, 030501 (2013).

[10] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[11] R. F. Werner, Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[12] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden–variable models for entangled quantum states, Journal of Physics A: Mathematical and Theoretical **47**, 424002 (2014).

[13] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, entanglement, nonlocality, and the einstein-podolsky-rosen paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[14] T. Moroder and O. Gittsovich, Calibration-robust entanglement detection beyond bell inequalities, Phys. Rev. A **85**, 032301 (2012).

[15] A. Tavakoli, A. A. Abbott, M.-O. Renou, N. Gisin, and N. Brunner, Semi-device-independent characterization of multipartite entanglement of states and measurements, Phys. Rev. A **98**, 052333 (2018).

[16] T. Moroder, O. Gittsovich, M. Huber, R. Uola, and O. Gühne, Steering maps and their application to dimension-bounded steering, Phys. Rev. Lett. **116**, 090403 (2016).

[17] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson, Experimental high-dimensional two-photon entanglement and violations of generalized bell inequalities, Nature Physics **7**, 677 (2011).

[18] M. Erhard, M. Krenn, and A. Zeilinger, Advances in high-dimensional quantum entanglement, Nature Reviews Physics **2**, 365 (2020).

[19] S. Ecker, F. Bouchard, L. Bulla, F. Brandt, O. Kohout, F. Steinlechner, R. Fickler, M. Malik, Y. Guryanova, R. Ursin, and M. Huber, Overcoming noise in entanglement distribution, Phys. Rev. X **9**, 041042 (2019).

[20] N. Herrera Valencia, V. Srivastav, M. Pivoluska, M. Huber, N. Friis, W. McCutcheon, and M. Malik, High-Dimensional Pixel Entanglement: Efficient Generation and Certification, Quantum **4**, 376 (2020).

[21] X.-M. Hu, W.-B. Xing, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, P. Erker, and M. Huber, Efficient generation of high-dimensional entanglement through multipath down-conversion, Phys. Rev. Lett. **125**, 090503 (2020).

[22] F. Bouchard, N. H. Valencia, F. Brandt, R. Fickler, M. Huber, and M. Malik, Measuring azimuthal and radial modes of photons, Opt. Express **26**, 31925 (2018).

[23] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23**, 880 (1969).

[24] C. Spengler, M. Huber, S. Brierley, T. Adaktylos, and B. C. Hiesmayr, Entanglement detection via mutually unbiased bases, Phys. Rev. A **86**, 022311 (2012).

[25] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, Experimental detection of multipartite entanglement using witness operators, Phys. Rev. Lett. **92**, 087902 (2004).

[26] R. A. Bertlmann and P. Krammer, Bloch vectors for qudits, Journal of Physics A: Mathematical and Theoretical **41**, 235303 (2008).

[27] G. ZAUNER, Quantum designs: Foundations of a noncommutative design theory, International Journal of Quantum Information **09**, 445 (2011), https://doi.org/10.1142/S0219749911006776.

[28] A. J. Scott, SICs: Extending the list of solutions (2017), arXiv:1703.03993v1, 1703.03993.

[29] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The sic question: History and state of play, Axioms **6**, 10.3390/axioms6030021 (2017).

[30] S. Burgdorf and I. Klep, The truncated tracial moment problem, Journal of Operator Theory **68**, 141 (2012).

[31] M. Navascués and T. Vértesi, Bounding the set of finite dimensional quantum correlations, Phys. Rev. Lett. **115**, 020501 (2015).

[32] A. Tavakoli, Semi-device-independent framework based on restricted distrust in prepare-and-measure experiments, Phys. Rev. Lett. **126**, 210503 (2021).

[33] A. Tavakoli, D. Rosset, and M.-O. Renou, Enabling computation of correlation bounds for finite-dimensional quantum systems via symmetrization, Phys. Rev. Lett. **122**, 070501 (2019).

[34] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Distinguishing separable and entangled states, Phys. Rev. Lett. **88**, 187904 (2002).

[35] T. Van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patrón, and S. Pironio, Semi-device-independent framework based on natural physical assumptions, Quantum **1**, 33 (2017).

[36] A. Tavakoli, E. Zambrini Cruzeiro, J. Bohr Brask, N. Gisin, and N. Brunner, Informationally restricted quantum correlations, Quantum **4**, 332 (2020).

[37] A. Tavakoli, E. Zambrini Cruzeiro, E. Woodhead, and S. Pironio, Informationally restricted correlations: a general framework for classical and quantum systems, Quantum **6**, 620 (2022).

[38] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C. C. W. Lim, Characterising the correlations of prepare-and-measure quantum networks, npj Quantum Information **5**, 17 (2019).

Appendix A: Simplest entanglement witness

Consider the entanglement witness $\mathcal{W} = \langle \sigma_X \otimes \sigma_X \rangle + \langle \sigma_Z \otimes \sigma_Z \rangle$ on a pair of qubits. We allow the lab observables to have an ε-deviation with respect to the target measurements $\{\sigma_X, \sigma_Z\}$ on both sites. This corresponds to the constraints

$$\mathrm{tr}\,(A_1\sigma_X) \geq 2 - 4\varepsilon, \qquad \mathrm{tr}\,(A_2\sigma_Z) \geq 2 - 4\varepsilon, \tag{A1}$$

$$\mathrm{tr}\,(B_1\sigma_X) \geq 2 - 4\varepsilon, \qquad \mathrm{tr}\,(B_2\sigma_Z) \geq 2 - 4\varepsilon, \tag{A2}$$

where we have chosen that all measurements are subject to the same magnitude of inaccuracy.

Due to the symmetry of $\mathcal{W}$ under a party swap, we can choose $A_1 = B_1$ and $A_2 = B_2$. Since the measurements are characterised by a pair of Bloch vectors, we can without loss of generality choose them in the XZ-plane of the Bloch sphere. We therefore write $A_k = B_k = \cos\theta_k \sigma_X + \sin\theta_k \sigma_Z$. In the relevant case of equality, the fidelity conditions then become

$$\theta_1 = -\arccos(1-2\varepsilon), \tag{A3}$$
$$\theta_2 = \arcsin(1-2\varepsilon). \tag{A4}$$

Due to the party symmetry, we can choose a product state on the form $|\phi\rangle \otimes |\phi\rangle$ where $|\phi\rangle = \cos z|0\rangle + \sin z|1\rangle$. Then we obtain

$$\mathcal{W} = 1 + 4(1-2\varepsilon)\sqrt{\varepsilon(1-\varepsilon)}\sin(4z), \tag{A5}$$

which is optimal at $z = \frac{\pi}{8}$ when $\varepsilon \leq \frac{1}{2}$. Hence

$$\mathcal{W}_{\text{sep}} = 1 + 4(1-2\varepsilon)\sqrt{\varepsilon(1-\varepsilon)}. \tag{A6}$$

Notice that this is only valid for $\varepsilon \leq \frac{1}{2} - \frac{1}{2\sqrt{2}}$. For larger ε we have $\mathcal{W}_{\text{sep}} = 2$.

Moreover, we note that the immediate generalisation of this witness, namely $\mathcal{W} = \langle \sigma_X \otimes \sigma_X \rangle + \langle \sigma_Y \otimes \sigma_Y \rangle + \langle \sigma_Z \otimes \sigma_Z \rangle$, in the presence of measurement inaccuracies, can by similar means be shown to admit the separable bound

$$\mathcal{W}_{\text{sep}} = 2 + 4\sqrt{2}(1-2\varepsilon)\sqrt{\varepsilon(1-\varepsilon)} - (1-2\varepsilon)^2, \tag{A7}$$

when $\varepsilon \leq \frac{3-\sqrt{3}}{6}$ and $\mathcal{W}_{\text{sep}} = 3$ otherwise.

Appendix B: Entanglement detection based on the CHSH quantity

Consider a pair of qubits, each of which is subject to two measurements. The target observables on both sites are σ_X and σ_Z. The lab observables all have the same inaccuracy bound ε. Thus we have

$$\operatorname{tr}(A_1\sigma_X) \geq 2 - 4\varepsilon, \qquad \operatorname{tr}(A_2\sigma_Z) \geq 2 - 4\varepsilon, \tag{B1}$$
$$\operatorname{tr}(B_1\sigma_X) \geq 2 - 4\varepsilon, \qquad \operatorname{tr}(B_2\sigma_Z) \geq 2 - 4\varepsilon. \tag{B2}$$

In case of perfect measurements, the CHSH quantity acts as a conventional entanglement witness,

$$\mathcal{W} = \langle A_1 \otimes B_1 \rangle + \langle A_1 \otimes B_2 \rangle + \langle A_2 \otimes B_1 \rangle - \langle A_2 \otimes B_2 \rangle \leq \sqrt{2}, \tag{B3}$$

which is respected by all separable states. Evidently, since $\mathcal{W} \leq 2$ for local hidden variable models, which in particular account for the statistics of any measurements performed on a separable state, it follows that entanglement can be detected for arbitrary ε.

We show the potential influence of small measurement inaccuracies through an explicit quantum model. Choose $A_1 = B_1$ and associate it to a Bloch vector $\vec{n}_1 = (\cos\alpha, 0, \sin\alpha)$ in the XZ-plane. Similarly choose $A_2 = B_2$ and associate it to the Bloch vector $\vec{n}_2 = (\cos\beta, 0, \sin\beta)$. Our strategy is to align the two Bloch vectors as much as possible under the constraints (B2). This implies the choice of

$$\alpha = \arccos(1-2\varepsilon), \qquad \beta = \arcsin(1-2\varepsilon). \tag{B4}$$

Then, we choose the product state $|\psi\rangle = |\phi\rangle \otimes |\phi\rangle$ with $|\phi\rangle = \cos z|0\rangle + \sin z|1\rangle$, where

$$z = -\frac{\pi}{4} + \frac{1}{4}\arctan\left(\frac{1}{8\varepsilon - 8\varepsilon^2 - 1}\right). \tag{B5}$$

The angle has been choosen so as to place the Bloch vector of $|\phi\rangle$ right in the middle of $\vec{n}_1$ and $\vec{n}_2$. This leads to the following value of the CHSH quantity,

$$\mathcal{W} = 4(1-2\varepsilon)\sqrt{\varepsilon(1-\varepsilon)} + \sqrt{2 - 16\varepsilon(1-\varepsilon)(1-2\varepsilon)^2}, \tag{B6}$$

when $\varepsilon \leq \frac{1}{2} - \frac{1}{2\sqrt{2}}$ and $\mathcal{W} = 2$ otherwise. The derivative diverges as $\varepsilon \to 0^+$, indicating the first-order impact of small measurement inaccuracies. For small ε, the value scales as $\mathcal{W} \sim \sqrt{2} + 4\sqrt{\varepsilon} - 4\sqrt{2}\varepsilon$. For example, if we choose $\varepsilon = 0.5\%$, the separable model achieves $\mathcal{W} = 1.67$ which is a perturbation comparable to that obtained in the main text for the simplest two-qubit entanglement witness.

