Comparison of different definitions of the geometric measure of entanglement

Lin Chen,[1,2,*] Martin Aulbach,[3,†] and Michal Hajdušek[2,4,‡]

[1]*Department of Pure Mathematics and Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada*

[2]*Singapore University of Technology and Design, 20 Dover Drive, Singapore 138682*

[3]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542*

[4]*Department of Physics, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received 22 January 2014; published 7 April 2014)

Several inequivalent definitions of the geometric measure of entanglement (GM) have been introduced and studied in the past. Here we review several known and new definitions, with the qualifying criterion being that for pure states the measure is a linear or logarithmic function of the maximal fidelity with product states. The entanglement axioms and properties of the measures are studied, and qualitative and quantitative comparisons are made between all definitions. Streltsov *et al.* [New J. Phys. **12**, 123004 (2010)] proved the equivalence of two linear definitions of GM, whereas we show that the corresponding logarithmic definitions are distinct. Certain classes of states such as "maximally correlated states" and isotropic states are particularly valuable for this analysis. A little-known GM definition is found to be the first one to be both normalized and weakly monotonous, thus being a prime candidate for future studies of multipartite entanglement. We also find that a large class of graph states, which includes all cluster states, have a "universal" closest separable state that minimizes the quantum relative entropy, the Bures distance, and the trace distance.

DOI: 10.1103/PhysRevA.89.042305 PACS number(s): 03.67.Mn, 03.65.Ud

I. INTRODUCTION

Entanglement measures lie at the heart of quantum information theory, because they assess the usefulness of quantum states for tasks such as quantum teleportation, quantum computation, and cryptography protocols [1,2]. Numerous entanglement measures have been defined in the past, each of which may capture different properties of a state as a resource for certain tasks. One well-known measure is the geometric measure of entanglement (GM). Originally introduced for pure bipartite states [3,4], the GM was subsequently generalized to multipartite and to mixed states [5–7]. Two key benefits of GM are that it is an inherently multipartite entanglement measure and that it is comparatively easy to compute for many states [6,8–19].

The GM has a variety of operational interpretations: It assesses the usability of initial states for Grover's algorithm [20,21], the discrimination of quantum states under LOCC [10], the additivity and output purity of quantum channels [22] and the usefulness of states as resources for one-way quantum computation [8,23–25]. Further uses of GM include the construction and study of entanglement witnesses [6,26], the derivation of a generalized Schmidt decomposition [27], and the study of condensed matter systems, such as ground-state characterization and detection of phase transitions [28–30].

Several inequivalent definitions of GM have surfaced in the literature [5–7,9–12,20]. Regarding pure states, GM is expressed either as a linear or logarithmic function of the maximal fidelity with product states. Regarding mixed states, the pioneering papers did not agree on a unique definition, which led to the emergence of several inequivalent extensions of GM to mixed states. Although some of the GM definitions have been compared to other entanglement measures [11], a detailed comparison of all the different definitions of GM to each other has not been done before. An important milestone towards this goal was achieved by Streltsov *et al.* [13], who proved the equivalency of two frequently used definitions of GM.

In this paper we compare and characterize several known and new definitions of GM. The only qualifying criterion for an entanglement quantity to be regarded a GM definition is that on the subset of pure states it coincides with the well-defined linear or logarithmic GM.

Five known definitions, one little-known, and one new definition of GM are studied in this paper, first with respect to their entanglement axioms. This is followed by a quantitative and qualitative comparison of the definitions to each other. The "maximally correlated states" (as defined in Definition 11) turn out to be a particularly helpful class of states for this purpose. We also discover that a large class of graph states, including all cluster states, have a "universal" closest separable state that minimizes the quantum relative entropy, the Bures distance, and the trace distance.

The paper is organized as follows. Section II reviews some basic concepts of quantum information theory for later usage. In Sec. III the definitions of GM are introduced, and some preliminary results, e.g., with regard to entanglement axioms, are obtained. The subsequent Sec. IV closely examines the relationship between all six distinct definitions of GM from a variety of perspectives, and a hierarchy that allows for a partitioning of state space is obtained. A common closest separable state with respect to three different distance measures is presented for a large class of graph states in Sec. V. The concluding Sec. VI summarizes our results. For convenience, Tables I and II and Figs. 1 and 2 list some of our findings in compact form.

*linchen0529@gmail.com
†aulbach.martin@gmail.com
‡michal_hajdusek@sutd.edu.sg

TABLE I. The inequalities (35) facilitate a partitioning of state space $\mathcal{S}(\mathcal{H})$ into four subsets with clear physical meaning, and the inequalities (36) allow for a further division of the genuinely mixed entangled states $D = D_1 \cup D_2 \cup D_3$ into three subsets. Theorem 7 and Theorem 20 describe necessary and sufficient conditions for states belonging to D_2 and D_3, respectively.

Subset	Characterization of states in subset	Linear measures	Logarithmic measures
A	Pure separable	$0 = G^{f/c} = G^{m}$	$0 = G_l^{f} = G_l^{c} = G_l^{m}$
B	Pure entangled	$0 < G^{f/c} = G^{m}$	$0 < G_l^{f} = G_l^{c} = G_l^{m}$
C	Genuinely mixed separable	$0 = G^{f/c} < G^{m}$	$0 = G_l^{f} = G_l^{c} < G_l^{m}$
D_1			$0 < G_l^{f} < G_l^{c} < G_l^{m}$
D D_2	Genuinely mixed entangled	$0 < G^{f/c} < G^{m}$	$0 < G_l^{f} = G_l^{c} < G_l^{m}$
D_3			$0 < G_l^{f} < G_l^{c} = G_l^{m}$

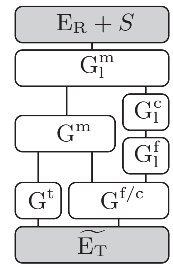

FIG. 1. The quantitative hierarchy of the different measures is shown, with the six distinct extensions of GM in white boxes. For a given state $\rho \in \mathcal{S}(\mathcal{H})$ the value of the measures increases monotonically from bottom to top along the vertical lines, and measures that are not vertically connected are not in an inequality relationship to each other. The quantities $\widetilde{E_T}$ and $E_R + S$ are not extensions of GM, but they provide lower and upper bounds, respectively.

II. PRELIMINARIES

First, we review some basic concepts for later usage, in particular axiomatic entanglement measures and distance measures. For more comprehensive reviews we refer to Refs. [1,2,31,32].

A. Axiomatic entanglement measures

Operationally motivated entanglement measures such as the entanglement cost and the distillable entanglement have clear physical meanings but tend to be difficult to study from a mathematical viewpoint, especially for multipartite systems. On the other hand, axiomatically motivated entanglement measures may not have operational implications.

Let $\mathcal{S}(\mathcal{H})$ be the space of operators acting on $\mathcal{H}$. Considering n parties $A_1,\ldots,A_n$ with joint Hilbert space $\mathcal{H} = \otimes_{j=1}^{n}\mathcal{H}_j$, a general n-partite state shared over the parties is described by a density matrix $\rho \in \mathcal{S}(\mathcal{H})$ acting on $\mathcal{H}$. Such a state is considered separable if it can be written in the form $\rho = \sum_i p_i \rho_i^1 \otimes \cdots \otimes \rho_i^n$, with $\sum_i p_i = 1$, and where ρ_i^j is a single-particle state of the j-th party. In the axiomatic approach [2,33,34], an entanglement measure is a functional $E : \mathcal{S}(\mathcal{H}) \to \mathbb{R}^+$ that satisfies two fundamental axioms as follows:

(1) $E(\rho) = 0$ if ρ is separable.

(2) E does not increase on average under local operations and classical communication (LOCC). Depending on which quantum operations are considered, this is defined as

(a) Weak monotonicity: $E(\rho) \geqslant E(\sigma)$, if $\rho \overset{\text{LOCC}}{\longmapsto} \sigma = \mathcal{E}(\rho) = \sum_i \vec{P}_i \rho \vec{P}_i^{\dagger}$.

TABLE II. Overview of the axioms fulfilled for the various definitions of GM. Subtable (a) lists the linear and logarithmic GM for pure states and whether axioms are fulfilled when considering quantum operations between pure states only. Subtable (b) lists the six distinct extensions of GM to mixed states. An axiom being fulfilled on pure states [as indicated in (a)] is necessary, but not sufficient for that axiom being fulfilled for mixed state extensions. The only exception is normalization, which is defined by pure states only. The properties of G^{t} and G_l^{f} have not been studied before, and we found that G_l^{c} satisfies Axiom 2(a) for two-qubit systems. For higher dimensions it is still unknown whether G_l^{c} satisfies Axiom 2(a).

(a) Pure states $|\Psi\rangle$

Properties	G	G_l
Axiom 1	✓	✓
Axiom 2(a)	✓	✓
Axiom 2(b)	✓	✗
Normalization	✗	✓

(b) Extensions to mixed states ρ

Properties	$G^{f/c}$	G^{m}	G^{t}	G_l^{f}	G_l^{c}	G_l^{m}
Axiom 1	✓	✗	✗	✓	✓	✗
Axiom 2(a)	✓	✗	✗	✓	?	✗
Axiom 2(b)	✓	✗	✗	✗	✗	✗
Convexity	✓	✗	✗	✓	✓	✗
Concavity	✗	✓	✗	✗	✗	✗

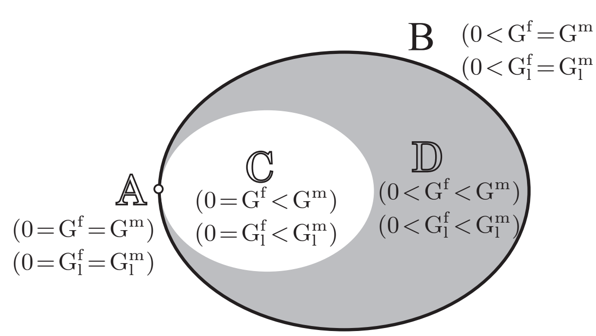

FIG. 2. Illustration of the partitioning of state space $\mathcal{S}(\mathcal{H}) = A \cup B \cup C \cup D$ into four pairwise disjoint sets. The pure states $(A \cup B)$ lie at the boundary, and the separable states $(A \cup C)$ form a closed, convex subset of $\mathcal{S}(\mathcal{H})$. The set a state ρ belongs to is uniquely determined by whether the inequalities in (35) are strict or equalities. By means of $\Lambda_f^2(\rho)$ and $\Lambda_m^2(\rho)$, the same partitioning can be facilitated by the logarithmic quantities G_l^{f} and G_l^{m}. When taking G_l^{c} into account, as well, the set of genuinely mixed entangled states $D = D_1 \cup D_2 \cup D_3$ is further subdivided into three pairwise disjoint sets.

(b) Strong monotonicity: $E(\rho) \geqslant \sum_i p_i E(\sigma_i)$, if $\rho \overset{\text{LOCC}}{\longmapsto} \sigma_i = \frac{\mathcal{E}_i(\rho)}{\operatorname{Tr}\mathcal{E}_i(\rho)}$ with probability $p_i = \operatorname{Tr}\mathcal{E}_i(\rho)$.

Here, the maps $\mathcal{E}$ and $\mathcal{E}_i$ stand for LOCC, and the elements $\{\vec{P}_i\}$ form a complete positive-operator valued measure, i.e., $\sum_i \vec{P}_i^\dagger \vec{P}_i = \mathbb{1}$. Weak monotonicity corresponds to trace-preserving quantum operations where the measurement outcome is unknown or discarded. Strong monotonicity corresponds to selective quantum operations, also known as measuring quantum operations [1]. Weak monotonicity is a special case of strong monotonicity that follows when a single outcome is obtained with probability 1, i.e., $\mathcal{E} = \mathcal{E}_1$. We refer to measures satisfying Axioms 1 and 2(a) as *weak entanglement measures* and measures that additionally satisfy Axiom 2(b) as *strong entanglement measures*.

Historically, strong monotonicity was required for entanglement measures [1,33,34], but in many cases weak monotonicity suffices [2]. Weak entanglement measures thus can be considered proper entanglement measures. Another historic requirement is that entanglement measures should coincide with the entropy of entanglement for pure bipartite states [1,34]. However, many popular measures fail this property, and the property cannot be easily extended to multipartite states, so it is not considered essential anymore. Invariance under local unitary (LU) operations is clearly important for entanglement measures, but it does not need to be stated as a separate axiom, because it automatically follows from weak monotonicity [1].

Apart from the axioms discussed above, many more desirable properties could be specified. Some common ones are the following:[1]

Normalization: $E(|\Phi\rangle^{\otimes n}) = n$ for 2 qubit Bell states $|\Phi\rangle$
Convexity: $E(\rho) \leqslant \sum_i p_i E(\rho_i)$ for all $\rho = \sum_i p_i \rho_i$
Additivity: $E(\rho^{\otimes 2}) = 2\,E(\rho)$ for all $\rho \in \mathcal{S}(\mathcal{H})$
Strong additivity: $E(\rho \otimes \sigma) = E(\rho) + E(\sigma)$ for all $\rho,\sigma \in \mathcal{S}(\mathcal{H})$.

The desirability of normalization is clear from the perception that Bell states carry 1 ebit of entanglement each. Convexity is motivated by the notion that entanglement should not increase under loss of information, namely when a selection of identifiable states ρ_i (right-hand side) is transformed into a mixture ρ (left-hand side) [1]. One may assume that this process can be physically realized by standard quantum operations, and thus strong monotonicity implies convexity. However, some additional properties (such as continuity) need to be satisfied, and the logarithmic negativity constitutes a counterexample by being a strong entanglement measure that is not convex [35]. An important consequence of convexity is that the measure can be maximized on the subset of pure states, i.e., there exist maximally entangled states (MES) that can be cast as pure states $\rho = |\psi\rangle\langle\psi|$.

If f and g are two convex functions and g is nondecreasing, then $g \circ f$ is also convex. For example, if $f(\rho) \geqslant 0$ is a convex measure, then $f^2(\rho)$ is also convex, using $g(x) = x^2$. In analogy to convexity, *concavity* is defined as $f(\rho) \geqslant \sum_i p_i f(\rho_i)$ for all $\rho = \sum_i p_i \rho_i$.

[1]For brevity, we abbreviate $f(|\psi\rangle\langle\psi|)$ as $f(|\psi\rangle)$ for functions $f(\rho)$ defined on $\mathcal{S}(\mathcal{H})$.

Regarding the additivity axioms, the tensor products in their definition have a specific physical meaning: Instead of enlarging the Hilbert space, $\rho \otimes \sigma$ refers to two states acting on the same Hilbert space [8]. If ρ and σ are both states of n d-level subsystems, then $\rho \otimes \sigma$ is a state of n d^2-level subsystems (instead of $2n$ d-level subsystems). Obviously, strong additivity implies additivity.

From a mathematical viewpoint, two entanglement measures, E_1 and E_2, are equivalent, if $E_1(\rho) = E_2(\rho)$ holds for all $\rho \in \mathcal{S}(\mathcal{H})$. A less restrictive, yet physically sound, criterion is the property of *ordering*: E_1 and E_2 have the same entanglement ordering if the same order is obtained when sorting all states by their amount of entanglement. This is the case if for all $\rho,\sigma \in \mathcal{S}(\mathcal{H})$ the two statements $E_1(\rho) > E_1(\sigma)$ and $E_2(\rho) > E_2(\sigma)$ are equivalent, i.e., they are either both true or both false. Entanglement measures that are equivalent will be denoted as $E_1 \equiv E_2$, and measures with the same ordering as $E_1 \cong E_2$.