Appendix C: Lower bounds: alternating convex search

Consider that we are given an arbitrary linear functional $\mathcal{W}$, arbitrary target measurements $\{\tilde{A}_{a|x}\}$ and $\{\tilde{B}_{b|y}\}$ and arbitrary measurement inaccuracies $\{\varepsilon_x^{\text{A}}, \varepsilon_y^{\text{B}}\}$. Consider a linear functional

$$\mathcal{W} = \sum_{a,b,x,y} c_{abxy} \operatorname{tr}\left[A_{a|x} \otimes B_{b|y}\rho\right], \tag{C1}$$

with some real coefficients c_{abxy}. We describe a numerical method, based on alternating convex search, to systematically establish lower bounds on both $\mathcal{W}_{\text{sep}}$ and $\mathcal{W}_{\text{ent}}$. To this end we consider latter case first.

In order to place a lower bound on $\mathcal{W}_{\text{ent}}$, we decompose the optimisation problem into three parts: one over the measurements on system A, one over the measurements on system B and one over the global shared state. To this end, we first choose a random set of measurements $\{B_{b|y}\}$ and a random pure state ρ. Then, we optimise $\mathcal{W}$ over the measurements $\{A_{a|x}\}$ under the constraint that $\mathcal{F}_x^{\text{A}} \geq 1 - \varepsilon_x^{\text{A}}$. This optimisation is a semidefinite program and can therefore be efficiently solved. Using the returned measurements $\{A_{a|x}\}$, we optimise $\mathcal{W}$ over the measurements $\{B_{b|y}\}$ under the constraint that $\mathcal{F}_y^{\text{B}} \geq 1 - \varepsilon_y^{\text{B}}$. This is again a semidefinite program. Finally, using the returned measurements $\{B_{b|y}\}$, we evaluate the Bell operator

$$\mathcal{B} = \sum_{a,b,x,y} c_{abxy} A_{a|x} \otimes B_{b|y} \tag{C2}$$

and compute its largest eigenvalue. The associated eigenvector is the optimal state, which corresponds to our choice of ρ. This routine of two semidefinite programs and one eigenvalue computation can then be iterated in order to find increasingly accurate lower bounds on $\mathcal{W}_{\text{ent}}$. The procedure depends on the initial starting point and ought therefore to be repeated several times independently.

To place a lower bound on $\mathcal{W}_{\text{sep}}$, we can proceed analogously to the above when treating the separate optimisations over the measurements $\{A_{a|x}\}$ and $\{B_{b|y}\}$. However, the optimisation over the state is now less straightforward since we require that $\rho = |\phi\rangle\langle\phi| \otimes |\psi\rangle\langle\psi|$. The optimisation over the state can be cast as another alternating convex search, treated as a sub-routine to the main alteranting convex search. In other words, we sample a random $|\phi\rangle$ and evaluate the semidefinite program optimising $\mathcal{W}$ over $|\psi\rangle$. Then, using the returned $|\psi\rangle$, we run a semidefinite program optimising $\mathcal{W}$ over $|\phi\rangle$. This procedure is iterated until desired convergence is obtained.

Appendix D: Bounds on witness

Let $\{\lambda_i\}_{i=1}^{d^2-1}$ be an orthonormal basis the space of operators acting on d-dimensional Hilbert space, with $\text{tr}\left(\lambda_i \lambda_j^\dagger\right) = d\delta_{ij}$. Then, every qudit state can be written as

$$\rho = \frac{1}{d}\left(\mathbb{1} + \vec{\mu}\cdot\vec{\lambda}\right), \tag{D1}$$

where $\vec{\mu}$ is some complex-valued Bloch vector with entries $\mu_i = \langle\lambda_i^\dagger\rangle = \text{tr}\left(\rho\lambda_i^\dagger\right)$. By checking the purity $\text{tr}\left(\rho^2\right)$, one finds that $\|\vec{\mu}\|^2 = \sum_{i=1}^{d^2-1}\langle\lambda_i^\dagger\rangle^2 \leq d-1$. In general, not every such Bloch vector corresponds to a valid density matrix.

Consider the witness

$$\mathcal{W}^{(d)} = \sum_{i=1}^{n}\langle\lambda_i\otimes\bar{\lambda}_i\rangle. \tag{D2}$$

For separable states, we can evaluate $\mathcal{W}_{\text{sep}}^{(d)}$ by restricting to product states. Then we have

$$\begin{aligned}\mathcal{W}^{(d)} &= \sum_{i=1}^{n}\langle\lambda_i\rangle_A\langle\bar{\lambda}_i\rangle_B \leq \sqrt{\sum_{i=1}^{n}\langle\lambda_i\rangle_A^2}\sqrt{\sum_{i=1}^{n}\langle\bar{\lambda}_i\rangle_B^2}\\ &\leq d-1 = \mathcal{W}_{\text{sep}}^{(d)}.\end{aligned} \tag{D3}$$

Notice that this is independent of n.

For entangled states, we have

$$\mathcal{W}^{(d)} \leq \sum_{i=1}^{n}\langle\lambda_i\otimes\bar{\lambda}_i\rangle = \nu_{\max}\left[\sum_{i=1}^{n}\lambda_i\otimes\bar{\lambda}_i\right] \tag{D4}$$

$$\leq n\max_i \nu_{\max}\left[\lambda_i\otimes\bar{\lambda}_i\right] = n\max_i\nu_{\max}\left[\lambda_i\right]^2 \leq n(d-1), \tag{D5}$$

where we used that $\nu_{\max}\left[\lambda_i\right] \leq \sqrt{d-1}$. However this, essentially trivial, bound is only tight for $d=2$, in which case it is algebraically maximal. To obtain a bound for $d>2$, we note that the entangled state lives in dimension d^2. Hence, its Bloch vector length is at most $\sqrt{d^2-1}$. In other words,

$$\sum_{i=1}^{n}\langle\lambda_i\otimes\bar{\lambda}_i\rangle^2 \leq d^2-1. \tag{D6}$$

Taking the case of equality, we obtain a bound on the largest value of the witness when all entries in the sum are equal. Thus we require

$$\langle\lambda_i\otimes\bar{\lambda}_i\rangle = \sqrt{\frac{d^2-1}{n}}, \tag{D7}$$

which gives

$$\mathcal{W}_{\text{ent}}^{(d)} \leq \sqrt{n}\sqrt{d^2-1}. \tag{D8}$$

This bound is not necessarily tight.

Consider now the case when we have separable states and inaccurate measurements. Expand $\mathcal{W}^{(d)}$ as follows,

$$\begin{aligned}\mathcal{W}^{(d)} =& \sum_{i=1}^{n}\langle A_i\otimes B_i\rangle = q^2\sum_{i=1}^{n}\langle\lambda_i\rangle_A\langle\bar{\lambda}_i\rangle_B\\ &+ q\sqrt{1-q^2}\sum_{i=1}^{n}\left(\langle\lambda_i\rangle_A\langle\bar{\lambda}_i^\perp\rangle_B + \langle\lambda_i^\perp\rangle_A\langle\bar{\lambda}_i\rangle_B\right)\\ &+ \left(1-q^2\right)\sum_{i=1}^{n}\langle\lambda_i^\perp\rangle_A\langle\bar{\lambda}_i^\perp\rangle_B.\end{aligned} \tag{D9}$$

We examine these sums one by one. From (D3), we see that the first sum is at most $d-1$. Next, we use the Cauchy-Schwarz inequality to write the second sum as

$$\begin{aligned}\sum_{i=1}^{n}\langle\lambda_i\rangle_A\langle\bar{\lambda}_i^\perp\rangle_B &\leq \sqrt{\sum_{i=1}^{n}\langle\lambda_i\rangle_A^2}\sqrt{\sum_{i=1}^{n}\langle\bar{\lambda}_i^\perp\rangle_B^2}\\ &\leq \sqrt{d-1}\sqrt{\sum_{i=1}^{n}\langle\bar{\lambda}_i^\perp\rangle_B^2} \leq (d-1)\sqrt{n-1}.\end{aligned} \tag{D10}$$

In the last step, we have used the following lemma. Let $\vec{u}\in\mathbb{R}^n$ and $\vec{v}^i\in\mathbb{R}^n$ be unit vectors such that the i'th component of $\vec{v}^i$ is zero, i.e. $\vec{v}_i^i = 0$. Then we have that

$$\sum_{i=1}^{n}\left(\vec{u}\cdot\vec{v}^i\right)^2 \leq \sum_{i=1}^{n}1-\vec{u}_i^2 = n-1. \tag{D11}$$

Again using the Cauchy-Schwarz inequality and this lemma also leads to

$$\sum_{i=1}^{n}\langle\lambda_i^\perp\rangle_A\langle\bar{\lambda}_i\rangle_B \leq (d-1)\sqrt{n-1}, \tag{D12}$$

$$\sum_{i=1}^{n}\langle\lambda_i^\perp\rangle_A\langle\bar{\lambda}_i^\perp\rangle_B \leq (d-1)(n-1). \tag{D13}$$

Putting it together, we arrive at the bound

$$\mathcal{W}_{\text{sep}} \leq (d-1)\left(n-1-q^2(n-2)+2q\sqrt{n-1}\sqrt{1-q^2}\right). \tag{D14}$$

Appendix E: Semidefinite relaxations

Consider the task of optimising an arbitrary linear functional over the set of projective quantum strategies with a given inaccuracy to a set of target measurements:

$$\begin{aligned}
&\mathcal{W}_{\text{ent}} = \max_{\{A_{a|x}\},\{B_{b|y}\},\rho} \mathcal{W}[p] \\
&\text{subject to} \quad \text{tr}\,(\rho) = 1, \qquad \rho \geq 0, \qquad \rho \in \mathcal{L}(\mathbb{C}^d) \\
&A_{a|x}A_{a'|x} = A_{a|x}\delta_{a,a'}, \qquad B_{b|y}B_{b'|y} = B_{b|y}\delta_{b,b'}, \\
&\sum_a A_{a|x} = \mathbb{1}_d, \qquad \sum_b B_{b|y} = \mathbb{1}_d \\
&\mathcal{F}_x^{\text{A}} \geq 1 - \varepsilon_x^{\text{A}}, \qquad \mathcal{F}_y^{\text{B}} \geq 1 - \varepsilon_y^{\text{B}} \\
&p(a,b|x,y) = \text{tr}\left[A_{a|x} \otimes B_{b|y}\rho\right], \qquad \text{(E1)}
\end{aligned}$$

where $\mathcal{L}(\mathbb{C}^d)$ is the set of linear operators of dimension d. This is generally a difficult optimisation problem. However, it can be relaxed into a hierarchy of increasingly precise criteria, each of which can be evaluated as a semidefinite program.

To this end, define the operator list

$$S = \{\mathbb{1}_{d^2}, \rho, \{A_{a|x}\}_{a,x}, \{B_{b|y}\}_{b,y}, \{\tilde{A}_{a|x}\}_{a,x}, \{\tilde{B}_{b|y}\}_{b,y}\}. \qquad \text{(E2)}$$

Here, the measurement operators are to be understood as spanning the full Hilbert space, e.g. $A_{a|x} \to A_{a|x} \otimes \mathbb{1}_d$. We let M_k denote the set of all monomials, taken from the list S, of degree at most k. We let $n(k)$ denote the size of the set M_k. Then, we define the $n(k) \times n(k)$ tracial moment matrix as

$$\Gamma(u;v) = \text{tr}\left(uv^\dagger\right), \qquad \text{(E3)}$$

for $u, v \in M_k$. A quantum model implies the positivity of Γ. Moreover, by including enough monomials, we can formulate the objective as a linear function in the moment matrix,

$$\mathcal{W}(\Gamma) = \sum_{a,b,x,y} c_{abxy}\Gamma(\rho A_{a|x}; B_{b|y}). \qquad \text{(E4)}$$

Similarly, the inaccuracy constraints can be formulated as the linear constraints

$$\begin{aligned}
&\frac{1}{d^2}\sum_{a=1}^{o}\Gamma(A_{a|x};\tilde{A}_{a|x}) \geq 1 - \varepsilon_x^{\text{A}}, \\
&\frac{1}{d^2}\sum_{b=1}^{o}\Gamma(B_{b|y};\tilde{B}_{b|y}) \geq 1 - \varepsilon_y^{\text{B}}. \qquad \text{(E5)}
\end{aligned}$$

In order to capture the constraints of d-dimensional Hilbert space and to fix the target measurements in the optimisation, we proceed as follows [31, 32]. We randomly sample ρ, $\{A_{a|x}\}_{a,x}$ and $\{B_{b|y}\}_{b,y}$ from a d-dimensional Hilbert space and construct the list S. Note that the target measurements are fixed at all times. Then, we evaluate the moment matrix and label it $\Gamma^{(1)}$. This process is repeated, leading to a list of sampled moment matrices $\{\Gamma^{(1)}, \ldots, \Gamma^{(m)}\}$. The sampling is terminated when the next moment matrix is found to be linearly dependent on all the previously sampled moment matrices. Thus, the sampled list constitutes a (non-orthonormal) basis of the space of moment matrices. We then define the total moment matrix as the affine combination

$$\Gamma = \sum_{i=1}^{m} s_i\Gamma^{(i)}, \qquad \sum_{i=1}^{m} s_i = 1, \qquad \text{(E6)}$$

where $\{s_i\}$ serve as optimisation variables.

We can now formulate our relaxation of the optimisation problem (E1) as $\mathcal{W}_{\text{ent}}(\varepsilon) \leq \mathcal{W}_{\text{ent}}^{\uparrow}(\varepsilon)$ where

$$\mathcal{W}_{\text{ent}}^{\uparrow} \equiv \max_{\{s_i\}} \mathcal{W}(\Gamma) \quad \text{subject to} \quad \Gamma \geq 0 \qquad \text{(E7)}$$

under the constraints (E5) and (E6). This can be evaluated as a semidefinite program. The relaxation becomes tighter as the list of monomials M_k is extended.

In order to instead obtain bounds of the form $\mathcal{W}_{\text{sep}}(\varepsilon) \leq \mathcal{W}_{\text{sep}}^{\uparrow}(\varepsilon)$, we can add the constraint

$$\sum_{a,b,x,y} c_{abxy}\Gamma(\rho\tilde{A}_{a|x};\tilde{B}_{b|y}) \leq \mathcal{W}_{\text{sep}}(0), \qquad \text{(E8)}$$

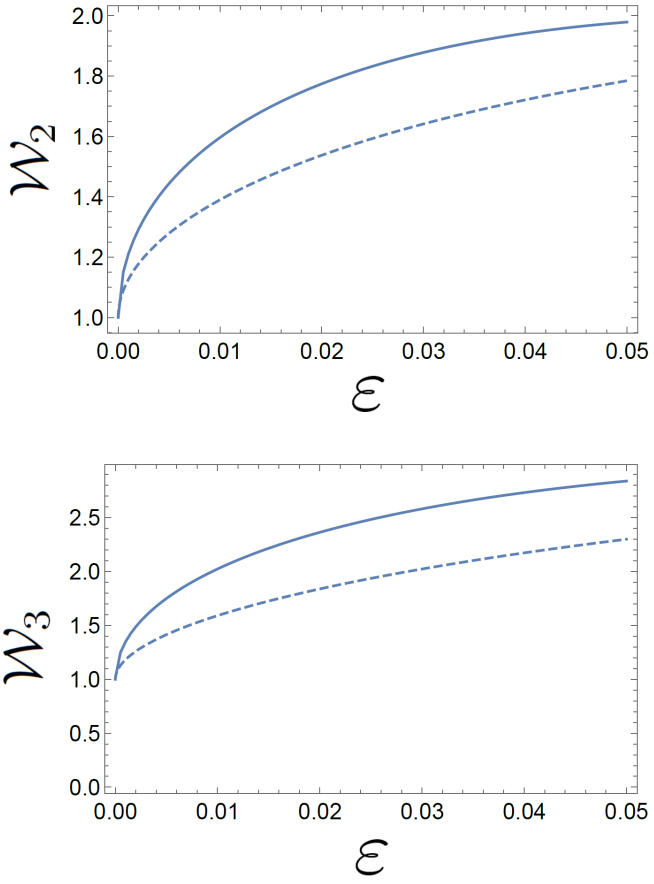

FIG. 2. **Solid lines.** Bounds $\mathcal{W}_{\text{sep}}^{\uparrow}(\varepsilon)$ obtained for the witness $\mathcal{W}_2$ and $\mathcal{W}_3$ via SDP relaxations where separability is relaxed to the set of entangled states obeying $\mathcal{W}_2(0) \leq 1$ and $\mathcal{W}_3(0) \leq 1$ respectively. These bounds can likely be made tighter by adding more ideal entanglement witness constraints to the SDP, in order to improve the relaxation of separability. **Dashed lines.** Optimal separable bound calculated analytically in Appendix A.

which corresponds to a standard entanglement witness. Note that we can introduce even more "target" measurements in the operator list S, thus extending the size $n(k)$ of the moment matrix, and then use them to build additional linear constraint like (E8) representing standard entanglement witnesses. The introduction of these shrinks the effective state space, thus improving the accuracy of the bound $\mathcal{W}_{\text{sep}}^{\uparrow}(\varepsilon)$, at the price of a larger SDP.

We exemplify a simple version of this method for the case of the two witnesses considered in Appendix A, namely $\mathcal{W}_2 = \langle \sigma_X \otimes \sigma_X \rangle + \langle \sigma_Z \otimes \sigma_Z \rangle \leq 1$ and $\mathcal{W}_3 = \langle \sigma_X \otimes \sigma_X \rangle + \langle \sigma_Y \otimes \sigma_Y \rangle + \langle \sigma_Z \otimes \sigma_Z \rangle \leq 1$, at inaccuracy ε. These are evaluated with monomial lists of length 46 and 89 respectively. The results are illustrated in Figure 2. As expected, the returned bounds are not tight, due to the basic relaxation of the separable set to all entangled states obeying $\mathcal{W} \leq 1$. Nevertheless, the bounds are non-trivial for relevant values of ε.

11 The shape of higher-dimensional state space: Bloch-ball analog for a qutrit

The following research article was published in *Quantum, Volume 5* the 29th June 2021 with the DOI 10.22331/q-2021-06-29-485.

11.1 Contribution

In this work, I created the self-dual model (Section 4). Further significant parts of the model properties were investigated by me, such as the behaviour under unitary transformations and the connection to classical probability distributions. Further I took care of the revising the manuscript, as substantial changes were necessary before acceptance.

The shape of higher-dimensional state space: Bloch-ball analog for a qutrit

Christopher Eltschka[1], Marcus Huber[2,3], Simon Morelli[3,2], and Jens Siewert[4,5]

[1]Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany

[2]Atominstitut, Technische Universität Wien, 1020 Vienna, Austria

[3]Institute for Quantum Optics and Quantum Information - IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

[4]Departamento de Química Física, Universidad del País Vasco UPV/EHU, E-48080 Bilbao, Spain

[5]IKERBASQUE Basque Foundation for Science, E-48009 Bilbao, Spain

22 June, 2021

Geometric intuition is a crucial tool to obtain deeper insight into many concepts of physics. A paradigmatic example of its power is the Bloch ball, the geometrical representation for the state space of the simplest possible quantum system, a two-level system (or qubit). However, already for a three-level system (qutrit) the state space has eight dimensions, so that its complexity exceeds the grasp of our three-dimensional space of experience. This is unfortunate, given that the geometric object describing the state space of a qutrit has a much richer structure and is in many ways more representative for a general quantum system than a qubit. In this work we demonstrate that, based on the Bloch representation of quantum states, it is possible to construct a three dimensional model for the qutrit state space that captures most of the essential geometric features of the latter. Besides being of indisputable theoretical value, this opens the door to a new type of representation, thus extending our geometric intuition beyond the simplest quantum systems.