Many different entanglement measures have been proposed, but here we only consider measures that are based on the distance to the set of separable states. The *relative entropy of entanglement* (REE) measures the minimum distance in terms of relative entropy between the given state ρ and the set of separable states (SEP) as follows:

$$E_{\mathrm{R}}(\rho) := \min_{\sigma \in \mathrm{SEP}} S(\rho|\sigma)\,, \tag{1}$$

where

$$S(\rho|\sigma) = \operatorname{Tr}\rho(\log\rho - \log\sigma) \tag{2}$$

is the *quantum relative entropy* [34]. Any state σ minimizing $S(\rho|\sigma)$ is called a *closest separable state* of ρ. Since the definition involves the minimization over all separable states, REE is known only for a few examples, such as bipartite pure states [34,36,37], Bell diagonal states [36,38,39], Werner states [40–42], maximally correlated states, isotropic states [38,39], generalized Dicke states [11,14,26], antisymmetric basis states [11,26], some graph states [10,43], the Smolin state, and Dür's multipartite entangled states [14,15]. A numeric method for computing REE of bipartite states has been proposed [34].

The REE can be applied to arbitrary multipartite states, pure or mixed. The central measure of this paper, the geometric measure of entanglement (GM)—to be discussed in Sec. III—is also an inherently multipartite measure, although its definition for mixed states is not unique.

B. Distance measures and fidelity

A good measure of distance $D(\rho,\sigma) : \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \to \mathbb{R}^+$ between two quantum states should be symmetric, zero if and only if $\rho = \sigma$, and observe weak monotonicity, which in this context means $D(\rho,\sigma) \geqslant D(\mathcal{E}(\rho),\mathcal{E}(\sigma))$ for any trace-preserving quantum operation $\mathcal{E}$ [2,31,36]. The last property, also known as *contractivity* under quantum operations, guarantees LU invariance: $D(\rho,\sigma) = D(U\rho U^\dagger, U\sigma U^\dagger)$. Any distance function with these properties is called a *distance measure*. One such distance measure is the trace distance [31],

$$D_{\mathrm{T}}(\rho,\sigma) = \frac{1}{2}\operatorname{Tr}|\rho-\sigma| = \frac{1}{2}\operatorname{Tr}\sqrt{(\rho-\sigma)^2} = \frac{1}{2}\sum_i |\lambda_i|\,, \tag{3}$$

where the λ_i are the eigenvalues of the matrix $(\rho - \sigma)$. For qubits $D_{\rm T}(\rho,\sigma)$ is equal to half the Euclidean distance between the corresponding Bloch vectors. The trace distance is convex in both arguments and, furthermore, is jointly convex [31],

$$D_{\rm T}\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) \leqslant \sum_i p_i D_{\rm T}(\rho_i,\sigma_i). \tag{4}$$

Another distance measure is the Bures distance and the closely related fidelity [31,44,45]. The Bures distance is

$$D_{\rm B}(\rho,\sigma) = \sqrt{2 - 2F(\rho,\sigma)}, \tag{5}$$

where $F(\rho,\sigma)$ is the fidelity between two states, defined as

$$F(\rho,\sigma) = {\rm Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}. \tag{6}$$

Alternative definitions in the literature are $\widetilde{D}_{\rm T} := D_{\rm T}^2$ for the trace distance, $\widetilde{D}_{\rm B} := D_{\rm B}^2$ for the Bures distance, and $\widetilde{F} := F^2$ for the fidelity.[2] The three necessary properties of distance measures outlined above remain invariant under exponentiation, and therefore $\widetilde{D}_{\rm T}$ and $\widetilde{D}_{\rm B}$ are also distance measures. Although closely related to the Bures distance, the fidelity is not a distance measure itself because $F(\rho,\sigma) = 1$ when $\rho = \sigma$. The fidelity is symmetric, unitarily invariant, concave in both arguments, jointly concave, and has codomain $F \in [0,1]$, with $F = 1$ if and only if $\rho = \sigma$ [31]. According to Uhlmann's theorem, the fidelity has a clear physical interpretation as the maximal overlap between all purifications of the input states [31,44].

If at least one of the two arguments of the fidelity is pure, then (6) simplifies to $F^2(\rho,\sigma) = {\rm Tr}(\rho\sigma)$, thus yielding

$$F^2(|\psi\rangle,\sigma) = |\langle\psi|\sigma|\psi\rangle|, \tag{7a}$$

$$F^2(|\psi\rangle,|\phi\rangle) = |\langle\psi|\phi\rangle|^2. \tag{7b}$$

In particular, for pure states the fidelity coincides with the Fubini-Study metric, the natural geometry on $\mathcal{H}$.

The fidelity also provides upper and lower bounds on the trace distance, with the lower bound increasing with the purity of the input states as follows [31]:

$$1 - F(\rho,\sigma) \leqslant D_{\rm T}(\rho,\sigma) \leqslant \sqrt{1 - F^2(\rho,\sigma)}, \tag{8a}$$

$$1 - F^2(|\psi\rangle,\sigma) \leqslant D_{\rm T}(|\psi\rangle,\sigma) \leqslant \sqrt{1 - F^2(|\psi\rangle,\sigma)}, \tag{8b}$$

$$D_{\rm T}(|\psi\rangle,|\phi\rangle) = \sqrt{1 - F^2(|\psi\rangle,|\phi\rangle)}. \tag{8c}$$

$$D_{\rm B}(|\psi\rangle,|\phi\rangle) = \sqrt{2 - 2F(|\psi\rangle,|\phi\rangle)}. \tag{8d}$$

Vedral *et al.* [36] found that from every distance measure $D(\rho,\sigma)$ a weak entanglement measure $E(\rho)$ can be constructed as

$$E(\rho) := \min_{\sigma \in {\rm SEP}} D(\rho,\sigma). \tag{9}$$

This construction directly yields weak entanglement measures from the trace distance (3) and Bures distance (5),

$$E_{\rm T}(\rho) = \min_{\sigma \in {\rm SEP}} D_{\rm T}(\rho,\sigma), \tag{10a}$$

$$E_{\rm B}(\rho) = \min_{\sigma \in {\rm SEP}} D_{\rm B}(\rho,\sigma), \tag{10b}$$

which we refer to as trace entanglement (TE) and Bures entanglement (BE), respectively. In analogy to $\widetilde{D}_{\rm T}$ and $\widetilde{D}_{\rm B}$, we define $\widetilde{E}_{\rm T} := E_{\rm T}^2$ and $\widetilde{E}_{\rm B} := E_{\rm B}^2$, and note that they are also weak entanglement measures. From (5) and the left-hand side of (8a), it follows that

$$\widetilde{E}_{\rm B}(\rho) \leqslant 2E_{\rm T}(\rho), \quad \text{for all } \rho \in \mathcal{S}(\mathcal{H}).$$

The quantum relative entropy (2) is not symmetric, and therefore not a proper distance measure. Nevertheless, by means of (9), it gives rise to the REE, which is a strong entanglement measure [34,47]. One could also ask whether the Hilbert-Schmidt distance $D_{\rm HS}(\rho,\sigma) = {\rm Tr}[(\rho - \sigma)^2]$, a metric in the mathematical sense, gives rise to an entanglement measure. However, this metric does not satisfy weak monotonicity, and it is an open question whether inserting $D_{\rm HS}$ in (9) yields an entanglement measure [48].

A simple, but important, mathematical inequality for this paper is

$$x - 1 \geqslant \log_b x \quad \forall\, x \in (0,1]\ \forall\, 1 < b \leqslant e, \tag{11}$$

where e denotes the base of the natural logarithm. We call (11) the *elementary inequality*. To demonstrate its usefulness, consider the two most common entropic quantities in quantum information theory, the *linear entropy* $M(\rho) = 1 - {\rm Tr}(\rho^2)$, and the *von Neumann entropy* $S(\rho) = -{\rm Tr}(\rho \log \rho)$. The linear entropy can be understood as an approximation of von Neumann entropy, obtained by the Taylor series $\log(\rho) \approx \rho - \mathbb{1}$, where $\mathbb{1}$ has the same range as ρ. For the commonly used logarithm bases 2 and e, (11) yields $\mathbb{1} - \rho \leqslant -\log\rho$, hence

$$M(\rho) \leqslant S(\rho), \quad \forall\, \rho \in \mathcal{S}(\mathcal{H}). \tag{12}$$

III. GEOMETRIC MEASURE OF ENTANGLEMENT

In this section we review the two common definitions of GM for pure states and introduce the known and unknown extensions to mixed states.

A. GM for pure states

The fundamental quantity for GM of pure states is

$$\Lambda^2(|\psi\rangle) := \max_{|\varphi\rangle \in {\rm PRO}} |\langle\varphi|\psi\rangle|^2, \tag{13}$$

where PRO denotes the set of fully product pure states of $\mathcal{H}$, henceforth referred to as product states. Comparing (13) with (7), we see that $\Lambda^2(|\psi\rangle)$ is the maximum fidelity between $|\psi\rangle$ and the set PRO. Furthermore, it is clear from (7) that the maximal value of $F(|\psi\rangle,\cdot)$ can be found among pure states as follows:

$$\Lambda^2(|\psi\rangle) = \max_{|\varphi\rangle \in {\rm PRO}} F^2(|\psi\rangle,|\varphi\rangle) = \max_{\sigma \in {\rm SEP}} F^2(|\psi\rangle,\sigma). \tag{14}$$

Any product state or separable state that maximizes the corresponding fidelity expression in (14) is called *closest product*

[2] This ambiguity of the definitions led to an incorrect definition of the Bures distance in Ref. [46].

state (CPS) or *closest separable state* (CSS), respectively. Note that the CPS or CSS is in general not unique. The relationship between the CSSs and CPSs of a given state $|\psi\rangle \in \mathcal{H}$ is seen from (7): If $\{|\phi_i\rangle\}$ is the set of CPSs, then any superposition $\sigma = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ with $\sum_i p_i = 1$ is a CSS. Conversely, if σ is a CSS, then there must exist a decomposition $\sigma = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ such that all $|\phi_i\rangle$ are CPSs.

The two common definitions of GM for pure states are [3,6]

$$G(|\psi\rangle) := 1 - \Lambda^2(|\psi\rangle), \tag{15a}$$

$$G_l(|\psi\rangle) := -\log \Lambda^2(|\psi\rangle), \tag{15b}$$

which we refer to as the *linear GM* and *logarithmic GM*, respectively. Unless denoted otherwise, the base 2 logarithm is used in this paper. Thanks to the elementary inequality (11), the results of this paper are also valid for any other logarithm base up to, and including, the base of the natural logarithm. For larger bases this is not the case, because (11) then no longer holds. Both G and G_l increase monotonically with Λ^2, and eliminating Λ^2 yields

$$G_l(|\psi\rangle) = -\log[1 - G(|\psi\rangle)] \quad \forall\, |\psi\rangle \in \mathcal{H}. \tag{16}$$

Due to this monotonic relationship, G and G_l have the same ordering for pure states. In particular, they have the same MESs. For bipartite states the MES (up to LU) is $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle$, yielding $G(|\Psi\rangle) = 1 - \frac{1}{d}$ and $G_l(|\Psi\rangle) = \log d$. For the simplest multipartite case of three qubits, the W state has been analytically determined as the MES for the GM [16]. For general multipartite systems, however, W states only yield low entanglement in terms of GM, and the identification of the MESs is an open problem. For the subset of permutation-symmetric states the MES are better understood [17–19], because the CPSs of symmetric states are symmetric themselves [49], thus considerably simplifying the optimization problem.

From (16) and the elementary inequality (11) it follows that

$$G(|\psi\rangle) \leqslant G_l(|\psi\rangle) \quad \forall\, |\psi\rangle \in \mathcal{H}. \tag{17}$$

Since $\Lambda^2(|\psi\rangle) = |\langle\psi|\psi\rangle|^2 = 1$ holds for all $|\psi\rangle \in$ PRO, entanglement Axiom 1 is satisfied for pure states for both definitions in (15). Regarding Axiom 2, an extension of G to mixed states that satisfies strong monotonicity is known [6], which automatically implies strong monotonicity of G on the subset of pure states. In contrast to this, an explicit counterexample ruling out strong monotonicity is known for G_l, and since this counterexample considers pure states only [11], no extension of G_l to mixed states can be strongly monotonous. However, we will later see that extensions of G_l with weak monotonicity do exist (cf. G_l^{f} defined in Sec. III B 3), and therefore G_l is weakly monotonous on the subset of pure states. The axiomatic properties of G and G_l are summarized in Table II in the left column.

Regarding the optional axioms, it is easy to verify that G_l is normalized, whereas G is not. This makes G_l the natural choice for quantitative studies of entanglement, such as scaling laws or comparison with other measures. The MES entanglement of n qubits ($n \geqslant 3$) scales linearly as $n - 2\log_2(n) - O(1) < G_l(|\Psi\rangle) < n - 1$ [8,23,50]. Restricting the computational coefficients to real values does not affect this scaling [8], but for *positive states* (i.e., states whose coefficients are all positive in the computational basis) the n-qubit MES are bounded by $G_l(|\Psi\rangle) \leqslant \frac{n}{2}$, and this bound is strict for even n (a trivial example being $n/2$ Bell pairs) [18]. On the other hand, symmetric n-qubit MESs scale logarithmically as $\log_2(n+1) - O(1) < G_l(|\Psi\rangle) < \log_2(n+1)$ [18,19]. These scaling laws readily generalize to qudits, leading to the conclusion that the MESs of sufficiently large multipartite systems are neither positive nor symmetric. Furthermore, since generic states are nearly maximally entangled with respect to GM [23], the above scaling laws can also be applied to random states.

Regarding the additivity axioms, neither G nor G_l are additive in general. For G this is obvious from its codomain [0,1], and for G_l it has been shown that states with a high amount of entanglement are not additive [8]. Nevertheless, many states of interest are additive or even strongly additive under G_l. In particular, it has been shown that positive states are strongly additive [8]:

Lemma 1. Let $|\psi\rangle \in \mathcal{H}$ be a positive state. Then $|\psi\rangle$ is strongly additive, i.e., $\Lambda^2(|\psi\rangle \otimes |\phi\rangle) = \Lambda^2(|\psi\rangle)\Lambda^2(|\phi\rangle)$ and $G_l(|\psi\rangle \otimes |\phi\rangle) = G_l(|\psi\rangle) + G_l(|\phi\rangle)$ holds for all $|\phi\rangle \in \mathcal{H}$.

Examples of positive states are multipartite Dicke states of arbitrary dimension and all bipartite pure states (by means of the Schmidt decomposition). Lemma 1 can be readily generalized to mixed states for an extension of GM that will be discussed in Sec. III B 2.

Note that the definition of the linear GM (15a) coincides with the Groverian entanglement measure $E_{\text{Gr}}(|\psi\rangle) = G(|\psi\rangle)^{1/2}$, which has a tangible operational interpretation by means of a quantum algorithm [7,20].

B. GM for mixed states

With the GM defined for pure states $|\psi\rangle \in \mathcal{H}$, we now consider the possible extensions to mixed states $\rho \in \mathcal{S}(\mathcal{H})$. Extensions of the linear GM will be labeled $G^{\text{x}}(\rho)$, and extensions of the logarithmic GM as $G_l^{\text{x}}(\rho)$, where x stands for a label to denote the extension. Any valid extension must coincide with (15a) or (15b) on the subset of pure states $\rho = |\psi\rangle\langle\psi|$. In other words, $G^{\text{x}}(|\psi\rangle\langle\psi|) = G(|\psi\rangle)$ and $G_l^{\text{x}}(|\psi\rangle\langle\psi|) = G_l(|\psi\rangle)$ must hold for all $|\psi\rangle \in \mathcal{H}$.