1 Introduction

Nowadays virtually every student of quantum mechanics learns about the Bloch sphere and the Bloch ball as the geometrical representations of pure and mixed states of qubits, respectively. These objects have become indispensable for developing an intuition of elementary concepts such as basic quantum operations and the action of decoherence [1], or more advanced topics like the Majorana representation of symmetric multi-qubit states [2]. Nonetheless, the more experienced practitioner in the field of quantum mechanics is aware of various shortcomings of the Bloch ball if systems of higher dimension $d > 2$ are to be discussed. For example, the entire surface of the Bloch ball is covered by pure states, whereas most of the boundary of higher-dimensional state spaces is formed by mixed states. Those parts of the boundary may either be flat or curved, however, their curvature is different from that of the pure-state surface parts, which actually represents a single unitary orbit. Another important fact not shown by the Bloch ball is that not every orthogonal transformation can be applied to any quantum state: Whenever a state is not an element of the inscribed sphere of the state space of maximum radius there exist rotations in Bloch space that take it outside the state space and therefore are not allowed. Consequently, it is a long-standing question whether it is possible to construct a consistent three-dimensional model for higher-dimensional state spaces $\mathcal{Q}_d$, $d \geq 3$, that captures at least a part of these important geometric features.

Jens Siewert: jens.siewert@ehu.eus

2 Bloch ball for qubits

Let us briefly recapitulate the properties of the Bloch ball for qubits, $d = 2$. According to Fano [3, 4] density operators of qubits are parametrized by using the Pauli matrices σ_x, σ_y, σ_z and the identity $\mathbb{1}_2$

$$\rho = \frac{1}{2}\left(s_0\mathbb{1}_2 + x\,\sigma_x + y\,\sigma_y + z\,\sigma_z\right) \tag{1a}$$

$$s = \mathrm{Tr}\left(\sigma_s\rho\right) \quad , \tag{1b}$$

with real numbers $s \in \{x, y, z\}$, $|s| \leq 1$, and the normalization $s_0 = 1$. The state space $\mathcal{Q}_2$ of all qubit density operators is represented by a ball of radius 1 about the origin of $\mathbb{R}^3$, that is, each point (x, y, z) of this ball corresponds to exactly one state ρ. Pure states lie on the surface (forming a connected set), whereas mixed states are located inside the Bloch sphere, with the fully mixed state at the center.

Interestingly, the convex combinations of two states, $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ $(0 \le \lambda \le 1)$, are given by the straight line connecting the points of ρ_1 and ρ_2. A common choice of computational basis states is $\{|0\rangle, |1\rangle\}$, which correspond to north and south pole of the sphere and form the regular simplex Δ_1, the special case of Δ_{d-1} for $d = 2$. This shows that the vectors in $\mathbb{R}^3$ belonging to the basis states are *not* orthogonal. These vectors are orthogonal in $\mathbb{R}^4$ including the direction of s_0, but their projections into $\mathbb{R}^3$ lose this property.

Every density matrix can be obtained by unitarily rotating a diagonal density matrix, that resembles a classical probability distribution. Since for $d = 2$ the special unitary group SU(2) is the universal cover group of the rotation group SO(3), every rotation of the Bloch ball $Q_2 \subset \mathbb{R}^3$ has a corresponding unitary rotation in the state space. From this perspective, the Bloch ball is obtained by all possible rotations of the simplex Δ_1 in $\mathbb{R}^3$.

To the best of our knowledge, the idea that the parametrization in Eq. (1) entails a useful visualization for the non-unitary dynamics of a spin $\frac{1}{2}$, e.g., in a situation of radiation damping was put forward by Feynman and co-workers [5].

3 A Bloch-ball analog for a qutrit

Over the years much work has been done to elucidate the geometry of the state space of higher level systems, with special focus on the qutrit state space $\mathcal{Q}_3$ [6–25]. To develop an intuition for the full high-dimensional geometry of the qutrit state space, subsets, cross sections and projections onto two and three dimensions were extensively studied [6, 7, 10, 12, 15, 17, 18, 20–23] and multi-parameter representations of qutrit states were developed [14, 21, 24–26]. While these approaches can reproduce many geometric properties correctly, they do not give a global view of the Bloch body.

In this section we review known facts about the higher dimensional state space, with focus on dimension $d = 3$, collecting a list of requirements we wish our model to reproduce. We then construct a three dimensional global model of the state space, that reproduces astonishingly many properties of $\mathcal{Q}_3$.

3.1 Qutrit geometry

In analogy to qubits, density matrices describing the state of a d-level quantum system can be parametrized by the identity $\mathbb{1}_d$ and $d^2 - 1$ traceless Hermitian matrices [4, 6, 27]. So the quantum state space $\mathcal{Q}_d$ can be represented by a subset in $\mathbb{R}^{d^2-1}$, constrained by inequalities that arise from the positivity of the density operators [10, 11]. While it is still true that the quantum state space $\mathcal{Q}_d$ is obtained by rotating the simplex Δ_{d-1} in $\mathbb{R}^{d^2-1}$, it is no longer a sphere as not all rotations are allowed. In fact, for $d > 2$ the special unitary group SU(d) is a proper subgroup of the rotation group SO(d^2-1), meaning that the quantum state space $\mathcal{Q}_d$ is a proper subset of the d^2-1 dimensional (Hilbert-Schmidt) ball.

For qutrits the density matrices are parametrized by the identity $\mathbb{1}_3$ and the normalized Gell-Mann matrices X_j, Y_k $(j,k = 1,2,3)$, and Z_1, Z_2 (see Appendix),

$$\rho = \frac{1}{3}\left(\mathbb{1}_3 + \sum_{j=1}^{3} x_j \mathsf{X}_j + \sum_{k=1}^{3} y_k \mathsf{Y}_k + \sum_{l=1}^{2} z_l \mathsf{Z}_l\right) \tag{2a}$$

$$x_j = \mathrm{Tr}\left(\mathsf{X}_j\,\rho\right)\ ,\ y_k = \mathrm{Tr}\left(\mathsf{Y}_k\,\rho\right)\ ,\ z_l = \mathrm{Tr}\left(\mathsf{Z}_l\,\rho\right)\ , \tag{2b}$$

with real numbers x_j, y_k, and z_l, hence their space has eight dimensions. A Euclidean metric, corresponding to that of our everyday geometric experience, is induced in this space by the Hilbert-Schmidt norm, $\|\rho_1 - \rho_2\|_2 \equiv \sqrt{3\,\mathrm{Tr}\,[(\rho_1-\rho_2)^2]}$.

The geometric properties of $\mathcal{Q}_3$ we wish for a global model to reproduce were thoroughly discussed by Bengtsson et al. in Ref. [18].

i) Most importantly, $\mathcal{Q}_3$ is a convex set with the topology of a ball, so the model should share this characteristics. There are no pieces of lower dimension attached to it ("no hair" condition).

ii) The actual Bloch body is neither a polytope nor a smooth object.

iii) The Bloch body has an outer sphere of radius $\sqrt{2}$ and an inner sphere of radius $1/\sqrt{2}$ [28].

iv) The pure states (rank 1) form a connected set on the surface at maximum distance $\sqrt{2}$ from the completely mixed state $\frac{1}{3}\mathbb{1}_3$. Its measure is zero compared to that of $\mathcal{Q}_3$. In particular we aim at prominently displaying the three pure states corresponding to the preferred basis for the model.

v) Density matrices on the surface of $\mathcal{Q}_3$ are of rank 1 or 2, whereas states inside $\mathcal{Q}_3$ are of full rank (rank 3).

Finally, there are some additional properties specifically related to the nature of $\mathcal{Q}_3$ as a convex set:

vi) The set of quantum states is self-dual.

vii) All cross sections of $\mathcal{Q}_3$ do not have non-exposed faces. All corners of two-dimensional projections of $\mathcal{Q}_3$ are polyhedral.

3.2 A three dimensional model for a qutrit

Surprisingly, it is indeed possible to find an object in $\mathbb{R}^3$ representing $\mathcal{Q}_3$ that obeys most of the requirements in this list. In fact, there are (at least) two solutions with slightly different advantages. First we construct an object that fulfills properties i–iv) and also partially v). It represents a valid model

for $\mathcal{Q}_3$ that we call $Q_3^{(1)}$. Then we show that this first solution can be extended to a model $Q_3^{(2)}$ that possesses in particular also the property vi). Yet that model does not obey vii).

It is well known that the computational basis states in d dimensions form the corners of a regular simplex Δ_{d-1} in $\mathcal{Q}_d$ [6, 18], that is, for a qutrit a basis will be represented by Δ_2, an equilateral triangle. We insist that our model for $\mathcal{Q}_3$ displays and emphasizes one particular basis $\{|0\rangle, |1\rangle, |2\rangle\}$, because the mapping between physical states and points in the model will depend on this choice. Hence, we use the coordinates [cf. Eq. (2b)]

$$z_1 \;=\; \mathrm{Tr}\,(\mathsf{Z}_1\,\rho) \quad , \qquad z_2 \;=\; \mathrm{Tr}\,(\mathsf{Z}_2\,\rho) \quad ,$$

to *faithfully* represent the diagonal matrices as a simplex Δ_2 in the horizontal coordinate plane. In particular, the basis state $|2\rangle$ corresponds to $(z_1, z_2) = (0, -\sqrt{2})$, whereas the states $|0\rangle$ and $|1\rangle$ are located in $(\pm\sqrt{\frac{3}{2}}, \sqrt{\frac{1}{2}})$, respectively. The completely mixed state lies at the origin $(0, 0)$.

For the remaining six coordinates we are left with only one direction in $\mathbb{R}^3$. Here, we propose to use the coordinate

$$w \;=\; \sqrt{\sum_{j=1}^{3} x_j^2 + y_j^2} \quad , \tag{3}$$

which assumes only non-negative values. Hence we have our first model for $\mathcal{Q}_3$,

$$Q_3^{(1)} = \{(z_1, z_2, w) \in \mathbb{R}^3 \text{ s.t. } \text{Eqs. (2b), (3) } \forall \rho \in \mathcal{Q}_3\}. \tag{4}$$

The interpretation for the coordinates of a point $P = (z_1, z_2, w)$ is simple. Imagine the state ρ corresponding to P written as a sum of its diagonal and offdiagonal parts,

$$\rho \;=\; \rho_{\mathrm{diag}} \;+\; \rho_{\mathrm{offdiag}} \quad .$$

While the distance of P from the origin in the plane equals the Hilbert-Schmidt length $\sqrt{3\,\mathrm{Tr}(\rho_{\mathrm{diag}}^2) - 1}$ of the diagonal part of the Bloch vector for ρ, the vertical distance of P from the plane is the Hilbert-Schmidt length of this Bloch vector's offdiagonal part, $\sqrt{3\,\mathrm{Tr}(\rho_{\mathrm{offdiag}}^2)}$. The result is shown in Fig. 1.

3.3 Global geometric properties

Let us analyze the properties of this first model for $\mathcal{Q}_3$ with respect to our list of requirements. The surface of this object corresponds to a hemisphere of radius $\sqrt{2}$, where the three spherical segments beyond the triangle in the plane are cut off. Evidently, it is both convex and simply connected, without anything else attached to it. The flat surfaces of the cuts are connected with the smooth upper boundary by a sharp corner, so the object is neither smooth nor a polytope. Since our model preserves the length of the Bloch vector, it is still circumscribed by an outer sphere of radius $\sqrt{2}$ and center at the origin. The radius of the inner (hemi-)sphere coincides with that of the in-circle of the simplex Δ_2 in the horizontal plane and equals $1/\sqrt{2}$.

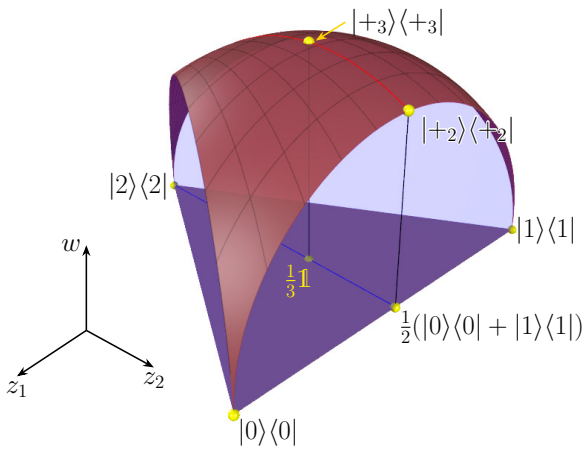

Figure 1: The model $Q_3^{(1)}$ of the qutrit Bloch body $\mathcal{Q}_3$ according to Eq. (4). The location of the basis states $|j\rangle$ is specified. The semicircular surface in the foreground is the image of the Bloch ball spanned by the states $|0\rangle$ and $|1\rangle$ and, in this representation, has the "north pole" $|+_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The point with the largest w coordinate is the image of the maximally coherent superposition $|+_3\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle)$. Note that all the pure states with $|\langle j|\psi\rangle|^2 = \frac{1}{3}$ get mapped to this point, in particular all the bases which are mutually unbiased with $\{|0\rangle, |1\rangle, |2\rangle\}$.

The spherical part of the surface above the ground plane corresponds to the set of pure states, hence to density matrices of rank 1. They form a simply connected surface of measure zero at maximal distance $\sqrt{2}$ from the origin. The cuts are half circles, in fact, these flat surfaces are the images of the three two-dimensional Bloch balls corresponding to the pairs of states $\{|0\rangle, |1\rangle\}$, $\{|0\rangle, |2\rangle\}$, and $\{|1\rangle, |2\rangle\}$ (cf. also Fig. 1). This means, these surfaces correspond to states of at most rank 2. It is understood that the base of $Q_3^{(1)}$ represents an artificial cut – similar to the base of the sculpture of a bust – that does not represent a boundary of $\mathcal{Q}_3$. The states on the other surfaces are of lower rank, all states of rank 3 reside in the interior of the model. However rank-2 states that have all of the basis vectors $|0\rangle$, $|1\rangle$, $|2\rangle$ in their span are located *inside* $Q_3^{(1)}$, although in $\mathcal{Q}_3$ they are part of the boundary. But the rank-2 states do not cover the entire interior of $Q_3^{(1)}$. Evidently, all points inside the inner sphere are of rank 3, since rank-2 states have purity of at least $\frac{1}{2}$. However the set of points corresponding only to rank-3 states is larger than that; see Sec. 3.4 for

details. Conversely, the interior points outside that set correspond both to rank-2 and rank-3 states. In particular, in agreement with property v), the images of rank-3 states cover the complete interior. Thus, we have established that our construction $Q_3^{(1)}$ has all the desired properties i–iv) of our list, property v) is at least partially satisfied.

We may ask how the coordinates in $Q_3^{(1)}$ are related to those of the actual state space $\mathcal{Q}_3$. Clearly our model is neither a projection nor a cross section. Instead, it resembles a representation in cylindrical coordinates: Our "diagonal coordinates" z_1, z_2 correspond to two longitudinal axes, whereas each $r_j = \sqrt{x_j^2 + y_j^2}$ (where $j = 1, 2, 3$) may be viewed as a radial coordinate belonging to mutually orthogonal directions. Our model does not display the polar angles and shows only the total radial distance.

3.4 Local algebraic properties

We have mentioned that the simplex Δ_2 represents the diagonal states faithfully, it is isomorphic to that set just as the Bloch ball is isomorphic to $\mathcal{Q}_2$. But our three-dimensional model $Q_3^{(1)}$ cannot represent all states of the eight-dimensional state space faithfully. In fact, a general point of $Q_3^{(1)}$ corresponds to infinitely many states and the simplex representing the basis states $|j\rangle\langle j|$, $j = 0, 1, 2$ is the only set of states with a unique preimage. States mapped onto the same point in the model belong to the equivalence class of states with the same diagonal entries and purity, they form a five-dimensional manifold in the original state space $\mathcal{Q}_3$. These subspaces are closed under the action of unitary operators that commute with Z_1 and Z_2, in particular diagonal unitary operators. As a consequence the model is invariant under these transformations.

The action of general unitary transformations, however, is displayed by our model. Since the purity remains unchanged, all points in the unitary orbit have the same distance from the origin. This orbit forms a subset of a sphere, its shape depends on the eigenvalues of the transformed state. As previously elucidated, a point in the interior represents a whole subspace of states with possibly different eigenvalues, so it makes little sense to talk about the orbit of a point in our model. However, diagonal states in the ground triangle are depicted faithfully and therefore uniquely identify a unitary orbit. One only needs to use the eigenbasis of the state to investigate which points can be reached by unitary transformations. The permutations of the diagonal entries correspond to SU(3) rotations, therefore the unitary orbit includes six points on the triangle (for non-degenerate eigenvalues). Since the eigenvalues of a matrix majorize the diagonal entries, the orbit remains on a sphere above this hexagon. Vice versa, every point on the sphere above the hexagon can

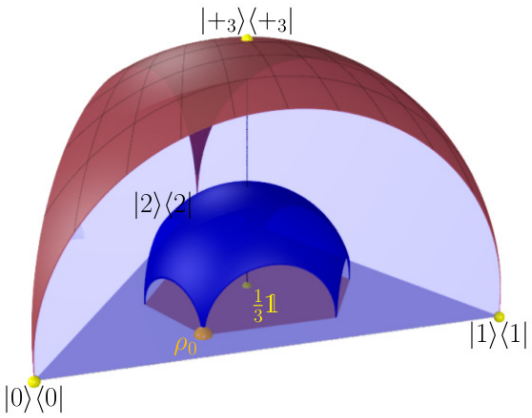

Figure 2: The SU(3) orbit of the state $\rho_0 = \frac{6}{10}|0\rangle\langle 0| + \frac{3}{10}|1\rangle\langle 1| + \frac{1}{10}|2\rangle\langle 2|$ is represented by the blue spherical surface. The image of ρ_0 together with the other five points in the z_1-z_2 plane forms the Birkhoff polytope (see text). The eigenvalue vectors of the diagonal states inside the polytope are majorized by that of ρ_0, $(\frac{6}{10}, \frac{3}{10}, \frac{1}{10})$.

be reached through a unitary transformation, see Fig. 2. If the eigenvalues are degenerate, the hexagon becomes a triangle. A special case are the pure states, the unitary orbit of a pure state coincides with the upper surface of $Q_3^{(1)}$.

This visualization establishes a direct connection to the action of doubly stochastic matrices on classical probability distributions. A classical probability distribution p majorizes exactly the probability distributions $M.\mathsf{p}$, where M is a doubly stochastic matrix. The set of distributions majorized by p is called the Birkhoff polytope [6]. The action of doubly stochastic matrices on classical probability distributions corresponds to the action of unitary transformations on normalized diagonal matrices. That is, for every doubly stochastic matrix M there exists a unitary matrix U, such that $M.\mathsf{p} = \mathrm{diag}(U.\mathsf{D}.U^\dagger)$, where D is a diagonal matrix with $\mathsf{p} = \mathrm{diag}(\mathsf{D})$. In the model this is visualized by unitarily transforming a diagonal state and then projecting it onto the ground triangle. This exactly reproduces the Birkhoff polytope, but it also generalizes it in the sense that the norm of the offdiagonal part remains visible.

The images of SU(3) orbits also allow us to identify the set of images of rank-2 states. The rank-2 diagonal states are exactly the sides of the base triangle. Therefore the rank-2 states are given by all the orbits of those points. It is readily seen that the boundary arcs connecting two triangle sides form half-circular cones that each have one of the basis states as apex, and are tangential to the insphere at their base. The points inside those cones, as well as the points inside the inner sphere, are not covered by those orbits, and therefore correspond only to states of rank 3. This rank-3 only region is convex; indeed it is the convex

hull of the interior of the insphere and the interior of the base triangle.

Note that the mirror image of this region and its boundary forms the lower part of our second model $Q_3^{(2)}$ which will be described in Sec. 4.