Since the expressions "pure" and "mixed" can be ambiguous, we briefly clarify their usage. From a mathematical viewpoint, $\sigma = |\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H})$ is a mixed state, but physically it is equivalent to the pure state $|\psi\rangle \in \mathcal{H}$. Therefore, we refer to both $|\psi\rangle$ and $\sigma = |\psi\rangle\langle\psi|$ as *pure* states. On the other hand, we refer to all states $\rho \in \mathcal{S}(\mathcal{H})$ as *mixed* states, so $\sigma = |\psi\rangle\langle\psi|$ can be regarded as pure and mixed. Mixed states that are not pure will be called *genuinely mixed*. Mathematically, a state $\rho \in \mathcal{S}(\mathcal{H})$ is genuinely mixed, if and only if $\text{rank}\,\rho \geqslant 2$.

One strategy to extend (15) to mixed states is to extend (13) to mixed states, i.e., to introduce a function $\Lambda_{\text{x}}^2(\rho) : \mathcal{S}(\mathcal{H}) \to \mathbb{R}^+$ with the property that $\Lambda_{\text{x}}^2(|\psi\rangle\langle\psi|) = \Lambda^2(|\psi\rangle)$ for all $|\psi\rangle \in \mathcal{H}$. The following lemma asserts the properties of extensions defined in that manner.

Lemma 2. Let $\Lambda_x^2(\rho):\mathcal{S}(\mathcal{H})\to\mathbb{R}^+$ be an extension of (13) to the set of all density matrices. Then the following holds for $G^x(\rho):=1-\Lambda_x^2(\rho)$ and $G_l^x(\rho):=-\log_2\Lambda_x^2(\rho)$:

(1) $G^x(\rho)\leqslant G_l^x(\rho)$ holds for all ρ.

(2) $G_l^x(\rho)=-\log_2(1-G^x(\rho))$ or, equivalently, $G^x(\rho)=1-2^{-G_l^x(\rho)}$ holds for all ρ. Furthermore, $G^x\cong G_l^x$.

(3) G^x satisfies Axiom 1 if and only if G_l^x does.

(4) G^x satisfies Axiom 2(a) if and only if G_l^x does.

(5) If G^x is convex, then G_l^x is also convex.

(6) If G_l^x is concave, then G^x is also concave.

Proof.

(1) This directly follows from the elementary inequality (11).

(2) The relationships between $G^{\mathrm{x}}(\rho)$ and $G_{\mathrm{l}}^{\mathrm{x}}(\rho)$ are obtained by eliminating $\Lambda_{\mathrm{x}}^2(\rho)$. Since $f(y)=-\log_2(1-y)$ and $g(y)=1-2^{-y}$ are monotonously increasing functions in $y\in[0,1]$, G^{x} and $G_{\mathrm{l}}^{\mathrm{x}}$ have the same ordering.

(3) For any $\rho\in$ SEP we have the following: $G^{\mathrm{x}}(\rho)=0\Leftrightarrow 1-\Lambda^2(\rho)=0\Leftrightarrow\Lambda^2(\rho)=1\Leftrightarrow-\log_2\Lambda^2(\rho)=0\Leftrightarrow G_{\mathrm{l}}^{\mathrm{x}}(\rho)=0$.

(4) For any $\rho\mapsto\sigma=\sum_i\vec{P}_i\rho\vec{P}_i^\dagger$ we have the following: $G^{\mathrm{x}}(\rho)\geqslant G^{\mathrm{x}}(\sigma)\Leftrightarrow 1-\Lambda^2(\rho)\geqslant 1-\Lambda^2(\sigma)\Leftrightarrow\Lambda^2(\rho)\leqslant\Lambda^2(\sigma)\Leftrightarrow-\log_2\Lambda^2(\rho)\geqslant-\log_2\Lambda^2(\sigma)\Leftrightarrow G_{\mathrm{l}}^{\mathrm{x}}(\rho)\geqslant G_{\mathrm{l}}^{\mathrm{x}}(\sigma)$.

(5) Let $G^{\mathrm{x}}(\rho)=1-\Lambda^2(\rho)$ be convex. Since $g(y):=-\log_2(1-y)$ is a convex nondecreasing function, $g(G^{\mathrm{x}}(\rho))=-\log_2\Lambda^2(\rho)$ is also convex.

(6) Let $G_{\mathrm{l}}^{\mathrm{x}}(\rho)=-\log_2\Lambda^2(\rho)$ be concave. Its additive inverse $-G_{\mathrm{l}}^{\mathrm{x}}(\rho)$ is therefore convex. Since $g(y):=2^y$ is a convex nondecreasing function, $g(-G_{\mathrm{l}}^{\mathrm{x}}(\rho))=\Lambda^2(\rho)$ is also convex. Therefore, $1-\Lambda^2(\rho)$ is concave. ■

With regard to items 5 and 6 of Lemma 2, it should be noted that convexity of G^{x} does not follow from convexity of $G_{\mathrm{l}}^{\mathrm{x}}$ and that concavity of $G_{\mathrm{l}}^{\mathrm{x}}$ does not follow from concavity of G^{x}. A counterexample for the latter case are the measures G^{m} and $G_{\mathrm{l}}^{\mathrm{m}}$ introduced in Sec. III B 2.

Several extensions of GM have been proposed in the past, and below we will introduce these as well as new ones. The first approach is based on a convex roof construction, akin to the entanglement of formation [1]. The second and third approach are based on extending the definition (13) to mixed states by means of the fidelity between the given state and the set of all pure or mixed states, respectively. Consequently, Lemma 2 applies to these two approaches. The fourth approach is to extend the linear GM by means of the trace distance (3).

1. Extension by convex roof: G^c/G_l^c

Based on definitions (13) and (15), the convex roofs of the linear and logarithmic GM are

$$G^{\mathrm{c}}(\rho):=\min_{\{p_i,|\psi_i\rangle\}}\sum_i p_i G(|\psi_i\rangle),\tag{18a}$$

$$G_{\mathrm{l}}^{\mathrm{c}}(\rho):=\min_{\{p_i,|\psi_i\rangle\}}\sum_i p_i G_{\mathrm{l}}(|\psi_i\rangle),\tag{18b}$$

where the minimum runs over all decompositions of $\rho=\sum_i p_i|\psi_i\rangle\langle\psi_i|$. Decompositions that minimize (18a) or (18b) will be called *optimal decompositions* and are labeled $\{P_i,|\Psi_i\rangle\}$. It is natural to ask whether for a given ρ the two functionals are minimized for the same decomposition. We will later show that for many states G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$ have the same optimal decompositions (e.g., for all isotropic states and two qubit states), but that there also exist states for which G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$ do not have a common optimal decomposition (e.g., for some maximally correlated states). Another open question is how many pure components $|\psi_i\rangle$ are necessary for an optimal decomposition of G^{c} or $G_{\mathrm{l}}^{\mathrm{c}}$. At least for G^{c} it is known that $(\dim\mathcal{H})^2$ pure components suffice [13].

Mathematically, any two decompositions $\{p_i,|\psi_i\rangle\}$ and $\{q_j,|\phi_j\rangle\}$ of the same ρ are related by a unitary matrix u_{ij} (i.e., $\sum_k u_{ki}^*u_{kj}=\delta_{ij}$), so

$$\sqrt{p_i}|\psi_i\rangle=\sum_j u_{ij}\sqrt{q_j}|\phi_j\rangle\tag{19}$$

holds for all i [31,51]. This identity will later be used in some proofs.

Regarding the entanglement axioms, Axiom 1 and convexity directly follow from the convex roof definitions for both G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$. The quantity G^{c} was first studied in detail in the seminal paper of Wei *et al.* [6] and found to be a strong entanglement measure. On the other hand, $G_{\mathrm{l}}^{\mathrm{c}}$ cannot be a strong entanglement measure [11]. However, it is an open question whether $G_{\mathrm{l}}^{\mathrm{c}}$ is weakly monotonous, i.e., whether $G_{\mathrm{l}}^{\mathrm{c}}(\rho)\geqslant G_{\mathrm{l}}^{\mathrm{c}}(\mathcal{E}(\rho))$ holds for all channels $\mathcal{E}$ and all $\rho\in\mathcal{S}(\mathcal{H})$. For $\rho\in$ SEP this is satisfied because of Axiom 1, and for all bipartite pure MES $|\Psi\rangle=\frac{1}{\sqrt{d}}\sum_{i=1}^d|ii\rangle$ this is also satisfied, because

$$G_{\mathrm{l}}^{\mathrm{c}}(\mathcal{E}(|\Psi\rangle))=\min_{\{p_i,|\psi_i\rangle\}}\sum_i p_i G_{\mathrm{l}}(|\psi_i\rangle)\leqslant\min_{\{p_i,|\psi_i\rangle\}}-\sum_i p_i\log\frac{1}{d}$$
$$=\log d=G_{\mathrm{l}}^{\mathrm{c}}(|\Psi\rangle).$$

In the two qubit case $G_{\mathrm{l}}^{\mathrm{c}}$ is a weak entanglement measure. In Appendix A we present a direct proof employing the concurrence and the optimal decomposition of the entanglement of formation [52]. The same result will later follow from a different line of argumentation.

Since the definitions (18) are not based on extending (13) to mixed states, Lemma 2 does not apply to G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$. In particular, we do not know if there exists an exact analytic relation between G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$ or if the two quantities have at least the same ordering (i.e., $G^{\mathrm{c}}\cong G_{\mathrm{l}}^{\mathrm{c}}$). The following lemma provides an analytic relation in the form of an inequality.

Lemma 3. For every state ρ the following holds:

(1) $G^c(\rho)\leqslant G_l^c(\rho)$.

(2) $G_l^c(\rho)\geqslant-\log_2[1-G^c(\rho)]$ or, equivalently, $G^c(\rho)\leqslant 1-2^{-G_l^c(\rho)}$.

Proof.

(1) This inequality readily follows from (17) and (18) for all ρ.

(2) Let $\{P_i,|\Psi_i\rangle\}$ be an optimal decomposition of ρ for $G_{\mathrm{l}}^{\mathrm{c}}(\rho)$. Then

$$G_{\mathrm{l}}^{\mathrm{c}}(\rho)=\sum_i P_i[-\log\Lambda^2(|\Psi_i\rangle)]\geqslant-\log\left[\sum_i P_i\Lambda^2(|\Psi_i\rangle)\right]$$
$$\geqslant-\log\left[\max_{\{p_i,|\psi_i\rangle\}}\sum_i p_i\Lambda^2(|\psi_i\rangle)\right]=-\log[1-G^{\mathrm{c}}(\rho)],$$

where the inequalities follow from that fact that $f(y) = -\log_2(y)$ is a convex and monotonically decreasing function. An analogous derivation yields $G^{\mathrm{c}}(\rho) \leqslant 1 - 2^{-G_{\mathrm{l}}^{\mathrm{c}}(\rho)}$. ■

Regarding the ordering of G^{c} and $G_{\mathrm{l}}^{\mathrm{c}}$, it will later be shown that the two measures do not have the same ordering in general (cf. Corollary 16).

2. Extension by trace inner product: G^m / G^m_l

Apart from the convex roof, the most widely studied extension of GM to mixed states is obtained by extending (13) to mixed states via the Hilbert-Schmidt inner product $\mathrm{Tr}(A^\dagger B)$, also known as trace inner product [31]. As proved below, this is equivalent to maximizing the fidelity between the input state and set of pure product states,

$$\Lambda_{\mathrm{m}}^2(\rho) := \max_{\sigma\in\mathrm{SEP}} \mathrm{Tr}(\rho\sigma) = \max_{|\varphi\rangle\in\mathrm{PRO}} \langle\varphi|\rho|\varphi\rangle = \max_{|\varphi\rangle\in\mathrm{PRO}} F^2(\rho,|\varphi\rangle). \tag{20}$$

Proof. The last equality is clear from (7a), and the $\geqslant$ part of the middle equality follows from the fact that the set of fully separable states contains the pure product states. The $\leqslant$ part of the middle equation follows as

$$\max_{\sigma\in\mathrm{SEP}} \mathrm{Tr}(\rho\sigma) = \mathrm{Tr}\left(\rho \sum_i p_i |\phi_i\rangle\langle\phi_i|\right) = \sum_i p_i \langle\phi_i|\rho|\phi_i\rangle$$
$$\leqslant \sum_i p_i \max_{|\varphi\rangle\in\mathrm{PRO}} \langle\varphi|\rho|\varphi\rangle = \max_{|\varphi\rangle\in\mathrm{PRO}} \langle\varphi|\rho|\varphi\rangle,$$

where $\sigma_{\mathrm{m}} = \sum_i p_i |\phi_i\rangle\langle\phi_i|$, with $|\phi_i\rangle \in \mathrm{PRO}$ for all i, is the separable state that maximizes $\mathrm{Tr}(\rho\sigma)$. ■

For pure states $\Lambda_{\mathrm{m}}^2(\rho)$ obviously coincides with $\Lambda^2(|\psi\rangle)$ from (13). Therefore, the functionals (15a) and (15b) are extended to mixed states as

$$G^{\mathrm{m}}(\rho) := 1 - \Lambda_{\mathrm{m}}^2(\rho), \tag{21a}$$

$$G_{\mathrm{l}}^{\mathrm{m}}(\rho) := -\log \Lambda_{\mathrm{m}}^2(\rho). \tag{21b}$$

Since Lemma 2 applies to these measures, there is some interdependence in their entanglement axioms. Indeed, neither G^{m} nor $G_{\mathrm{l}}^{\mathrm{m}}$ is an entanglement measure. This can be readily seen from that fact that the two measures attain their maximum for the maximally mixed state $\frac{\mathbb{1}}{\dim(\mathcal{H})}$, a separable state [12], thus violating Axiom 1 and Axiom 2(a).

With regard to convexity, the maximally mixed state is also a counterexample, because it can be decomposed into pure product states, $\mathbb{1} = \sum_i |i\rangle\langle i|$, with $G(|i\rangle) = G_{\mathrm{l}}(|i\rangle) = 0$. To check whether G^{m} or $G_{\mathrm{l}}^{\mathrm{m}}$ are concave, consider an arbitrary decomposition $\rho = \sum_i p_i\rho_i$. We have

$$\Lambda_{\mathrm{m}}^2(\rho) = \max_{|\varphi\rangle\in\mathrm{PRO}} \sum_i p_i \langle\varphi|\rho_i|\varphi\rangle \leqslant \sum_i p_i \max_{|\varphi_i\rangle\in\mathrm{PRO}} \langle\varphi_i|\rho_i|\varphi_i\rangle$$
$$= \sum_i p_i \Lambda_{\mathrm{m}}^2(\rho_i).$$

Therefore, $\Lambda_{\mathrm{m}}^2(\rho)$ is convex. From this it directly follows that $G^{\mathrm{m}}(\rho) = 1 - \Lambda_{\mathrm{m}}^2(\rho)$ is concave. It remains to investigate whether $G_{\mathrm{l}}^{\mathrm{m}}$ is also concave. Let us consider the isotropic state $\rho_{\mathrm{iso}} := p\mathbb{1}/d^2 + (1-p)|\Psi\rangle\langle\Psi|$, where $p \in [0,1]$ and $|\Psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |ii\rangle$ [53]. We easily see that $\Lambda_{\mathrm{m}}^2(\rho_{\mathrm{iso}}) = p/d^2 + (1-p)/d$. The concavity of the logarithm yields $G_{\mathrm{l}}^{\mathrm{m}}(\rho_{\mathrm{iso}}) \leqslant pG_{\mathrm{l}}^{\mathrm{m}}(\mathbb{1}/d^2) + (1-p)G_{\mathrm{l}}^{\mathrm{m}}(|\Psi\rangle)$, and the inequality is strict for $p \in (0,1)$. So the isotropic state is a counterexample to the concavity of $G_{\mathrm{l}}^{\mathrm{m}}$. To conclude, $G_{\mathrm{l}}^{\mathrm{m}}$ is neither convex nor concave.