The map that takes the qutrit state space to our model is decidedly nonlinear, but amazingly retains some linear properties of the state space for the image. Any convex combination of two diagonal qutrit states $\delta = \lambda\delta_1 + (1-\lambda)\delta_2$ with $\lambda \in (0,1)$ is again a diagonal state and its image lies on the straight line connecting the image points of δ_1 and δ_2. But even the convex combinations $\rho = \lambda\sigma + (1-\lambda)\delta$ of an arbitrary qutrit state σ with a diagonal state δ are located on the straight line connecting σ and δ. The mixture of the diagonal parts is faithfully represented, but in this case, this holds also for the offdiagonal parts, because δ_{offdiag} simply vanishes. Consequently, for any problem involving two arbitrary states we can find a visualization of their mixture by means of a straight line: It is enough to consider the problem in the eigenbasis of one of the states.

Another noteworthy aspect of our model $Q_3^{(1)}$ is the separate treatment of diagonal and offdiagonal parts of the state. The diagonal matrices may be viewed as states in a classical probability space [18]. They become nonclassical by adding coherences, i.e., an offdiagonal part. The w coordinate in our model is the 2-norm for the offdiagonal part of the state. As the 1-norm of the offdiagonal part is established as a coherence measure [29] and the 1-norm is lower-bounded by the 2-norm, our model explicitly shows a lower bound to the coherence of a given quantum state. Correspondingly, the maximally coherent state $|+_3\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle)$ is located at the north pole of the hemisphere.

4 Self-dual qutrit geometry

The model we have discussed so far does not meet the requirements v) and vi) of our list. In particular it is not self-dual, that is, our three-dimensional Bloch vectors do not fulfill the following condition. Let $\vec{\xi}$ be the Bloch vector belonging to a state on the boundary of the state space. Then for the set of Bloch vectors $\vec{\eta}$ defining the dual hyperplanes enveloping the state space on the side opposite to $\vec{\xi}$ we have [18]

$$\vec{\xi}\cdot\vec{\eta} \;=\; -1 \quad . \tag{5}$$

Our model so far is not self-dual simply because for vectors $\vec{\xi}$ on the surface with $w > 0$ the dual planes do not touch the lower boundary of our Bloch body model. We will now demonstrate that the model can be extended by adding a "lower part" with coordinates $w < 0$ so that the entire object becomes self-dual.

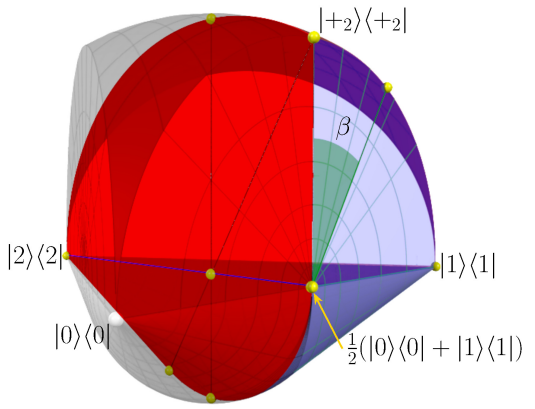

Figure 3: Constructing the dual of the surface of the model $Q_3^{(1)}$. Red: Cross section for determining the dual states $\tilde{\rho}(p)$ of the mixtures $\rho(p)$ in Eq. (8). Gray: We have included the part that has been cut from the full Bloch body (see Fig. 4, which shows the complete object $Q_3^{(2)}$ from the same point of view). By rotating the section highlighted in red by the angle β (see text) one obtains the lower conical part of $Q_3^{(2)}$.

4.1 Self-dual extension of $Q_3^{(1)}$

To achieve self-duality, consider first the pure states Π with $\mathrm{Tr}(\Pi^2) = \mathrm{Tr}(\Pi) = 1$. If we denote the vector of Gell-Mann matrices by $\vec{\mathsf{h}}$ and the Bloch vector of the state Π by $\vec{\pi}$, we can write $\Pi = \frac{1}{3}(\mathbb{1}_3 + \vec{\pi}\cdot\vec{\mathsf{h}})$. Recall that reversing the sign of all coordinates in the qubit Bloch sphere amounts to a point reflection operation at the maximally mixed state. This fact suggests that we might try reversing the sign of w in a process of reversing all of the coordinate signs, for example in an operation of the kind $\rho \longrightarrow \alpha\mathbb{1}_3 - \rho$ (where $\alpha > 0$). We have to make sure that the result is again positive. In Ref. [30] the so-called universal state inversion map $\mathcal{S}(\rho) = \nu(\mathbb{1} - \rho)$ was introduced; it guarantees that the result actually is a state. We choose the prefactor $\nu = \frac{1}{2}$ so as to normalize the resulting qutrit state and write

$$\mathcal{S}(\rho) \;\equiv\; \tilde{\rho} \;=\; \frac{1}{2}(\mathbb{1}_3 - \rho) \quad . \tag{6}$$

The inverted state $\tilde{\rho}$ is positive and for pure states we have $\mathrm{Tr}(\Pi\tilde{\Pi}) = 0$. Therefore,

$$0 \;=\; 1 \;+\; \vec{\pi}\cdot\vec{\tilde{\pi}} \quad , \tag{7}$$

that is, the Bloch vector $\vec{\tilde{\pi}}$ defines the dual plane for $\vec{\pi}$ (and vice versa). As the image of $\vec{\pi}$ determines a point in the spherical part of the boundary of our model, also the image of $\vec{\tilde{\pi}}$ lies on a spherical surface, however, with half the radius because of the prefactor $\nu = \frac{1}{2}$. The result is that the spherical part of our model $Q_3^{(1)}$ gets replicated in the region $w < 0$, whereat it is point-reflected at the origin and scaled down by a factor $\frac{1}{2}$.

This reasoning cannot be applied for the mixed boundary states of the model. Rather, we apply

state inversion $x_j \to -x_j$, $y_k \to -y_k$, $z_l \to -z_l$ and determine the boundary states by explicit calculation. Consider, for example, the mixtures (cf. Fig. 3)

$$\rho(p) \;=\; p\,|{+_2}\rangle\langle{+_2}| \;+\; \frac{1-p}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \qquad (8)$$

with $0 \le p \le 1$. It is straightforward to show that the dual states are given by

$$\begin{aligned}\tilde\rho(p) \;&=\; \frac{1}{3}\left(\mathbb{1}_3 - \frac{\sqrt{2}}{1+3p}\,\mathsf{Z}_2 - \frac{\sqrt{6}\,p}{1+3p}\,\mathsf{X}_1\right)\\ &=\; \frac{1}{3}\left(\mathbb{1}_3 - (1-q)\sqrt{2}\,\mathsf{Z}_2 - q\left[\frac{\sqrt{2}}{4}\,\mathsf{Z}_2 + \frac{\sqrt{6}\,p}{4}\,\mathsf{X}_1\right]\right)\end{aligned} \qquad (9)$$

with $q = \frac{4p}{1+3p}$. That is, also $\tilde\rho(p)$ lies on a straight line connecting $|2\rangle\langle 2|$ and the dual of $|{+_2}\rangle\langle{+_2}|$. An analogous calculation applies to all mixtures of $\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ with pure states on the circumference of the $\{|0\rangle, |1\rangle\}$ "Bloch sphere". Instead of considering the cross section in the z_2-w plane for the states $\tilde\rho(p)$ we can rotate this plane about the z_2 axis by an angle β (see Fig. 3), so that the pure state on the circumference becomes $\cos\left(\frac{\pi-2\beta}{4}\right)|1\rangle + \sin\left(\frac{\pi-2\beta}{4}\right)|0\rangle$.

The shape of the corresponding parts of the cross section remains the same as in the one shown in Fig. 3. Consequently, the dual states of the qubit Bloch ball $\{|0\rangle, |1\rangle\}$ are located on a half circular cone with apex in $|2\rangle\langle 2|$ (cf. Fig. 4). Due to the three-fold symmetry of $Q_3^{(1)}$ the regions near the other pure states $|0\rangle\langle 0|$, $|1\rangle\langle 1|$ can be constructed analogously.

Once we have found the shape [i.e., the boundary coordinates $w_{\max}(z_1, z_2)$ and $w_{\min}(z_1, z_2)$] of the self-dual object, we also need to provide a rule to decide whether a given state ρ from inside $\mathcal{Q}_3$ is mapped to the "upper" ($w > 0$) or the "lower" ($w < 0$) part of the object. In other words, from ρ we can determine the coordinates z_1, z_2, $|w|$ as well as $w_{\min}(z_1, z_2)$, but how can we decide about the sign of w?

In the following we describe a procedure to define this mapping $\rho \longrightarrow P(z_1, z_2, w)$. It is both straightforward and consistent (however, there might be other choices). Let ρ be an arbitrary qutrit state with smallest eigenvalue $\lambda_{\min}$. First we subtract as much of the fully mixed state so that one eigenvalue vanishes, i.e.,

$$\rho' \;=\; \frac{1}{1-3\lambda_{\min}}\,(\rho - \lambda_{\min}\mathbb{1}_3) \qquad (10)$$

is a rank-2 state with Bloch coordinates $(z_1', z_2', |w'|)$. Then

$$\left.\begin{aligned}|w'| &> |w_{\min}(z_1', z_2')|\\ |w'| &\le |w_{\min}(z_1', z_2')|\end{aligned}\right\} \;\Longrightarrow\; \begin{cases} w > 0\\ w < 0\end{cases}. \qquad (11)$$

This implies that in the case of equality the state ρ' is located at the lower boundary in $Q_3^{(2)}$. Moreover, together with ρ' also all its mixtures with $\frac{1}{3}\mathbb{1}_3$ belong to the lower part of $Q_3^{(2)}$, ensuring its compactness. The meaning of this definition is clear: In comparison to the structure of $Q_3^{(1)}$, in $Q_3^{(2)}$ we redistribute the location of the rank-2 states ρ'. If the purity of ρ' is large enough for its diagonal coordinates (z_1', z_2') it needs to remain in the upper part, $w > 0$, otherwise it goes to the lower part, $w < 0$.