Although not entanglement measures, the quantities G^{m} and $G_{\mathrm{l}}^{\mathrm{m}}$ are easier to calculate than other definitions of GM and have received a considerable amount of attention [8,10,12,14,26,50]. The quantity $G_{\mathrm{l}}^{\mathrm{m}}$ has been found to be closely related to the relative entropy of entanglement and the logarithmic global robustness of entanglement [8,11,18,26]. Furthermore, $G_{\mathrm{l}}^{\mathrm{m}}$ has been employed for the construction of optimal entanglement witnesses [26] and for the study of state discrimination under LOCC [12,26]. Zhu *et al.* [8] calculated $\Lambda_{\mathrm{m}}^2(\rho)$ for many states of interest, and Jung *et al.* [50] found that tracing out one subsystem of an n-partite pure state does not change this quantity, i.e., $\Lambda_{\mathrm{m}}^2(|\psi\rangle) = \Lambda_{\mathrm{m}}^2(\rho)$, with $\rho = \mathrm{Tr}_i(|\psi\rangle\langle\psi|)$, holds for all $|\psi\rangle \in \mathcal{H}$ and all $1 \leqslant i \leqslant n$.

The quantity $G_{\mathrm{l}}^{\mathrm{m}}$ allows to generalize Lemma 1 to mixed states [8]. A density matrix is called positive if all its entries in the computational basis are positive.

Lemma 4. Let $\rho \in \mathcal{S}(\mathcal{H})$ be a positive state. Then ρ is strongly additive, i.e., $\Lambda^2(\rho\otimes\sigma) = \Lambda^2(\rho)\Lambda^2(\sigma)$ and $G_l^m(\rho\otimes\sigma) = G_l^m(\rho) + G_l^m(\sigma)$ holds for all $\sigma \in \mathcal{S}(\mathcal{H})$.

This lemma has been employed to show the strong additivity of many mixed states, such as mixtures of Dicke states, Bell diagonal states, isotropic states, multiqubit Dür states, and the Smolin state [8]. The additivity problem of $G_{\mathrm{l}}^{\mathrm{m}}$ is closely related to that of the relative entropy and the logarithmic global robustness [26], which facilitated the study of additivity under these two entanglement measures, as well [8].

3. Extension by fidelity: G^f / G^f_l

We can also extend Λ^2 to mixed states by means of the fidelity as follows:

$$\Lambda_{\mathrm{f}}^2(\rho) := \max_{\sigma\in\mathrm{SEP}} F^2(\rho,\sigma). \tag{22}$$

This quantity has been previously studied [7,9,13,20,46,54] and has been described as *fidelity of separability* [13]. In the bipartite case (22) is equivalent to the so-called maximum k-extendible fidelity of a state in the limit $k \to \infty$ [54]. The maximum k-extendible fidelity has an operational interpretation as the maximum probability with which one party can convince another party that ρ is separable in a specific protocol [54].

It is easy to verify that for pure states Λ_{f}^2 coincides with Λ^2,

$$\Lambda_{\mathrm{f}}^2(|\psi\rangle) = \max_{\sigma\in\mathrm{SEP}} F^2(|\psi\rangle,\sigma) = \max_{\sigma\in\mathrm{SEP}} |\langle\psi|\sigma|\psi\rangle|$$
$$= \max_{|\varphi\rangle\in\mathrm{PRO}} |\langle\varphi|\psi\rangle|^2 = \Lambda^2(|\psi\rangle).$$

The corresponding extensions of the linear and logarithmic GM are

$$G^{\mathrm{f}}(\rho) := 1 - \Lambda_{\mathrm{f}}^2(\rho), \tag{23a}$$

$$G_{\mathrm{l}}^{\mathrm{f}}(\rho) := -\log \Lambda_{\mathrm{f}}^2(\rho). \tag{23b}$$

As seen for G^{m} and $G_{\mathrm{l}}^{\mathrm{m}}$, Lemma 2 applies to these measures. G^{f} is intimately related to the Groverian entanglement measure [which is equal to $\sqrt{G^{\rm f}(\rho)}$ for pure states ρ only] [7,20], thus giving it an operational interpretation by a quantum algorithm. $G^{\rm f}$ has been shown to be a weak entanglement measure [7] and has also been studied in Ref. [46]. On the other hand, little is known about $G^{\rm f}_{\rm l}$. It has been touched upon in the context of additivity in Ref. [9], but to our knowledge its properties have not been studied before.

Intriguingly, it was discovered that $G^{\rm f}(\rho)$ is equivalent to its convex roof [13], and since the convex roof is precisely $G^{\rm c}(\rho)$, the definitions (23a) and (18a) are equivalent:

Proposition 5. $G^f \equiv G^c$, i.e., $G^f(\rho) = G^c(\rho)$ holds for all states ρ.

We will jointly refer to these two definitions as $G^{\rm f/c}$ in the following, except in situations where the emphasis is on their formal definitions (i.e., fidelity-based versus convex roof-based). The relationship among $\sigma_{\rm f}$, the CSS of ρ in terms of (23a), and the optimal decomposition $\{P_i,|\Psi_i\rangle\}$ of ρ in terms of (18a) is also fully understood and outlined in Ref. [13]. Since $G^{\rm c}$ is known to be a strong entanglement measure with convexity, the same is true for $G^{\rm f}$. In particular, the convexity of (23a) implies that $\Lambda^2_{\rm f}(\rho)$ is concave. There are many states for which $G^{\rm c}(\rho)$ has been computed [6], and from Proposition 5 and Lemma 2 the values of $G^{\rm f}(\rho)$ and $G^{\rm f}_{\rm l}(\rho)$ directly follow.

With the known properties of $G^{\rm f}$, it follows from Lemma 2 that $G^{\rm f}_{\rm l}$ is a weak entanglement measure with convexity. From the convexity of $G^{\rm f}_{\rm l}$ it then follows that $G^{\rm f}_{\rm l}(\rho) \leqslant G^{\rm c}_{\rm l}(\rho)$ holds for all ρ as follows:

$$G^{\rm f}_{\rm l}(\rho) \leqslant \sum_i P_i G_{\rm l}(|\Psi_i\rangle) = \min_{\{p_i,|\psi_i\rangle\}} \sum_i p_i G_{\rm l}(|\psi_i\rangle) = G^{\rm c}_{\rm l}(\rho), \tag{24}$$

where $\{P_i,|\Psi_i\rangle\}$ is an optimal decomposition for $G^{\rm c}_{\rm l}(\rho)$. The question whether $G^{\rm f}_{\rm l}(\rho) \leqslant G^{\rm c}_{\rm l}(\rho)$ is strict for some states will be extensively studied in Sec. IV.

4. Extension by trace distance: G^t

It is tempting to introduce another mixed extension of GM, based on the trace distance defined in (3). From (8c) we obtain $D^2_{\rm T}(|\psi\rangle,|\phi\rangle) = 1 - F^2(|\psi\rangle,|\phi\rangle) = 1 - |\langle\psi|\phi\rangle|^2$, an expression with the form of (15a). We therefore define

$$G^{\rm t}(\rho) := \min_{|\varphi\rangle\in{\rm PRO}} D^2_{\rm T}(\rho,|\varphi\rangle) = \tfrac{1}{4}\min_{|\varphi\rangle\in{\rm PRO}} ({\rm Tr}\,|\rho - |\varphi\rangle\langle\varphi||)^2, \tag{25}$$

and call this measure the *trace extension of GM*. For pure input states $G^{\rm t}$ obviously coincides with (15a), so (25) is an extension of the linear GM. Note that the related definition,

$$\widetilde{E_{\rm T}}(\rho) = \min_{\sigma\in{\rm SEP}} D^2_{\rm T}(\rho,\sigma), \tag{26}$$

was already introduced (up to a square operation) as the trace entanglement in (10a) and shown to be a weak entanglement measure. From (25) and (26) it immediately follows that $\widetilde{E_{\rm T}}(\rho) \leqslant G^{\rm t}(\rho)$ for all ρ. To see whether $\widetilde{E_{\rm T}}$ also coincides with (15a) for the subset of pure states, we need to answer the question whether for pure input states the closest separable state $\sigma_{\rm t} \in \mathcal{S}(\mathcal{H})$ in terms of the trace distance can always be chosen to be pure, i.e., $\sigma_{\rm t} = |\phi\rangle\langle\phi|$. The cluster states provide a counterexample for this as follows.

Corollary 6. There exist pure states $|\psi\rangle$ for which $\widetilde{E_{\rm T}}(|\psi\rangle) < G^{\rm t}(|\psi\rangle)$ holds.

Proof. From Theorem 28 and the succeeding paragraph, together with (8c), it follows that $\widetilde{E_{\rm T}}(|{\rm C}_n\rangle) = D^2_{\rm T}(|{\rm C}_n\rangle,\delta) = (1-2^{-\frac{n}{2}})^2 < 1 - 2^{-\frac{n}{2}} = 1 - \Lambda^2(|{\rm C}_n\rangle) = G^{\rm t}(|{\rm C}_n\rangle)$ holds for all n qubit cluster states $|{\rm C}_n\rangle$ with even n. ■

As a consequence, $\widetilde{E_{\rm T}}$ is not an extension of the linear GM, because $D_{\rm T}(|\psi\rangle,\cdot)$ is in general not minimized by pure states. However, $\widetilde{E_{\rm T}}$ is an interesting quantity on its own, because it is a weak entanglement measure and a lower bound to $G^{\rm t}$. Furthermore, we will see in Sec. IV C that $\widetilde{E_{\rm T}}$ is also a lower bound to $G^{\rm f/c}$, which makes it a joint lower bound to all the GM definitions discussed in this paper.

The convexity of $E_{\rm T}$, and thus the convexity of $\widetilde{E_{\rm T}} = E^2_{\rm T}$, can be proved with the joint convexity of the trace distance. For any $\rho = \sum_i p_i\rho_i$ we have

$$\begin{aligned}\sum_i p_i E_{\rm T}(\rho_i) &= \sum_i p_i \min_{\sigma_i\in{\rm SEP}} D_{\rm T}(\rho_i,\sigma_i)\\ &= \min_{\{\sigma_i\}\in{\rm SEP}}\left[\sum_i p_i D_{\rm T}(\rho_i,\sigma_i)\right]\\ &\geqslant \min_{\{\sigma_i\}\in{\rm SEP}}\left[D_{\rm T}\big(\rho,\textstyle\sum_i p_i\sigma_i\big)\right] = \min_{\sigma\in{\rm SEP}} D_{\rm T}(\rho,\sigma)\\ &= E_{\rm T}(\rho),\end{aligned}$$

where the inequality follows from (4).

On the other hand, $G^{\rm t}$ is not an entanglement measure. For this, note that $G^{\rm t}(|\varphi\rangle) = 0$ holds for all $|\varphi\rangle \in {\rm PRO}$ and that $G^{\rm t}(\rho) > 0$ holds for all genuinely mixed $\rho \in {\rm SEP}$. From this it is not only clear that Axiom 1 is violated, but one can also immediately construct counterexamples for the concavity and the weak monotonicity (e.g., with the depolarizing channel [31]). Using the isotropic state as a counterexample, it is shown in Appendix B that $G^{\rm t}$ is not concave either.

IV. RELATIONSHIPS BETWEEN THE GM DEFINITIONS

In the previous section we introduced and discussed seven different definitions of GM for mixed states (and only two of them differ for pure states), of which two are equivalent. In the following we analyze the relationship between these different definitions.

For an arbitrary mixed state $\rho \in \mathcal{S}(\mathcal{H})$ the quantities $G^{\rm m}(\rho)$ and $G^{\rm m}_{\rm l}(\rho)$ correspond to the same closest product state $|\phi_{\rm m}\rangle \in {\rm PRO}$, and the quantities $G^{\rm f}(\rho)$ and $G^{\rm f}_{\rm l}(\rho)$ correspond to the same closest separable state $\sigma_{\rm f} \in {\rm SEP}$. In contrast to this, $G^{\rm c}(\rho)$ and $G^{\rm c}_{\rm l}(\rho)$ correspond to optimal decompositions $\{P_i,|\Psi_i\rangle\}$ that may differ for the two measures. The quantity $G^{\rm t}(\rho)$ corresponds to a closest product state $|\phi_{\rm t}\rangle \in {\rm PRO}$. In total, with the exception of the convex roof-based measures, the different values of GM for a given state are determined by two product states $|\phi_{\rm m}\rangle$, $|\phi_{\rm t}\rangle$ and one separable state $\sigma_{\rm f}$.

A. Comparison among $G^{\rm f/c}$, $G^{\rm c}_{\rm l}$, and $G^{\rm f}_{\rm l}$

From the previous discussion we already know that $G^{\rm f} \equiv G^{\rm c}$, $G^{\rm f/c} \cong G^{\rm f}_{\rm l}$, and that $G^{\rm c}_{\rm l}(\rho) \geqslant G^{\rm f}_{\rm l}(\rho) = -\log_2[1 -$

$G^{\text{f/c}}(\rho)$]. In this subsection, we study the connection between the fidelity-based and convex roof-based extensions in further detail by addressing some open problems. For example, it is neither obvious nor known whether G_1^{c} and G_1^{f} are equivalent ($G_1^{\text{c}} \equiv G_1^{\text{f}}$), and, if not, whether they have at least the same ordering ($G_1^{\text{c}} \cong G_1^{\text{f}}$). For this purpose, we will first derive necessary and sufficient conditions for $G_1^{\text{c}}(\rho) = G_1^{\text{f}}(\rho)$ and then investigate optimal decompositions for specific classes of states (e.g., maximally correlated states, isotropic states, two-qubit states).

Theorem 7. For any state ρ the following four conditions are equivalent:

(1) $G_l^c(\rho) = G_l^f(\rho)$ holds.

(2) $G_l^c(\rho) = -\log_2[1 - G^{f/c}(\rho)]$ holds.

(3) There exists a decomposition $\{P_i, |\Psi_i\rangle\}$, so

  (a) $\{P_i, |\Psi_i\rangle\}$ is optimal for $G^c(\rho)$ and $G_l^c(\rho)$, and

  (b) the $|\Psi_i\rangle$ are all equally entangled.

(4) For every optimal decomposition $\{P_i, |\Psi_i\rangle\}$ of $G_l^c(\rho)$ the following holds:

  (a) $\{P_i, |\Psi_i\rangle\}$ is also optimal for $G^c(\rho)$, and

  (b) the $|\Psi_i\rangle$ are all equally entangled.

Here, the meaning of 3(b) and 4(b) is that $\Lambda^2(|\Psi_i\rangle) = \Lambda^2(|\Psi_j\rangle)$ holds for all i, j.

Proof. Let $\{P_i, |\Psi_i\rangle\}$ be some optimal decomposition of $G_1^{\text{c}}(\rho)$. Using Lemma 2, Proposition 5, and (18), we have

$$
\begin{aligned}
G_1^{\text{f}}(\rho) &= -\log[1 - G^{\text{f/c}}(\rho)] = -\log\left[\max_{\{p_i, |\psi_i\rangle\}} \sum_i p_i \Lambda^2(|\psi_i\rangle)\right] \\
&\leqslant -\log\left[\sum_i P_i \Lambda^2(|\Psi_i\rangle)\right] \leqslant -\sum_i P_i \log \Lambda^2(|\Psi_i\rangle) \\
&= G_1^{\text{c}}(\rho),
\end{aligned}
\tag{27}
$$

where the second inequality follows from the concavity of the logarithm.

1. $\Leftrightarrow$ 2.: This equivalency follows immediately from $G_1^{\text{f}}(\rho) = -\log_2[1 - G^{\text{f/c}}(\rho)]$.

1. $\Leftrightarrow$ 4.: Apparently, $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho)$ holds if and only if both inequalities in (27) become equalities. Regarding the first inequality in (27), this inequality becomes an equality if and only if $\{P_i, |\Psi_i\rangle\}$ is also an optimal decomposition for $G^{\text{c}}(\rho)$. Regarding the second inequality in (27), the strict concavity of the logarithm implies that this inequality becomes an equality if and only if $\Lambda^2(|\Psi_i\rangle) = \Lambda^2(|\Psi_j\rangle)$ holds for all i, j. Together, this yields that condition 1 holds if and only if condition 4 holds.