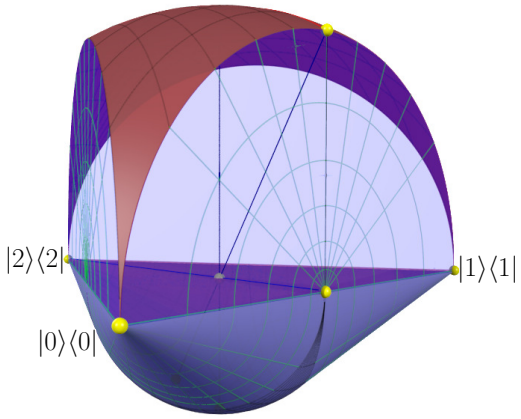

Figure 4: The self-dual model $Q_3^{(2)}$ of the qutrit state space $\mathcal{Q}_3$.

This concludes the construction of the self-dual model $Q_3^{(2)}$ for the qutrit Bloch body, see Fig. 4:

$$\begin{aligned}Q_3^{(2)} = \big\{&(z_1, z_2, w) \in \mathbb{R}^3 \text{ s.t.}\\ &\text{Eqs. (2b), (3), (11)} \forall \rho \in \mathcal{Q}_3\big\}\,. \end{aligned} \qquad (12)$$

We note that the extension to a self-dual object can be applied as well to the half-circle model $Q_2^{(1)}$ of the $d = 2$ Boch ball. Then, however, the rule given in Eq. (11) to determine the sign of w (without any addendum) does not improve the model. This is because in $d = 2$ all offdiagonal elements can be made real and positive by applying a single appropriate diagonal unitary to the state. In contrast, for $d > 2$ at most $2(d-1)$ offdiagonal matrix elements of a generic state can simultaneously be made positive by applying a diagonal unitary.

4.2 More convexity properties and relation to previous work

Convexity properties of the state space in the context of the Bloch-ball representation for $d = 2$ (and even for $d = 3$) were discussed early on, e.g., by Bloore [7] who analyzed the stratification of the state space with respect to the matrix rank. For a qutrit, all states of reduced rank $r < 3$ are part of the boundary. The four-dimensional set of rank-1 states $|\psi\rangle\langle\psi|$ forms the extremal states of the convex state space, as they do not have a convex decomposition into other states ρ_1 and ρ_2, $|\psi\rangle\langle\psi| \neq \lambda\rho_1 + (1-\lambda)\rho_2$ with $0 \le \lambda \le 1$.

Geometrically they are part of the outer sphere, because they have distance $\sqrt{2}$ from the origin.

As described in the previous section, the lower part of our self-dual model $Q_3^{(2)}$ consists of three half-circular cones, each with a basis state as their apex, which end at a half-sized point mirror image of the pure-state spherical section, that is, at the insphere. This description matches exactly the one of the internal boundary separating the points corresponding to rank-2 states from those corresponding only to rank-3 states in the model $Q_3^{(1)}$, except that it is below the base triangle rather than above. Also in the previous section, it was established that rank-2 states are mapped to the lower half exactly if doing so does not locate them outside the model.

This allows us to determine the image of the set of rank-2 states in $Q_3^{(2)}$. The only rank-2 states that fit into the lower part are exactly those which in $Q_3^{(1)}$ lie on the internal boundary between rank-2 states and rank-3 only states. Therefore all rank-2 states that get mapped into the lower half of the self-dual model are mapped onto the model's surface, while the interior of the lower part consists only of rank-3 states, just as criterion v) requires.

However, the same is not true in the upper part. Almost all the rank-2 states in the interior of $Q_3^{(1)}$ are not moved to the lower part and therefore are also in the interior of $Q_3^{(2)}$. Only the points on the internal rank boundary in $Q_3^{(1)}$ are moved to the boundary of $Q_3^{(2)}$, leaving only rank-3 states on the internal boundary. That is, criterion v) is still not perfectly satisfied. However, the new model $Q_3^{(2)}$ is self-dual by construction and hence obeys the criterion vi).

The last criterion vii) is critical because also the self-dual model $Q_3^{(2)}$ cannot reproduce these properties. This can be observed in the highlighted cross section in Fig. 3. The point corresponding to $\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ is non-exposed and counts as a non-exposed face of the two-dimensional cross section [18]. As these non-exposed faces occur at points where the self-dual extension is attached to $Q_3^{(1)}$ one might speculate that their occurrence is a price to pay for having such an extension. Moreover, the highlighted cross section in Fig. 3 coincides with the projection of $Q_3^{(2)}$ in the direction of the z_1 axis. Neither the corner corresponding to $|2\rangle\langle 2|$ nor that of $|+_2\rangle\langle +_2|$ is polyhedral.

The self-duality of $Q_3^{(2)}$ is particularly interesting, because it enables us to analyze how findings from previous work are featured in our model. As a first example we consider the duality of boundary states on outer and inner spheres. As was elucidated in Refs. [11, 20], boundary states on the outer sphere have their dual counterparts on the inner sphere, and vice versa. This property can be observed in Figs. 3 and 4 and is guaranteed by construction via the universal state inversion, Eq. (6), (7).

It is noteworthy that the universal state inversion (or reduction [31]) map finds an explicit geometric application here. Until now it was mainly associated with entanglement properties of multi-party states, although the relevance of geometrical concepts was implicit also in that context (cf., e.g., [32]).

One of the most notable qutrit three-sections is the so-called *obese tetrahedron* (or three-dimensional elliptope). It was already noted by Bloore [7] and thoroughly studied by Goyal and co-workers [20] (see Sec. 5.4 and Fig. 5 in Ref. [20]). Its barycenter is the completely mixed state and the corners are given by the four non-orthogonal states $\frac{1}{\sqrt{3}}(|0\rangle \pm |1\rangle \pm |2\rangle)$. The faces are bulgy, because the tetrahedron contains with its corner points also their duals on the opposite side. The image of the tetrahedron in $Q_3^{(2)}$ is the part of the vertical axis belonging to the model: The corner points are all mapped to the north pole, the barycenter to the origin and the centers of the faces to the point with $z_1 = z_2 = 0$ and $w = -\frac{1}{\sqrt{2}}$.

The last example we mention here is the conical three-section, Sec. 5.1 and Fig. 2 in Ref. [20]. Instead of the [128] section we consider the (unitarily equivalent) [138] section which indeed has a three-dimensional image in $Q_3^{(2)}$. The states of the cone are formed by all convex combinations of the real states of the $\{|0\rangle, |1\rangle\}$ Bloch ball (a circular disk) and the state $|2\rangle\langle 2|$. Therefore, one might expect the image to include for $w > 0$ the corresponding half-cone with the semicircular surface at $z_2 = \frac{1}{\sqrt{2}}$ and the apex at $z_2 = -\sqrt{2}$. However, this is only partially correct, because a part of the half-cone gets mapped to values $w < 0$ according to our rule, Eq. (11). This includes the complete half-cone attached to $|2\rangle\langle 2|$ for values $w < 0$ and $z_2 \leq -\frac{\sqrt{2}}{4}$. Moreover, also the half-cone with apex in the origin and base at $z_2 = -\frac{\sqrt{2}}{4}$ with base radius $\frac{\sqrt{6}}{4}$ gets mapped to $w < 0$. The union of both parts at $w < 0$ and $w \geq 0$ corresponds to the half-cone mentioned before, that is, to one half of the actual conic three-section.

5 Quantum channels

In order to give yet another demonstration that our method directly connects to the well-known properties of the Bloch sphere/ball visualization we show the action of three standard decohering channels (cf. Ref. [1]) on a qutrit. The depolarizing channel corresponds to driving a state towards the completely mixed state $\frac{1}{3}\mathbb{1}_3$, whereas phase-damping amounts to mixing a state with its diagonal part. Amplitude-damping describes the relaxation to one of the basis states. Following the discussion of $Q_3^{(1)}$ the action of these channels on the qutrit Bloch body

is evident and completely analogous to what is known for the $d = 2$ Bloch ball (cf. Fig. 5).

The channels are characterized by their Kraus operators K_j that obey the relation $\sum_j K_j^\dagger K_j = \mathbb{1}_3$. For the depolarizing channel we have $K_0 = \sqrt{1-8\gamma/9}\,\mathbb{1}_3$, $K_j = \sqrt{\gamma/9}\,\mathsf{h}_j$, where h_j $(j = 1 \ldots 8)$ stands as a shortcut for all the Gell-Mann matrices, so that

$$\rho'(\gamma) \;=\; (1-\gamma)\;\rho \;+\; \gamma\;\mathbb{1}_3 \quad . \tag{13}$$

The parameter $\gamma \in [0,1]$ describes the strength of the channel action.

The Kraus operators for phase damping (or pure dephasing) in its simplest version are $K_0 = \sqrt{1-2\gamma/3}\,\mathbb{1}_3$, $K_1 = \sqrt{\gamma/3}\;\mathsf{Z}_1$, $K_2 = \sqrt{\gamma/3}\;\mathsf{Z}_2$.

Under the action of the depolarizing and the phase-damping channel, the shrinking of the Bloch body model $Q_3^{(1)}$ occurs proportionally along straight lines, and the images of states that are represented by the same point will again be represented by the same point in the shrunken model. The same is, however, not true for the amplitude-damping channel.

The amplitude damping channel considered is of the form given by Grassl et al. [33], with $\gamma_{01} = \gamma_{02} =: \gamma$ and $\gamma_{12} = 0$, so that the states $|1\rangle$ and $|2\rangle$ are damped equally. That is,

$$\begin{aligned} K_0 &= |0\rangle\langle 0| + \sqrt{1-\gamma}\,|1\rangle\langle 1| + \sqrt{1-\gamma}\,|2\rangle\langle 2| \\ K_1 &= \sqrt{\gamma}\,|0\rangle\langle 1| \\ K_2 &= \sqrt{\gamma}\,|0\rangle\langle 2| \quad . \end{aligned} \tag{14}$$

This results in non-equal scaling of the non-diagonal elements. Whereas ρ_{01} and ρ_{02} scale with a prefactor of $\sqrt{1-\gamma}$, the element ρ_{12} gets a prefactor of $(1-\gamma)$.

This does not matter for diagonal states ($w = 0$), which are mapped into the model uniquely, and which are therefore again scaled linearly, by

$$z_1 \mapsto z_1' = \sqrt{\frac{3}{2}} + \left(z_1 - \sqrt{\frac{3}{2}}\right)(1-\gamma) \tag{15}$$

$$z_2 \mapsto z_2' = \frac{1}{\sqrt{2}} + \left(z_2 - \frac{1}{\sqrt{2}}\right)(1-\gamma) \tag{16}$$

Note that these two coordinates are scaled the same way for all the states.