4. $\Rightarrow$ 3.: Obvious, since $G_1^{\text{c}}(\rho)$ has at least one optimal decomposition.

3. $\Rightarrow$ 1.: Using the decomposition $\{P_i, |\Psi_i\rangle\}$ postulated by condition 3, the two inequalities in (27) turn into equalities. Therefore, $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho)$. ■

Note that for states ρ that fall under Theorem 7any optimal decomposition for $G_1^{\text{c}}(\rho)$ is also optimal for $G^{\text{c}}(\rho)$, but the converse is not necessarily true. In other words, for states ρ that satisfy $G_1^{\text{c}}(\rho) = G_1^{\text{f}}(\rho)$, the set of optimal decompositions for $G_1^{\text{c}}(\rho)$ is a nonempty subset of the set of optimal decompositions for $G^{\text{c}}(\rho)$. It is an open question whether states ρ with $G_1^{\text{c}}(\rho) = G_1^{\text{f}}(\rho)$ exist, for which the set of optimal decompositions of $G^{\text{c}}(\rho)$ is strictly larger than that of $G_1^{\text{c}}(\rho)$. Without the condition $G_1^{\text{c}}(\rho) = G_1^{\text{f}}(\rho)$, $G^{\text{c}}(\rho)$ and $G_1^{\text{c}}(\rho)$ may not even have any common optimal decomposition, as shown later in Corollary 15.

The following corollary helps to understand the relationship between optimal decompositions of $G^{\text{c}}(\rho)$ and $G_1^{\text{c}}(\rho)$.

Corollary 8. Let $\{P_i, |\Psi_i\rangle\}$ be some decomposition of ρ. The following two conditions are equivalent:

(1) $\{P_i, |\Psi_i\rangle\}$ is an optimal decomposition for $G^c(\rho)$, and the $|\Psi_i\rangle$ are all equally entangled.

(2) $\{P_i, |\Psi_i\rangle\}$ is an optimal decomposition for $G_l^c(\rho)$, and $G_l^c(\rho) = G_l^f(\rho)$ holds.

Proof.

2 $\Rightarrow$ 1: This easily follows from items 1 and 4 of Theorem 7.

1 $\Rightarrow$ 2: Let $\{P_i, |\Psi_i\rangle\}$ be an optimal decomposition of $G^{\text{c}}(\rho)$, where the $|\Psi_i\rangle$ are all equally entangled. In analogy to (27), we have

$$
\begin{aligned}
G_1^{\text{f}}(\rho) &= -\log[1 - G^{\text{f/c}}(\rho)] = -\log\left\{\sum_i P_i \Lambda^2(|\Psi_i\rangle)\right\} \\
&= \sum_i P_i[-\log \Lambda^2(|\Psi_i\rangle)] \geqslant \min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i G_1(|\psi_i\rangle) \\
&= G_1^{\text{c}}(\rho),
\end{aligned}
\tag{28}
$$

where the third equality follows from the fact that the $|\Psi_i\rangle$ are all equally entangled. As shown in (24), $G_1^{\text{f}}(\rho) \leqslant G_1^{\text{c}}(\rho)$ holds for all ρ, so the inequality in (28) must be an equality. Therefore, $\{P_i, |\Psi_i\rangle\}$ is also an optimal decomposition for $G_1^{\text{c}}(\rho)$, and $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho)$ holds. ■

Following the derivation of general results, we next investigate classes of states for whom $G_1^{\text{f}}(\rho)$ and $G_1^{\text{c}}(\rho)$ coincide. We will see that for two-qubit systems $G_1^{\text{f}} \equiv G_1^{\text{c}}$ holds and that for general bipartite systems $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho)$ holds for the subset of isotropic states.

Proposition 9. $G_l^f \equiv G_l^c$ holds for two qubits.

Proof. According to Proposition 4 of Ref. [13], if $f(x)$ is a non-negative convex function for $x \geqslant 0$ and obeys $f(0) = 0$, then for two-qubit systems $f(G^{\text{c}}(\rho))$ is equal to its convex roof. The function $f(x) := -\log_2(1 - x)$ satisfies the requirements, so $f(G^{\text{c}}(\rho)) = f(G^{\text{f}}(\rho)) = G_1^{\text{f}}(\rho)$ is equal to its convex roof, which is precisely $\min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i G_1(|\psi_i\rangle) = G_1^{\text{c}}(\rho)$. Therefore, $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho)$ holds for all ρ. ■

Proposition 10. $G_l^f(\rho) = G_l^c(\rho)$ holds for isotropic states in two qudits.

Proof. The isotropic states are $\rho_{\text{iso}} = p\mathbb{1}/d^2 + (1 - p)|\Psi\rangle\langle\Psi|$, with $p \in [0,1]$ and $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |ii\rangle$. The parametrization employed in Ref. [6] translates to ours as $p = \frac{d^2}{d^2-1}(1 - F)$, where $F \in [0,1]$. For $F \in [0, \frac{1}{d}]$, ρ is separable [6], which implies $G_1^{\text{f}}(\rho) = G_1^{\text{c}}(\rho) = 0$. Next, consider the entangled region $F \in (\frac{1}{d}, 1]$. From Ref. [6] [Eq. (36) and (54)] it follows that $E_{\sin^2} = G^{\text{f/c}}$ and $G^{\text{f/c}}(\rho_{\text{iso}}(F)) = 1 - \frac{1}{d}[\sqrt{F} + \sqrt{(d-1)(1-F)}]^2$. From Lemma 2 it follows that $G_1^{\text{f}}(\rho_{\text{iso}}(F)) = -\log \frac{1}{d}[\sqrt{F} + \sqrt{(d-1)(1-F)}]^2$. To compute $G_1^{\text{c}}(\rho_{\text{iso}}(F))$, we follow the idea of Ref. [6] to obtain the first and second equalities of the following equation:

$$G_1^c(\rho_{iso}(F)) = \mathcal{C}_{conv}\left\{-\max_{\{\mu_i\}}\left[\log \mu_i \middle| F = \frac{1}{d}\left(\sum_{i=1}^{d}\sqrt{\mu_i}\right)^2\right]\right\}$$
$$= \mathcal{C}_{conv}\left\{-\log \frac{1}{d}[\sqrt{F} + \sqrt{(d-1)(1-F)}]^2\right\}$$
$$= \mathcal{C}_{conv}\left[G_1^f(\rho_{iso}(F))\right] = G_1^f(\rho_{iso}(F)). \quad (29)$$

Here $\mathcal{C}_{conv}$ is the convex hull. The last equality in (29) follows from the fact that $G_1^f(\rho_{iso}(F))$ is convex. This completes the proof. ∎

Next, we investigate maximally correlated states, because necessary and sufficient conditions for $G_1^c(\rho) = G_1^f(\rho)$ can be derived for these states. The maximally correlated states have been extensively studied in terms of the entanglement of formation and distillable entanglement [42,55]. In this paper we focus on a special type of maximally correlated states defined as follows:

Definition 11. Given a bipartite d-level system $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$, let $\{n_i\}_{i=0}^r$ with $r < d$ be a set of integers with $0 = n_0 < \cdots < n_r = d$, and let $|\Theta_i\rangle = \frac{1}{\sqrt{n_i - n_{i-1}}}\sum_{k=n_{i-1}+1}^{n_i}|kk\rangle$ be the $(n_i - n_{i-1})$-level MES. Then $\rho = \sum_{i=1}^r q_i|\Theta_i\rangle\langle\Theta_i|$, with $\sum_{i=1}^r q_i = 1$ and $q_i \in (0,1)$, is a $d\times d$ maximally correlated state.

From now on, we refer to maximally correlated states as the states defined above. The integer r can be readily identified as the rank of the maximally correlated state, i.e., rank $\rho = r$.

Lemma 12. Let $\rho = \sum_{i=1}^r q_i|\Theta_i\rangle\langle\Theta_i|$ be a $d\times d$ maximally correlated state. Then

(1) The unique optimal decomposition for $G^c(\rho)$, up to overall phases, is $\{q_i,|\Theta_i\rangle\}$.

(2) $G^c(\rho) = G^f(\rho) = 1 - \sum_{i=1}^r \frac{q_i}{n_i - n_{i-1}}$.

(3) $G_l^c(\rho) \geqslant G_l^f(\rho) = -\log(\sum_{i=1}^r \frac{q_i}{n_i - n_{i-1}})$.

(4) $G_l^c(\rho) = G_l^f(\rho)$ if and only if the $|\Theta_i\rangle$ are all equally entangled.

(5) $G_l^c(\rho) = G_l^f(\rho)$ if and only if $n_i - n_{i-1}$ is the same for all $i = 1,\ldots,r$.

(6) If $G_l^c(\rho) = G_l^f(\rho)$, then $G_l^c(\rho)$ has the same optimal decompositions as $G^c(\rho)$.

Proof. First we prove item 1. Let $\{P_i,|\Psi_i\rangle\}$ be some optimal decomposition of $G^c(\rho)$. Then

$$1 - G^c(\rho) = \sum_i P_i\Lambda^2(|\Psi_i\rangle) \geqslant \sum_{j=1}^r q_j\Lambda^2(|\Theta_j\rangle)$$
$$= \sum_{j=1}^r \frac{q_j}{n_j - n_{j-1}}. \quad (30)$$

According to (19), the relationship between the decompositions $\{P_i,|\Psi_i\rangle\}$ and $\{q_j,|\Theta_j\rangle\}$ is $\sqrt{P_i}|\Psi_i\rangle = \sum_{j=1}^r u_{ij}\sqrt{q_j}|\Theta_j\rangle$, where u_{ij} is some unitary matrix. Because of the form of the $|\Theta_j\rangle$, this immediately yields the Schmidt decomposition of $|\Psi_i\rangle$. The GM of pure bipartite states is determined by their largest Schmidt coefficient, so we have

$$P_i\Lambda^2(|\Psi_i\rangle) = \max_{j\in\{1,\ldots,r\}}\left\{\left|\frac{u_{ij}\sqrt{q_j}}{\sqrt{n_j - n_{j-1}}}\right|^2\right\} \leqslant \sum_{j=1}^r \frac{u_{ij}^* u_{ij}\, q_j}{n_j - n_{j-1}}. \quad (31)$$

By summing over all i, we obtain $\sum_i P_i\Lambda^2(|\Psi_i\rangle) \leqslant \sum_{j=1}^r \frac{q_j}{n_j - n_{j-1}}$. Comparing this inequality to (30), we see that the inequalities must become equalities, and therefore $\{q_j,|\Theta_j\rangle\}$ is also optimal for $G^c(\rho)$. Since the inequality in (31) becomes an equality, all but one $u_{i1},\ldots,u_{ir}$ are zero. Hence, the state $|\Psi_i\rangle$ is identical to one of the states $|\Theta_j\rangle$, $j = 1,\ldots,r$, up to an overall phase. Therefore, $\{q_i,|\Theta_i\rangle\}$ is the unique optimal decomposition for $G^c(\rho)$, up to overall phases.

With item 1 proved, the other items easily follow:

Item 2 and 3: These follow directly from item 1 and $\Lambda^2(|\Theta_i\rangle) = \frac{1}{n_i - n_{i-1}}$, together with Proposition 5 and (24), respectively.

Item 4: If $G_1^c(\rho) = G_1^f(\rho)$, then it follows from item 1 and Theorem 7 that the $|\Theta_i\rangle$ are all equally entangled. Conversely, if the $|\Theta_i\rangle$ are all equally entangled, then it follows from item 1 and Corollary 8 that $G_1^c(\rho) = G_1^f(\rho)$.

Item 5: Because of $\Lambda^2(|\Theta_i\rangle) = \frac{1}{n_i - n_{i-1}}$, items 4 and 5 are equivalent.

Item 6: If $G_1^c(\rho) = G_1^f(\rho)$, then it follows from item 1 and Theorem 7 that every optimal decomposition of $G_1^c(\rho)$ must be of the form $\{q_i,|\Theta_i\rangle\}$, up to overall phases. Since overall phases do not change the value of (18a) or (18b), $G^c(\rho)$ and $G_1^c(\rho)$ have the same optimal decompositions. ∎

One may wonder whether the optimal decomposition for $G^c(\rho)$ in Lemma 12 is also optimal for $G_1^c(\rho)$ when $G_1^c(\rho) \neq G_1^f(\rho)$. In the following we show that this is the case for all maximally correlated qutrit states. Only rank-2 states need to be considered, because for $d = 3$ this is the only nontrivial case.

Proposition 13. Let $q \in (0,1)$ and $|\psi\rangle = \frac{1}{\sqrt{2}}(|11\rangle + |22\rangle)$. The maximally correlated two-qutrit state $\rho = q|00\rangle\langle 00| + (1-q)|\psi\rangle\langle\psi|$ has $G_l^c(\rho) = 1 - q$, with $\{q,|00\rangle; 1-q,|\psi\rangle\}$ being an optimal decomposition.

Proof. Let $\{P_i,|\Psi_i\rangle\}$ be some optimal decomposition of ρ for $G_1^c(\rho)$. According to (19), there exists a unitary u_{ij}, so $\sqrt{P_i}|\Psi_i\rangle = u_{i1}\sqrt{q}|00\rangle + u_{i2}\sqrt{1-q}|\psi\rangle$ for each i. Setting $x_i := |u_{i1}\sqrt{q}|^2$ and $y_i := |u_{i2}\sqrt{1-q}|^2$, we have $P_i = x_i + y_i$ and $\Lambda^2(|\Psi_i\rangle) = \max\{\frac{x_i}{P_i},\frac{y_i}{2P_i}\}$. Therefore,

$$G_1^c(\rho) = -\sum_{i=1}^r P_i \log\left[\max\left\{\frac{x_i}{P_i},\frac{y_i}{2P_i}\right\}\right]. \quad (32)$$

Without loss of generality we assume $2x_i \geqslant y_i$ for $i \in [1,k]$ and $2x_i \leqslant y_i$ for $i \in [k+1,r]$. We define $Y := \sum_{i=1}^k y_i \geqslant 0$ and $X := \sum_{i=1}^k x_i > 0$. Then

$$G_1^c(\rho) = \sum_{i=1}^k P_i \log\left(\frac{P_i}{x_i}\right) + \sum_{j=k+1}^r P_j \log\left(\frac{2P_j}{y_j}\right)$$
$$= \sum_{i=1}^k (x_i + y_i)\log\left(1 + \frac{y_i}{x_i}\right)$$
$$+ \sum_{j=k+1}^r (x_j + y_j)\log\left(2 + \frac{2x_j}{y_j}\right)$$

$$\geqslant Y\left[\left(1+\frac{X}{Y}\right)\log\left(1+\frac{Y}{X}\right)\right]+\sum_{j=k+1}^{r} y_j$$

$$\geqslant \frac{Y}{\ln 2}+\sum_{j=k+1}^{r} y_j \geqslant \sum_{i=1}^{r} y_i = 1-q.$$

The first inequality follows by applying Lemma 30 from Appendix C to the first sum and using $x_j \geqslant 0$, $y_j > 0$ in the second sum. The application of Lemma 30 is possible despite the restriction $2x_i \geqslant y_i$, because the minimum in (C1) cannot be smaller with additional restrictions than without. The second inequality follows from $\inf_{x>0}(1+\frac{1}{x})\log_2(1+x) = 1/\ln 2$.