For pure states, the absolute value of the offdiagonal elements is uniquely determined by the diagonal elements, and since only the absolute value enters into the w coordinate, again each point of the original model gets mapped to only one point of the shrunken model. In particular, the w coordinate of pure states gets mapped according to

$$\begin{aligned} w'^2 = (1-\gamma)^2 &\left[\frac{2}{3}\left(1 - z_2^2 - \frac{z_2}{\sqrt{2}}\right) \;+\; z_1\left(\frac{2}{\sqrt{3}}z_2 - \sqrt{\frac{2}{3}}\right)\right] \\ &+\; (1-\gamma)\left[\frac{4}{3} \;-\; z_1^2 + \left(\sqrt{\frac{2}{3}} - \frac{2}{\sqrt{3}}z_2\right)z_1 - \frac{z_2^2}{3} + \frac{\sqrt{2}z_2}{3}\right] \quad . \end{aligned} \tag{17}$$

For non-diagonal mixed states, states with different distribution of absolute values in the offdiagonal matrix elements get mapped to the same point in the model. Therefore states that are represented by the same point will get mapped by the channel to states represented by different points. It is obvious that the value of w will always be mapped somewhere into the range $[(1-\gamma)^2 w, (1-\gamma)w)$, but that range will not be exhausted for all source points.

6 Conclusion

We have presented two ways of visualizing the eight-dimensional state space of a qutrit in $\mathbb{R}^3$, thereby preserving a number of essential geometric properties of $\mathcal{Q}_3$, cf. Ref. [18]. Depending on the context one or the other representation may appear more useful. Our findings are relevant, because the properties of $Q_3^{(1)}$ and $Q_3^{(2)}$ are much closer to those of arbitrary higher-dimensional state space than the Bloch ball for qubits.

Thus we hope our results are helpful to develop a more precise intuition for objects in higher dimensions, which are commonly encountered in quantum information science.

It is conceivable that there are more possibilities for such visualizations along the lines of this work. The essential features of this visualization persist even if we use it to represent state spaces of dimensions $d > 3$ (splitting the generalized Gell-Mann into a diagonal and off-diagonal part in an identical fashion). Finding the most useful geometrical representation of those and categorizing them is a direction that we believe to yield further fruitful insight in the future.

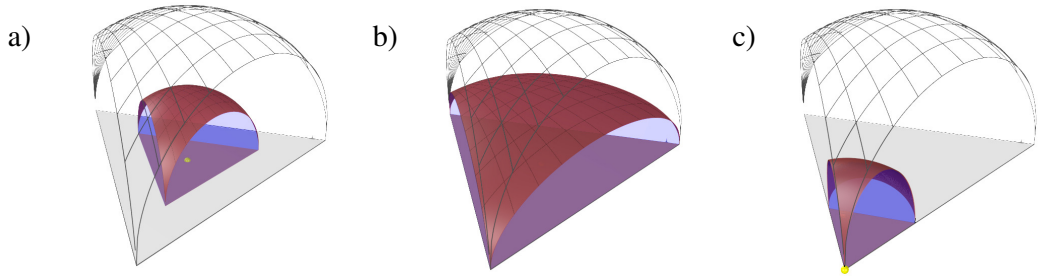

Figure 5: Action of quantum channels on the qutrit state space in $Q_3^{(1)}$ representation: a) depolarizing, b) phase-damping, and c) amplitude-damping channel.

Acknowledgments

This work was funded by the German Research Foundation Project EL710/2-1 (C.E., J.S.), by grant PGC2018-101355-B-100 (MCIU/FEDER,UE) and Basque Government grant IT986-16 (J.S.). C.E. and J.S. acknowledge Klaus Richter's support for this project. M.H. and S.M. acknowledge funding from the Austrian Science Fund (FWF) through the START project Y879-N27 and the stand-alone project P 31339-N27.

7 Appendix

For completeness, we provide the definition for the Gell-Mann matrices that we use in Eqs. (2a) and (9). Please note that we use a normalization to the dimension $d = 3$ (as opposed to the normalization to 2 conventionally used in high-energy physics). Moreover, also the enumeration differs from the usual one and is adapted to our coordinate description:

$$\mathsf{X}_1 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathsf{Y}_1 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathsf{X}_2 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \mathsf{Y}_2 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{bmatrix}$$

$$\mathsf{X}_3 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathsf{Y}_3 = \sqrt{\frac{3}{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{bmatrix}$$

$$\mathsf{Z}_1 = \sqrt{\frac{3}{2}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathsf{Z}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

References

[1] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[2] J. Martin, O. Giraud, P.A. Braun, D. Braun, and T. Bastin, *Multiqubit symmetric states with high geometric entanglement,* Physical Review A **81**, 062347 (2010).

[3] U. Fano, *A Stokes-Parameter Technique for the Treatment of Polarization in Quantum Mechanics,* Phys. Rev. **93**, 121 (1954).

[4] U. Fano, *Description of States in Quantum Mechanics by Density Matrix and Operator Techniques,* Rev. Mod. Phys. **29**, 74 (1957).

[5] R.P. Feynman, F.L. Vernon, and R.W. Hellwarth, *Geometrical Representation of the Schrödinger Equation for Solving Maser Problems,* J. Appl. Phys. **28**, 49 (1957).

[6] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement, 2nd edition* (Cambridge University Press, 2017).

[7] F.J. Bloore, *Geometrical description of the convex sets of states for systems with spin-1/2 and spin-1,* J. Phys. A **9**, 2059 (1976).

[8] G. Ramachandran and M.V.N. Murthy, *A new representation for the density matrix,* Nucl. Phys. A **323**, 403 (1979).

[9] M.S. Byrd and N. Khaneja, *Characterization of the positivity of the density matrix in terms of the coherence vector representation,* Phys. Rev.A **68**, 062322 (2003).

[10] G. Kimura, *The Bloch vector for N-level systems,* Phys. Lett. A **314**, 339 (2003).

[11] G. Kimura and A. Kossakowski, *The Bloch-vector space for N-level systems: the spherical-coordinate point of view,* Open Syst. Inf. Dyn. **12**, 207 (2005).

[12] I.P. Mendaš, *The classification of three-parameter density matrices for a qutrit,* J. Phys. A: Math. .Gen. **39**, 11313 (2006).

[13] L.J. Boya and K. Dixit, *Geometry of density matrix states,* Phys. Rev. A **78**, 042108 (2008).

[14] S. Vinjanampathy and A.R.P. Rau, *Bloch sphere like construction of SU(3) Hamiltonians using Unitary Integration,* J. Phys. A: Math. Theor. **42**, 425303 (2009).

[15] C.F. Dunkl, P. Gawron, J.A. Holbrook, J.A.

Miszczak, Z. Puchała, and K. Życzkowski, *Numerical shadow and geometry of quantum states,* J. Phys. A: Math. Theor. **44**, 335301 (2011).
[16] P. Kurzyński, *Multi-Bloch vector representation of the qutrit,* Quantum Inf. Comp. **11**, 361 (2011).
[17] G. Sarbicki and I. Bengtsson, *Dissecting the qutrit,* J. Phys. A: Math. Theor. **46**, 035306 (2012).
[18] I. Bengtsson, S. Weis, and K. Życzkowski, *Geometry of the Set of Mixed Quantum States: An Apophatic Approach* (Springer Basel, 2012) p. 175.
[19] G.N.M. Tabia and D.M. Appleby, *Exploring the geometry of qutrit state space using symmetric informationally complete probabilities,* Phys. Rev. A **88**, 012131 (2013).
[20] S. Goyal, B.N. Simon, R. Singh, and S. Simon, *Geometry of the generalized Bloch sphere for qutrits,* J. Phys. A: Math. Theor. **49**, 165203 (2016).
[21] P. Kurzyński, A. Kołodziejski, W. Laskowski, and M. Markiewicz, *Three-dimensional visualization of a qutrit,* Phys. Rev. A **93**, 062126 (2016).
[22] K. Szymański, S. Weis, and K. Życzkowski, *Classification of joint numerical ranges of three hermitian matrices of size three,* Lin. Alg. Appl. **545**, 148 (2018).
[23] J. Xie, A. Zhang, N. Cao, H. Xu, K. Zheng, Y.T. Poon, N.S. Sze, P. Xu, B. Zeng, and L. Zhang, *Observing Geometry of Quantum States in a Three-Level System,* Phys. Rev. Lett. **125**, 150401 (2020).
[24] S. Dogra, A. Vepsäläinen, and G.S. Paraoanu, *Majorana representation of adiabatic and superadiabatic processes in three-level systems,* Phys. Rev. Research **2**, 043079 (2020).
[25] G. Sharma and S. Ghosh, *Four-dimensional Bloch sphere representation of qutrits using Heisenberg-Weyl Operators,* arXiv:2101.06408 [quant-ph] (2021).
[26] E. Serrano-Ensástiga and D. Braun, *Majorana representation for mixed states,* Phys. Rev. A **101**, 022332 (2020).
[27] R.A. Bertlmann and P. Krammer, *Bloch vectors for qudits,* Journal of Physics A: Mathematical and Theoretical **41**, 235303 (2008).
[28] Note that we use a different normalization compared to that in Ref. [18] because of a different choice of prefactor for the Hilbert-Schmidt norm.
[29] T. Baumgratz, M. Cramer, and M.B. Plenio, *Quantifying Coherence,* Phys. Rev. Lett. **113**, 140401 (2014).
[30] P. Rungta, V. Bužek, C.M. Caves, M. Hillery, and G.J. Milburn, *Universal state inversion and concurrence in arbitrary dimensions,* Phys. Rev. A **64**, 042315 (2001).
[31] M. Horodecki and P. Horodecki, *Reduction criterion of separability and limits for a class of distillation protocols,* Phys. Rev. A **59**, 4206 (1999).
[32] C. Eltschka and J. Siewert, *Distribution of entanglement and correlations in all finite dimensions,* Quantum **2**, 64 (2018).
[33] M. Grassl, L. Kong, Z. Wei, Z. Yin, and B. Zeng, *Quantum Error-Correcting Codes for Qudit Amplitude Damping,* IEEE Transactions on Information Theory **64**, 4674 (2018).