On the other hand, the decomposition $\{q,|00\rangle; 1-q,|\psi\rangle\}$ yields $G_1^c(\rho) \leqslant 1-q$. This completes the proof. ■

One may conjecture that this proposition can be generalized to higher dimensions, with $|\psi\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|ii\rangle$ being a pure MES of any dimension, i.e., that $\{q,|00\rangle; 1-q,|\psi\rangle\}$ is an optimal decomposition of $\rho = q|00\rangle\langle 00| + (1-q)|\psi\rangle\langle\psi|$ for $G_l^c(\rho)$, yielding $G_l^c(\rho) = (1-q)\log n$. From Proposition 14 it will follow that this is not the case for $n > 2$. We will see that—compared to the qutrit case—the optimal decomposition for $G_l^c(\rho)$ of higher-dimensional maximally correlated states is more complex, even in the comparatively easy rank-2 case. In the following, e denotes the base of the natural logarithm.

Proposition 14. Let $m,n \in \mathbb{N}$ with $\frac{m}{n} \leqslant 1$ and $q \in (0,1)$ be constants that define the rank-2 maximally correlated qudit state $\rho = q|\psi_m\rangle\langle\psi_m| + (1-q)|\psi_n\rangle\langle\psi_n|$, with $|\psi_m\rangle = \frac{1}{\sqrt{m}}\sum_{i=1}^{m}|ii\rangle$ and $|\psi_n\rangle = \frac{1}{\sqrt{n}}\sum_{j=m+1}^{m+n}|jj\rangle$. Depending on the constants m, n, and q, an optimal decomposition and the value of $G_l^c(\rho)$ are

$\frac{m}{n} \geqslant \frac{1}{e}$: $\{q,|\psi_m\rangle; 1-q,|\psi_n\rangle\}$, yielding $G_l^c(\rho) = q\log m + (1-q)\log n$.

$\frac{m}{n} < \frac{1}{e}$ and $q \geqslant \frac{em}{n}$: $\{\frac{1}{2},\sqrt{q}|\psi_m\rangle \pm \sqrt{1-q}|\psi_n\rangle\}$, yielding $G_l^c(\rho) = \log(\frac{m}{q})$.

$\frac{m}{n} < \frac{1}{e}$ and $q < \frac{em}{n}$: $\{1-\frac{nq}{em},|\psi_n\rangle; \frac{nq}{2em},\sqrt{\frac{em}{n}}|\psi_m\rangle \pm \sqrt{1-\frac{em}{n}}|\psi_n\rangle\}$, yielding $G_l^c(\rho) = \log n - q\frac{n\log e}{me}$.

Proof. The above decompositions provide trivial upper bounds [note that in the second case $\Lambda^2(\sqrt{q}|\psi_m\rangle \pm \sqrt{1-q}|\psi_n\rangle) = \frac{q}{m}$ follows from $\frac{q}{m} > \frac{q(1-q)}{m} \geqslant \frac{e(1-q)}{n} > \frac{1-q}{n}$, and in the third case $\Lambda^2(\sqrt{\frac{em}{n}}|\psi_m\rangle \pm \sqrt{1-\frac{em}{n}}|\psi_n\rangle) = \frac{e}{n}$ follows from $\frac{e}{n} > \frac{1}{n} \geqslant \frac{1}{n}(1-\frac{em}{n})$]. Below we show that these are also lower bounds.

Let $\{P_i,|\Psi_i\rangle\}$ be some optimal decomposition for $G_1^c(\rho)$. According to (19), there exists a unitary u_{ij}, so $\sqrt{P_i}|\Psi_i\rangle = u_{i1}\sqrt{q}|\psi_m\rangle + u_{i2}\sqrt{1-q}|\psi_n\rangle$ for each i. Setting $x_i := |u_{i1}\sqrt{q}|^2$, and $y_i := |u_{i2}\sqrt{1-q}|^2$, we have $P_i = x_i + y_i$ and $\Lambda^2(|\Psi_i\rangle) = \max\{\frac{x_i}{mP_i},\frac{y_i}{nP_i}\}$. Therefore,

$$G_1^c(\rho) = -\sum_{i=1}^{r} P_i \log\left[\max\left\{\frac{x_i}{mP_i},\frac{y_i}{nP_i}\right\}\right].$$

First, we rule out $nx_i = my_i\ \forall i$ by showing that $G_1^c(\rho) = \sum_i P_i\log(mP_i/x_i) = \sum_i P_i\log(m+n) = \log(m+n)$ surpasses the upper bounds outlined above: For all $\frac{m}{n} \leqslant 1$ we have $\log(m+n) > \log n \geqslant q\log m + (1-q)\log n$, as well as $\log(m+n) > \log n \geqslant \log n - q\frac{n\log e}{me}$. Furthermore, for $\frac{m}{n} < \frac{1}{e}$ and $q \geqslant \frac{em}{n}$ we have $\log(m+n) > \log n > \log(\frac{mn}{me}) \geqslant \log(\frac{m}{q})$. In the following we therefore assume that $nx_i \neq my_i$ holds for at least one $i \in [1,r]$.

Without loss of generality we assume $nx_i \geqslant my_i$ for $i \in [1,k]$ and $nx_i \leqslant my_i$ for $i \in [k+1,r]$. We define $Y := \sum_{i=1}^{k} y_i \geqslant 0$, $X := \sum_{i=1}^{k} x_i > 0$, $\widetilde{Y} := \sum_{i=k+1}^{r} y_i > 0$, and $\widetilde{X} := \sum_{i=k+1}^{r} x_i \geqslant 0$, as well as $h := \frac{Y}{X} \geqslant 0$ and $s := \frac{\widetilde{X}}{\widetilde{Y}} \geqslant 0$. Note that $0 \leqslant hs < 1$, because of $hs = \frac{Y}{X}\frac{\widetilde{X}}{\widetilde{Y}} < \frac{n}{m}\frac{m}{n} = 1$ (the inequality is strict, because $nx_i > my_i$ or $nx_i < my_i$ holds for at least one i). Using $Y + \widetilde{Y} = 1-q$ and $X + \widetilde{X} = q$, it is easy to verify that $X = \frac{q-s(1-q)}{1-hs}$ and $\widetilde{Y} = \frac{(1-q)-hq}{1-hs}$. From $X > 0$ and $\widetilde{Y} > 0$, it then follows that $s \in [0,\frac{q}{1-q}]$ and $h \in [0,\frac{1-q}{q}]$, respectively.

$$\begin{aligned} G_1^c(\rho) &= \sum_{i=1}^{k} P_i\log\left(\frac{mP_i}{x_i}\right) + \sum_{i=k+1}^{r} P_i\log\left(\frac{nP_i}{y_i}\right) \\ &= \sum_{i=1}^{k}(x_i+y_i)\log\left[m\left(1+\frac{y_i}{x_i}\right)\right] \\ &\quad + \sum_{i=k+1}^{r}(x_i+y_i)\log\left[n\left(1+\frac{x_i}{y_i}\right)\right] \\ &\geqslant X(1+h)\log[m(1+h)] + \widetilde{Y}(1+s)\log[n(1+s)], \end{aligned} \tag{33}$$

where the inequality follows from applying Lemma 30 to each of the two sums. In Lemma 31 of Appendix C we show that the last line of (33) is an upper bound to the values outlined in the proposition. Hence, the upper and lower bounds coincide. This completes the proof. ■

Note that for qutrits ($m+n=3$) Proposition 14 simplifies to Proposition 13, because $\frac{1}{e} < \frac{1}{2} \leqslant \frac{m}{n}$. From the symmetry of $\rho = q|\psi_m\rangle\langle\psi_m| + (1-q)|\psi_n\rangle\langle\psi_n|$, it can be seen that Proposition 14 can be extended to the case $\frac{m}{n} > 1$ simply by swapping q and $1-q$. Importantly, Proposition 14 yields necessary and sufficient conditions for rank-2 maximally correlated qudit states to have common optimal decompositions for $G^c(\rho)$ and $G_1^c(\rho)$.

Corollary 15. Rank-2 maximally correlated states $\rho = q|\psi_m\rangle\langle\psi_m| + (1-q)|\psi_n\rangle\langle\psi_n|$ have common optimal decompositions for $G^c(\rho)$ and $G_l^c(\rho)$ if and only if $\frac{1}{e} \leqslant \frac{m}{n} \leqslant e$.

Proof. According to Lemma 12, the unique optimal decomposition for $G^c(\rho)$ is $\{q,|\psi_m\rangle; 1-q,|\psi_n\rangle\}$, up to overall phases. For symmetry reasons it suffices to consider $\frac{m}{n} \leqslant 1$. For $\frac{m}{n} \geqslant \frac{1}{e}$ the statement immediately follows from Proposition 14. For $\frac{m}{n} < \frac{1}{e}$ it is seen in the proof of Lemma 31 that for all $q \in (0,1)$ the minimum of $f(h,s)$ is strictly smaller than $f(0,0) = q\log m + (1-q)\log n$. Therefore, $\{q,|\psi_m\rangle; 1-q,|\psi_n\rangle\}$ cannot be an optimal decomposition of $G_1^c(\rho)$ for $\frac{m}{n} < \frac{1}{e}$. ■

Let us sum up the preceding findings. Theorem 7 gives necessary and sufficient conditions for $G_1^c(\rho) = G_1^f(\rho)$. Apart from the trivial classes of pure states and separable states, this includes all two-qubit states (Proposition 9) and all isotropic states (Proposition 10). Further examples are the four-qubit

Smolin state and multiqubit Dür states, for whom $G_1^{\mathrm{c}}(\rho) = -\log_2[1 - G^{\mathrm{c}}(\rho)]$ can be easily verified from Ref. [15].

According to Lemma 12, $G_1^{\mathrm{c}}(\rho) > G_1^{\mathrm{f}}(\rho)$ holds for maximally correlated states if and only if there are two states $|\Theta_i\rangle$ and $|\Theta_j\rangle$ in the optimal decomposition of $G^{\mathrm{c}}(\rho)$ that are not equally entangled. This is the case for the two-qutrit states of Proposition 13. Despite this, Proposition 13 shows that $G^{\mathrm{c}}(\rho)$ and $G_1^{\mathrm{c}}(\rho)$ still have common optimal decompositions, i.e., item 3(a) of Theorem 7can be true, while item 1 is false. In this case, the first inequality of (27) is an equality, while the second inequality is strict.

Optimal decompositions of two-qutrit isotropic states were found in Ref. [56] for the entanglement of formation, a convex roof-based entanglement measure. In all of these optimal decompositions *some* of the pure states are not equally entangled (although it is unknown whether [56] exhausts all optimal decompositions for two-qutrit isotropic states). Using Lemma 12, one can easily construct a state ρ in whose optimal decomposition $\{P_i,|\Psi_i\rangle\}$ for $G^{\mathrm{c}}(\rho)$ *any* two states $|\Psi_i\rangle$ and $|\Psi_j\rangle$ are not equally entangled. In some sense, this is a stronger result than the one of Ref. [56].

Next we address the question whether G_1^{c} has the same ordering as $G^{\mathrm{f/c}}$ or G_1^{f}. Because of $G^{\mathrm{f/c}} \cong G_1^{\mathrm{f}}$, the statement $G_1^{\mathrm{c}} \cong G^{\mathrm{f/c}}$ is equivalent to $G_1^{\mathrm{c}} \cong G_1^{\mathrm{f}}$. The two-qubit case is trivial, because of $G_1^{\mathrm{c}} \equiv G_1^{\mathrm{f}}$ (cf. Proposition 9), so we need to consider higher-dimensional systems.

Corollary 16. In general, $G^{f/c}$ and G_l^c do not have the same ordering ($G^{f/c} \ncong G_l^c$). Equivalently, G_l^f and G_l^c do not have the same ordering ($G_l^f \ncong G_l^c$).

Proof. A simple counterexample are the two maximally correlated six-level states $\rho = \frac{1}{2}|\Psi_{123}\rangle\langle\Psi_{123}| + \frac{1}{2}|\Psi_{456}\rangle\langle\Psi_{456}|$ and $\sigma = \frac{1}{3}|\Psi_{12}\rangle\langle\Psi_{12}| + \frac{2}{3}|\Psi_{3456}\rangle\langle\Psi_{3456}|$, where $|\Psi_{a\ldots z}\rangle := \frac{1}{\sqrt{z-a+1}}(|aa\rangle + \cdots + |zz\rangle)$. From items 2 and 4 of Lemma 12 it follows that $G^{\mathrm{f/c}}(\rho) = G^{\mathrm{f/c}}(\sigma) = \frac{2}{3}$ but $G_1^{\mathrm{c}}(\rho) < G_1^{\mathrm{c}}(\sigma)$. ■

We remark that Corollary 16 can be easily verified for a much wider range of systems, e.g., all bipartite d-level systems with $d \geqslant 4$, by considering any rank-2 maximally correlated state ρ belonging to the second class outlined in Proposition 14, together with a suitably chosen isotropic state σ, yielding $G^{\mathrm{f/c}}(\rho) < G^{\mathrm{f/c}}(\sigma)$ and $G_1^{\mathrm{c}}(\rho) > G_1^{\mathrm{c}}(\sigma)$.

B. Comparison between G^{m} and G_1^{m} and between G^{c} and G_1^{c}

In contrast to the convex roof-based extensions, G^{m} and G_1^{m} are demonstrably not convex, and they attain their maximum for the maximally mixed state. It is therefore intuitive to expect that $G^{\mathrm{m}}(\rho) \geqslant G^{\mathrm{c}}(\rho)$ and $G_1^{\mathrm{m}}(\rho) \geqslant G_1^{\mathrm{c}}(\rho)$ hold for all ρ. To prove these statements, we need the following lemma.

Lemma 17. Let $\rho,|\varphi\rangle$ be two arbitrary states, and $\langle\varphi|\rho|\varphi\rangle = g$. Then there exists a decomposition $\rho = \sum_{i=1}^r p_i|\psi_i\rangle\langle\psi_i|$, such that $r = \operatorname{rank}\rho$, and $|\langle\varphi|\psi_i\rangle|^2 = g$ for all i.

Proof. We use induction on the rank of ρ. The claim is trivial for rank $\rho = 1$. Suppose it is true for rank $\rho = r$. Consider a general state ρ with rank $\rho = r + 1$ and spectral decomposition $\rho = \sum_{i=1}^{r+1} p_i|\psi_i\rangle\langle\psi_i|$, with $\langle\psi_i|\psi_j\rangle = 0$ for $i \neq j$. Since $\langle\varphi|\rho|\varphi\rangle = g$, we can assume $|\langle\varphi|\psi_1\rangle|^2 \leqslant g \leqslant |\langle\varphi|\psi_2\rangle|^2$ without loss of generality. Denote $\langle\varphi|\psi_j\rangle = s_j e^{i\theta_j}$, $s_j \geqslant 0$ for $j = 1,2$. Using (19), we rewrite the sum of the first two terms as

$$p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2| = q_1|\phi_1\rangle\langle\phi_1| + q_2|\phi_2\rangle\langle\phi_2|,$$

with $\sqrt{q_i}|\phi_i\rangle = u_{i1}\sqrt{p_1}|\psi_1\rangle + u_{i2}\sqrt{p_2}|\psi_2\rangle$ for $i = 1,2$, and where we define the unitary matrix as

$$U(\vartheta) = [u_{ij}] = \begin{pmatrix} \mathrm{c}_\vartheta & \mathrm{s}_\vartheta e^{i(\theta_1-\theta_2)} \\ -\mathrm{s}_\vartheta e^{-i(\theta_1-\theta_2)} & \mathrm{c}_\vartheta \end{pmatrix}, \quad \vartheta \in (0,\pi),$$

with $\mathrm{c}_\vartheta := \cos\frac{\vartheta}{2}$ and $\mathrm{s}_\vartheta := \sin\frac{\vartheta}{2}$. From $\langle\psi_1|\psi_2\rangle = 0$ and $\langle\phi_1|\phi_1\rangle = 1$, we obtain $q_1 = \mathrm{c}_\vartheta^2 p_1 + \mathrm{s}_\vartheta^2 p_2$, hence

$$|\phi_1\rangle = \frac{\mathrm{c}_\vartheta\sqrt{p_1}|\psi_1\rangle + \mathrm{s}_\vartheta e^{i(\theta_1-\theta_2)}\sqrt{p_2}|\psi_2\rangle}{(\mathrm{c}_\vartheta^2 p_1 + \mathrm{s}_\vartheta^2 p_2)^{\frac{1}{2}}} \quad \text{and}$$

$$|\langle\varphi|\phi_1\rangle| = \frac{\mathrm{c}_\vartheta\sqrt{p_1}s_1 + \mathrm{s}_\vartheta\sqrt{p_2}s_2}{(\mathrm{c}_\vartheta^2 p_1 + \mathrm{s}_\vartheta^2 p_2)^{\frac{1}{2}}}.$$

We see that $\lim_{\vartheta\to 0}|\langle\varphi|\phi_1\rangle| = s_1$ and $\lim_{\vartheta\to\pi}|\langle\varphi|\phi_1\rangle| = s_2$. Since $s_1^2 \leqslant g \leqslant s_2^2$ and $|\langle\varphi|\phi_1\rangle|$ is continuous in ϑ, there must be some ϑ such that $|\langle\varphi|\phi_1\rangle|^2 = g$. Denoting the corresponding $\{q_i,|\phi_i\rangle\}$ as $\{Q_i,|\Phi_i\rangle\}$, we see that the state

$$\rho_1 := \frac{\rho - Q_1|\Phi_1\rangle\langle\Phi_1|}{1 - Q_1} = \frac{Q_2|\Phi_2\rangle\langle\Phi_2| + \sum_{i=3}^{r+1} p_i|\psi_i\rangle\langle\psi_i|}{1 - Q_1}$$

is a state of rank r that satisfies $\langle\varphi|\rho_1|\varphi\rangle = g$. Using the induction assumption on ρ_1, there is a decomposition $\rho_1 = \sum_{i=1}^r p_i'|\psi_i'\rangle\langle\psi_i'|$, such that $|\langle\varphi|\psi_i'\rangle|^2 = g$ for all i. Now the claim follows for $\rho = (1 - Q_1)\rho_1 + Q_1|\Phi_1\rangle\langle\Phi_1| = (1 - Q_1)\sum_{i=1}^r p_i'|\psi_i'\rangle\langle\psi_i'| + Q_1|\Phi_1\rangle\langle\Phi_1|$. This completes the proof. ■

Theorem 18. $G_l^m(\rho) \geqslant G_l^c(\rho)$ holds for all states ρ.

Proof. Let $|\varphi\rangle \in \mathrm{PRO}$ be a closest product state of ρ in accordance with (20), and let rank $\rho = r$. By virtue of Lemma 17, there exists a decomposition $\rho = \sum_{i=1}^r P_i|\Psi_i\rangle\langle\Psi_i|$ such that $|\langle\varphi|\Psi_i\rangle|^2 = \langle\varphi|\rho|\varphi\rangle$ for all i. Then we have

$$\begin{aligned} G_1^{\mathrm{m}}(\rho) &= -\log\langle\varphi|\rho|\varphi\rangle = -\log\sum_{i=1}^r P_i|\langle\varphi|\Psi_i\rangle|^2 \\ &= -\sum_{i=1}^r P_i \log|\langle\varphi|\Psi_i\rangle|^2 \geqslant -\sum_{i=1}^r P_i \log\Lambda^2(|\Psi_i\rangle) \\ &\geqslant \min_{\{p_i,|\psi_i\rangle\}} -\sum_i p_i \log\Lambda^2(|\psi_i\rangle) = G_1^{\mathrm{c}}(\rho), \end{aligned} \tag{34}$$

where the third equality follows from the fact that the $|\langle\varphi|\Psi_i\rangle|$ have the same value for all i. ■

Corollary 19. $G^m(\rho) \geqslant G^{f/c}(\rho)$ holds for all states ρ.

Proof. Using Lemma 2, Theorem 18 and Lemma 3, we obtain $G^{\mathrm{m}}(\rho) = 1 - 2^{-G_1^{\mathrm{m}}(\rho)} \geqslant 1 - 2^{-G_1^{\mathrm{c}}(\rho)} \geqslant G^{\mathrm{c}}(\rho)$. ■

The following Theorem 20 establishes necessary and sufficient conditions for $G_1^{\mathrm{m}}(\rho) = G_1^{\mathrm{c}}(\rho)$ in form of a straightforward relationship between the CPS for $\Lambda_{\mathrm{m}}^2(\rho)$ and the optimal decomposition of $G_1^{\mathrm{c}}(\rho)$. Hence, this theorem bears resemblance to Theorem 7, as well as to Proposition 5 of Ref. [13].

Theorem 20. For any state ρ the following two conditions are equivalent:

(1) $G_l^m(\rho) = G_l^c(\rho)$ holds.

(2) For every CPS $|\phi_m\rangle$ of $\Lambda_m^2(\rho)$ there exists an optimal decomposition $\{P_i,|\Psi_i\rangle\}$ of $G_l^c(\rho)$ for which $\Lambda_m^2(\rho) = \Lambda^2(|\Psi_i\rangle) = |\langle\phi_m|\Psi_i\rangle|^2$ holds for all i.

Proof.

1. $\Rightarrow$ 2.: Because of $G_{\rm l}^{\rm c}(\rho) = G_{\rm l}^{\rm m}(\rho)$, the two inequalities in (34) must become equalities, from which it follows that for every CPS $|\phi_{\rm m}\rangle$ there exists a decomposition $\{P_i,|\Psi_i\rangle\}$ of ρ which is optimal for $G_{\rm l}^{\rm c}(\rho)$ and for which $\Lambda^2(|\Psi_i\rangle) = |\langle\phi_{\rm m}|\Psi_i\rangle|^2 = \langle\phi_{\rm m}|\rho|\phi_{\rm m}\rangle = \Lambda_{\rm m}^2(\rho)$ holds for all i.

2. $\Rightarrow$ 1.: Let $|\phi_{\rm m}\rangle$ be a CPS for $\Lambda_{\rm m}^2(\rho)$ and let $\{P_i,|\Psi_i\rangle\}$ be an optimal decomposition of $G_{\rm l}^{\rm c}(\rho)$ for which $\Lambda_{\rm m}^2(\rho) = \Lambda^2(|\Psi_i\rangle) = |\langle\phi_{\rm m}|\Psi_i\rangle|^2$ holds for all i. Then, $G_{\rm l}^{\rm c}(\rho) = -\sum_i P_i \log\Lambda_{\rm m}^2(\rho) = -\log\Lambda_{\rm m}^2(\rho) = G_{\rm l}^{\rm m}(\rho)$. ■

In Sec. IV E this theorem will be demonstrated by means of the maximally correlated states. With regard to the linear measures, it will be shown in Theorem 25 that $G^{\rm m}(\rho) = G^{\rm f/c}(\rho)$ holds if and only if ρ is pure.

C. Comparison between $G^{\rm t}$ and $E_{\rm T}$

Theorem 21. $G^{f/c}(\rho) \geqslant \widetilde{E_T}(\rho)$ holds for all states ρ.

Proof. $G^{\rm f}(\rho) = 1 - F^2(\rho,\sigma_{\rm f}) \overset{(8)}{\geqslant} D_{\rm T}^2(\rho,\sigma_{\rm f}) \geqslant \min_{\sigma\in{\rm SEP}} D_{\rm T}^2(\rho,\sigma) = \widetilde{E_{\rm T}}(\rho)$. ■

Theorem 22. $G^t(\rho) \leqslant G^m(\rho)$ holds for all states ρ.

Proof. Let $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ be an arbitrary decomposition of ρ. Then

$$G^{\rm t}(\rho) = \min_{|\varphi\rangle\in{\rm PRO}} D_{\rm T}^2(\rho,|\varphi\rangle) \leqslant \min_{|\varphi\rangle\in{\rm PRO}}\left[\sum_i p_i D_{\rm T}^2(|\psi_i\rangle,|\varphi\rangle)\right]$$

$$\overset{(8)}{=} \min_{|\varphi\rangle\in{\rm PRO}}\left[\sum_i p_i(1-|\langle\varphi|\psi_i\rangle|^2)\right]$$

$$= 1 - \max_{|\varphi\rangle\in{\rm PRO}}\langle\varphi|\rho|\varphi\rangle = G^{\rm m}(\rho),$$

where the inequality follows from the convexity of $D_{\rm T}^2$. ■

The inequality $G^{\rm t}(\rho) \leqslant G^{\rm m}(\rho)$ can be strict, as seen for $\rho = \mathbb{1}/d^2$, the maximally mixed state of two qudits: $G^{\rm t}(\rho) = (1-\frac{1}{d^2})^2$ (cf. Appendix B), which is always smaller than $G^{\rm m}(\rho) = 1-\frac{1}{d^2}$.

D. Inequalities and hierarchies

Using the results from the preceding sections, we find the following inequality chains that include all the GM definitions considered in this paper. These inequality hierarchies are summarized and visualized in Fig. 1.

Theorem 23. The following inequalities hold for all states ρ:

(1) $\widetilde{E_T}(\rho) \leqslant G^t(\rho) \leqslant G^m(\rho)$

(2) $\widetilde{E_T}(\rho) \leqslant G^{f/c}(\rho) \leqslant G^m(\rho)$

(3) $G^m(\rho) \leqslant G_l^m(\rho) \leqslant E_R(\rho) + S(\rho)$

(4) $G^{f/c}(\rho) \leqslant G_l^f(\rho) \leqslant G_l^c(\rho) \leqslant G_l^m(\rho)$

Proof.

1: The first inequality was shown in Sec. III B 4 and the second one in Theorem 22.

2: The first inequality was shown in Theorem 21, and the second one in Corollary 19.

3: The first inequality follows from Lemma 2, and the second one was shown in Refs. [11,12].

4: The first inequality follows from Lemma 2, the second was shown in (24), and the third in Theorem 18. ■

To verify that no inequality relationship exists for measures that are not vertically connected in Fig. 1, e.g., $G^{\rm t}$ and $G_{\rm l}^{\rm c}$, we need to find $\rho_1,\ \rho_2 \in \mathcal{S}(\mathcal{H})$ so $G^{\rm t}(\rho_1) < G_{\rm l}^{\rm c}(\rho_1)$, and $G^{\rm t}(\rho_2) > G_{\rm l}^{\rm c}(\rho_2)$. The absence of an inequality relationship will be denoted as $G^{\rm t} \gtrless G_{\rm l}^{\rm c}$.

Proposition 24. $G^m \gtrless G_l^c$, $G^m \gtrless G_l^f$, $G^t \gtrless G_l^c$, $G^t \gtrless G_l^f$, and $G^t \gtrless G^{f/c}$.

Proof. Let ρ_1 be a genuinely mixed separable state and $\rho_2 = |\psi\rangle\langle\psi|$ a pure entangled state. Then, $G^{\rm m}(\rho_1) > G_{\rm l}^{\rm c}(\rho_1) = G_{\rm l}^{\rm f}(\rho_1) = 0$ and $G^{\rm m}(\rho_2) = G(|\psi\rangle) < G_{\rm l}(|\psi\rangle) = G_{\rm l}^{\rm c}(\rho_2) = G_{\rm l}^{\rm f}(\rho_2)$, from which it follows that $G^{\rm m} \gtrless G_{\rm l}^{\rm c}$ and $G^{\rm m} \gtrless G_{\rm l}^{\rm f}$.

Equivalently, $G^{\rm t}(\rho_1) > G_{\rm l}^{\rm c}(\rho_1) = G_{\rm l}^{\rm f}(\rho_1) = 0$ and $G^{\rm t}(\rho_2) = G(|\psi\rangle) < G_{\rm l}(|\psi\rangle) = G_{\rm l}^{\rm c}(\rho_2) = G_{\rm l}^{\rm f}(\rho_2)$, from which it follows that $G^{\rm t} \gtrless G_{\rm l}^{\rm c}$ and $G^{\rm t} \gtrless G_{\rm l}^{\rm f}$.

Let ρ_1 again be a genuinely mixed separable state, and $\rho_2 = q|00\rangle\langle 00| + (1-q)|\psi\rangle\langle\psi|$ the two-qutrit mixed entangled state of Proposition 13. Then, $G^{\rm t}(\rho_1) > G^{\rm f/c}(\rho_1) = 0$, and from item 2 of Lemma 12 it follows that $G^{\rm f/c}(\rho_2) = \frac{1-q}{2}$. By choosing $|\phi\rangle = |00\rangle$ in (25), we have $G^{\rm t}(\rho_2) \leqslant (1-q)^2$. So $G^{\rm t}(\rho_2) < G^{\rm f/c}(\rho_2)$ for $q \in (\frac{1}{2},1)$. Hence, $G^{\rm t} \gtrless G^{\rm f/c}$. ■

From Theorem 23 and Fig. 1 we see that $G^{\rm m}$ and $G_{\rm l}^{\rm m}$ are upper bounds for all linear and logarithmic GM definitions, respectively. Since $\Lambda_{\rm m}^2(\rho)$ can be computed for many prominent states (see, e.g., Ref. [8]), these upper bounds are readily accessible. For pure states the bounds are strict, and the more mixed a given state ρ is, the weaker the bounds are.

E. Partitioning of state space

As shown in the previous subsection, the inequalities

$$0 \leqslant G^{\rm f/c}(\rho) \leqslant G^{\rm m}(\rho), \tag{35}$$

$$0 \leqslant G_{\rm l}^{\rm f}(\rho) \leqslant G_{\rm l}^{\rm c}(\rho) \leqslant G_{\rm l}^{\rm m}(\rho), \tag{36}$$

hold for all states. Here we will see that these inequalities provide a physically meaningful partitioning of state space. For this we first prove that the measures in the above inequalities coincide if and only if ρ is pure.

Theorem 25. For any state ρ the following three conditions are equivalent:

(1) $G^{f/c}(\rho) = G^m(\rho)$

(2) $G_l^f(\rho) = G_l^c(\rho) = G_l^m(\rho)$

(3) ρ is pure, i.e., $\rho = |\psi\rangle\langle\psi|$

Proof.

1 $\Rightarrow$ 2: If $G^{\rm f}(\rho) = G^{\rm m}(\rho)$, then $\Lambda_{\rm f}^2(\rho) = \Lambda_{\rm m}^2(\rho)$ and $G_{\rm l}^{\rm f}(\rho) = G_{\rm l}^{\rm m}(\rho)$. From (36) it follows that $G_{\rm l}^{\rm f}(\rho) = G_{\rm l}^{\rm c}(\rho) = G_{\rm l}^{\rm m}(\rho)$.

2 $\Rightarrow$ 1: If $G_{\rm l}^{\rm f}(\rho) = G_{\rm l}^{\rm m}(\rho)$, then $\Lambda_{\rm f}^2(\rho) = \Lambda_{\rm m}^2(\rho)$ and $G^{\rm f}(\rho) = G^{\rm m}(\rho)$.

3 $\Rightarrow$ 2: Obvious from the definition of GM for pure states.

2 $\Rightarrow$ 3: Let $\{p_i,|\psi_i\rangle\}$ be some decomposition of ρ, and let $|\phi_{\rm m}\rangle$ be a CPS of $\Lambda_{\rm m}^2(\rho)$. Using the concavity of $\Lambda_{\rm f}^2(\rho)$, we

have

$$\begin{aligned}\Lambda_{\mathrm{f}}^2(\rho) &\geqslant \sum_i p_i \Lambda^2(|\psi_i\rangle) = \sum_i p_i \max_{|\varphi_i\rangle\in\mathrm{PRO}} |\langle\varphi_i|\psi_i\rangle|^2 \\ &\geqslant \sum_i p_i |\langle\phi_{\mathrm{m}}|\psi_i\rangle|^2 = \langle\phi_{\mathrm{m}}|\rho|\phi_{\mathrm{m}}\rangle \\ &= \max_{|\varphi\rangle\in\mathrm{PRO}} \langle\varphi|\rho|\varphi\rangle = \Lambda_{\mathrm{m}}^2(\rho)\,. \end{aligned} \quad (37)$$

From $G_1^{\mathrm{f}}(\rho) = G_1^{\mathrm{m}}(\rho)$ it follows that $\Lambda_{\mathrm{f}}^2(\rho) = \Lambda_{\mathrm{m}}^2(\rho)$, and therefore the inequalities in (37) must become equalities. From the second inequality of (37) it follows that there exists a $|\phi_{\mathrm{m}}\rangle \in \mathrm{PRO}$ that is a CPS for all $|\psi_i\rangle$. Furthermore, since the choice of decomposition was arbitrary, this $|\phi_{\mathrm{m}}\rangle$ must be a CPS for all $|\psi_i\rangle$ of all conceivable decompositions $\{p_i,|\psi_i\rangle\}$ of ρ. As shown below, this is possible only for pure states, i.e., rank $\rho = 1$.
Assume $r = \mathrm{rank}\,\rho \geqslant 2$. Because of $G_1^{\mathrm{f}}(\rho) = G_1^{\mathrm{c}}(\rho)$, it follows from Theorem 7 that there exists a decomposition $\{P_i,|\Psi_i\rangle\}$ of ρ where the $|\Psi_i\rangle$ are all equally entangled. Since $|\phi_{\mathrm{m}}\rangle$ is a common CPS, we can write $|\Psi_1\rangle$ and $|\Psi_2\rangle$ as

$$|\Psi_1\rangle = \alpha|\phi_{\mathrm{m}}\rangle + \sqrt{1-\alpha^2}|\phi_1^{\perp}\rangle,$$
$$|\Psi_2\rangle = \alpha|\phi_{\mathrm{m}}\rangle + \sqrt{1-\alpha^2}|\phi_2^{\perp}\rangle,$$

with $\alpha \in (0,1)$. Here $|\phi_1^{\perp}\rangle$ and $|\phi_2^{\perp}\rangle$ are states orthogonal to $|\phi_{\mathrm{m}}\rangle$, and because of rank $\rho \geqslant 2$, we can assume that $|\phi_1^{\perp}\rangle - |\phi_2^{\perp}\rangle$ is a nonzero vector. Clearly, $\Lambda^2(|\Psi_1\rangle) = \Lambda^2(|\Psi_2\rangle) = \alpha^2 > 0$. Starting with $\{P_i,|\Psi_i\rangle\}$, we use (19) to construct a new decomposition $\{q_i,|\Theta_i\rangle\}$ of ρ by using the $r \times r$ unitary matrix

$$U = \begin{pmatrix} \sqrt{\frac{P_2}{P_1+P_2}} & -\sqrt{\frac{P_1}{P_1+P_2}} & 0 \\ \sqrt{\frac{P_1}{P_1+P_2}} & \sqrt{\frac{P_2}{P_1+P_2}} & 0 \\ 0 & 0 & \mathbb{1}_{r-2} \end{pmatrix},$$

where $\mathbb{1}_{r-2}$ denotes the $(r-2)$-dimensional unit matrix. With this we have

$$\begin{aligned}|\Theta_1\rangle &\propto \sqrt{P_2}\sqrt{P_1}|\Psi_1\rangle - \sqrt{P_1}\sqrt{P_2}|\Psi_2\rangle \\ &= \sqrt{P_1P_2}\sqrt{1-\alpha^2}(|\Phi_1^{\perp}\rangle - |\Phi_2^{\perp}\rangle),\end{aligned}$$

and, therefore, $\langle\phi_{\mathrm{m}}|\Theta_1\rangle = 0$. This is a contradiction to the requirement that $|\phi_{\mathrm{m}}\rangle$ is also a CPS for $|\Theta_1\rangle$. This completes the proof. ■

According to Theorem 25, G^{m} (or G_1^{m}) coincides with the proper entanglement measure $G^{\mathrm{f/c}}$ (or G_1^{f}) only for pure states, thus further reinforcing the observation that $\Lambda_{\mathrm{m}}^2(\rho)$ assesses the entanglement as well as the mixedness of a state ρ. Together with the known fact that $\Lambda_{\mathrm{f}}^2(\rho) < 1$ if and only if ρ is entangled, we can use (35) to partition the state space into four subsets, $\mathcal{S}(\mathcal{H}) = A \cup B \cup C \cup D$, corresponding to pure separable, pure entangled, mixed separable, and mixed entangled states, respectively. As shown in Table I and Fig. 2, this partitioning is done by determining whether the inequalities in (35) are strict or become equalities. The inequalities (35) between the logarithmic measures can also be used for partitioning $\mathcal{S}(\mathcal{H})$, and in that case, the subset of mixed entangled states is further divided into three subsets, $D = D_1 \cup D_2 \cup D_3$, because $G^{\mathrm{f/c}}(\rho) < G^{\mathrm{m}}(\rho)$ corresponds to the three possible cases $G_1^{\mathrm{f}}(\rho) < G_1^{\mathrm{c}}(\rho) < G_1^{\mathrm{m}}(\rho)$, $G_1^{\mathrm{f}}(\rho) = G_1^{\mathrm{c}}(\rho) < G_1^{\mathrm{m}}(\rho)$, and $G_1^{\mathrm{f}}(\rho) < G_1^{\mathrm{c}}(\rho) = G_1^{\mathrm{m}}(\rho)$.

Since $G_1^{\mathrm{f}}(\rho) = G_1^{\mathrm{c}}(\rho)$ holds for isotropic states and all two-qubit states, these states belong to the set D_2. In particular, for the special case of two qubits ($\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$) we have $D = D_2$, i.e., D_1 and D_3 are empty. From this one could conjecture that generic mixed entangled states belong to D_2. However, for maximally correlated states, it is clear from item 5 of Lemma 12 that most states do not belong to D_2. The following theorem elucidates the relationship between the parameters of rank-2 maximally correlated states and the subgroups D_1, D_2, and D_3. For this we note that $\Lambda_{\mathrm{m}}^2(\rho)$ can be easily calculated for maximally correlated states of the form in Proposition 14 as $\Lambda_{\mathrm{m}}^2(\rho) = \max\{\frac{q}{m},\frac{1-q}{n}\}$.

Theorem 26. Depending on the value of the parameters $m,n \in \mathbb{N}$ with $\frac{m}{n} \leqslant 1$ and $q \in (0,1)$, the rank-2 maximally correlated states of Proposition 14 belong to either of the three subsets of genuinely mixed entangled states:

D_2, i.e., $G_l^f(\rho) = G_l^c(\rho) < G_l^m(\rho)$: for $\frac{m}{n} = 1$,

D_3, i.e., $G_l^f(\rho) < G_l^c(\rho) = G_l^m(\rho)$: for $\frac{m}{n} < \frac{1}{e}$ and $q \geqslant \frac{em}{n}$,

D_1, i.e., $G_l^f(\rho) < G_l^c(\rho) < G_l^m(\rho)$: for all other parameter values.

Proof. From item 5 of Lemma 12 it follows that $G_1^{\mathrm{f}}(\rho) = G_1^{\mathrm{c}}(\rho)$ if and only if $m = n$. Hence, states belong to D_2 if and only if $\frac{m}{n} = 1$. For $\frac{m}{n} < 1$ one can distinguish between $\rho \in D_1$ and $\rho \in D_3$ by determining whether $G_1^{\mathrm{c}}(\rho) \leqslant G_1^{\mathrm{m}}(\rho)$ is strict.

Recall that $G_1^{\mathrm{m}}(\rho) = \min\{\log(\frac{m}{q}),\log(\frac{n}{1-q})\}$. If $q \geqslant \frac{em}{n}$, then $\frac{m}{q} \leqslant \frac{n}{e} < \frac{n}{1-q}$, yielding $G_1^{\mathrm{m}}(\rho) = \log(\frac{m}{q})$. Therefore, if $\frac{m}{n} < \frac{1}{e}$ and $q \geqslant \frac{em}{n}$ (second case of Proposition 14), then $G_1^{\mathrm{c}}(\rho) = G_1^{\mathrm{m}}(\rho) = \log(\frac{m}{q})$, yielding $\rho \in D_3$.

Regarding the first and third cases of Proposition 14, it is seen from its proof (including Lemma 31 and its proof) that for parameter values in the interior of the domain (i.e., excluding $\frac{m}{n} = 1$) the value of $G_1^{\mathrm{c}}(\rho)$ is strictly smaller than both $\log(\frac{m}{q})$ and $\log(\frac{n}{1-q})$. Therefore, $G_1^{\mathrm{c}}(\rho) < G_1^{\mathrm{m}}(\rho) = \min\{\log(\frac{m}{q}),\log(\frac{n}{1-q})\}$ holds, so $\rho \in D_1$. ■

From Theorem 26 we see that the states belonging to D_3 precisely coincide with those outlined in the second case of Proposition 14. This allows us to demonstrate Theorem 20: Let $\rho \in D_3$ be a rank-2 maximally correlated state. Then every CPS $|\phi_{\mathrm{m}}\rangle$ of $\Lambda_{\mathrm{m}}^2(\rho)$ necessarily has the form $|\phi_{\mathrm{m}}\rangle = (\sum_{i=1}^{m} q_i|i\rangle) \otimes (\sum_{i=1}^{m} q_i^*|i\rangle)$ with $\sum_i |q_i|^2 = 1$. According to Proposition 14, an optimal decomposition for $G_1^{\mathrm{c}}(\rho)$ is $\rho = \frac{1}{2}|\Psi_+\rangle\langle\Psi_+| + \frac{1}{2}|\Psi_-\rangle\langle\Psi_-|$ with $|\Psi_\pm\rangle = \sqrt{q}|\psi_m\rangle \pm \sqrt{1-q}|\psi_n\rangle$. Using this decomposition, we obtain $\Lambda_{\mathrm{m}}^2(\rho) = \Lambda^2(|\Psi_\pm\rangle) = |\langle\phi_{\mathrm{m}}|\Psi_\pm\rangle|^2 = \frac{q}{m}$, thus verifying Theorem 20.

For maximally correlated qutrit states the only possible value of $\frac{m}{n}$ is $\frac{1}{2}$, so all states lie in D_1. In contrast to this, for four levels ($d = 4$), there are rank-2 maximally correlated states in each of the three sets D_1, D_2, and D_3.

In the following we determine whether the various subsets of the genuinely mixed states, $C \cup D = C \cup D_1 \cup D_2 \cup D_3$, are convex. C and $C \cup D$ are clearly convex sets, while D is not. All other subsets are investigated in the following lemma.

Proposition 27. The following sets are not convex: D_1, D_2, D_3, $D_1 \cup D_2$, $D_1 \cup D_3$, $D_2 \cup D_3$, $C \cup D_1$, $C \cup D_2$, $C \cup D_3$, and $C \cup D_1 \cup D_3$. Regarding $C \cup D_1 \cup D_2$ and $C \cup D_2 \cup D_3$, at least one of the two sets is not convex.

Proof. To prove that D_1, $D_1 \cup D_3$, $C \cup D_1$, and $C \cup D_1 \cup D_3$ are not convex, it suffices to find $\rho_1,\rho_2 \in D_1$ so $\frac{1}{2}(\rho_1 + \rho_2) \in D_2$. Using the notation $|\Psi_{a...z}\rangle := \frac{1}{\sqrt{z-a+1}}(|aa\rangle + \cdots + |zz\rangle)$ for MES, we choose ρ_1,ρ_2 to be

$$\rho_\pm = \frac{1}{2}|\Psi_{12}\rangle\langle\Psi_{12}| + \frac{1}{2}|\Psi^\pm_{3456}\rangle\langle\Psi^\pm_{3456}|, \quad \text{with}$$
$$|\Psi^\pm_{3456}\rangle := \frac{1}{\sqrt{2}}(|\Psi_{34}\rangle \pm |\Psi_{56}\rangle).$$

Evidently, ρ_+ and ρ_- are LU equivalent, so they lie in the same set D_i ($i = 1,2,3$). Since $\frac{m}{n} = \frac{1}{2}$ for ρ_+, it follows from Theorem 26 that $\rho_\pm \in D_1$. On the other hand,

$$\rho = \tfrac{1}{2}(\rho_+ + \rho_-) = \tfrac{1}{2}|\Psi_{12}\rangle\langle\Psi_{12}| + \tfrac{1}{4}|\Psi_{34}\rangle\langle\Psi_{34}| + \tfrac{1}{4}|\Psi_{56}\rangle\langle\Psi_{56}|,$$

so it follows from item 5 of Lemma 12 that $\rho \in D_2$.

Next, we prove that D_3 and $C \cup D_3$ are not convex by finding $\sigma_1,\sigma_2,\sigma_3 \in D_3$ that yield $\sigma = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) \in D_2$. For this, consider

$$\sigma_i = q|\Psi_{12}\rangle\langle\Psi_{12}| + (1-q)|\Psi^i_{345678}\rangle\langle\Psi^i_{345678}|, \quad \text{with}$$
$$\frac{e}{3} \leqslant q < 1, \quad \text{and}$$
$$|\Psi^i_{345678}\rangle := \frac{1}{\sqrt{3}}\big(|\Psi_{34}\rangle + e^{i\frac{2\pi i}{3}}|\Psi_{56}\rangle + e^{i\frac{4\pi i}{3}}|\Psi_{78}\rangle\big),$$

for $i = 1,2,3$. Evidently, the σ_i are LU equivalent, and since $\frac{m}{n} = \frac{1}{3}$ and $q \geqslant \frac{e}{3}$, it follows from Theorem 26 that $\sigma_i \in D_3$. On the other hand,

$$\sigma = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) = q|\Psi_{12}\rangle\langle\Psi_{12}| + \frac{1-q}{3}(|\Psi_{34}\rangle\langle\Psi_{34}| + |\Psi_{56}\rangle\langle\Psi_{56}| + |\Psi_{78}\rangle\langle\Psi_{78}|),$$

so it follows from item 5 of Lemma 12 that $\sigma \in D_2$.

To prove that D_2, $D_1 \cup D_2$, and $D_2 \cup D_3$ are not convex, we consider $\rho_\pm = \frac{1}{2}|\Psi^\pm_{12}\rangle\langle\Psi^\pm_{12}| + \frac{1}{2}|\Psi^\pm_{34}\rangle\langle\Psi^\pm_{34}| \in D_2$, which yields $\rho = \frac{1}{2}(\rho_+ + \rho_-) = \frac{1}{4}\mathbb{1} \in C$.

To prove that $C \cup D_2$ is not convex, we consider $\rho_\pm = \frac{1}{2}|\Psi^\pm_{12}\rangle\langle\Psi^\pm_{12}| + \frac{1}{2}|\Psi_{34}\rangle\langle\Psi_{34}| \in D_2$, and the genuinely mixed entangled state $\rho = \frac{1}{2}(\rho_+ + \rho_-) = \frac{1}{4}|11\rangle\langle 11| + \frac{1}{4}|22\rangle\langle 22| + \frac{1}{2}|\Psi_{34}\rangle\langle\Psi_{34}| \in D$. From item 5 of Lemma 12 it follows that $\rho \notin D_2$, hence $\rho \in D_1 \cup D_3$. Although we do not know whether $\rho \in D_1$ or $\rho \in D_3$, we can ascertain that no more than one of the two sets $C \cup D_1 \cup D_2$ and $C \cup D_2 \cup D_3$ can be convex. ■

Although partially answered by Proposition 27, it is still unknown whether $C \cup D_1 \cup D_2$ or $C \cup D_2 \cup D_3$ are convex.

V. GRAPH STATES AND CLUSTER STATES

Graph states are an important class of states for quantum information [57]. A subset of them, the cluster states, are the central ingredient for one-way quantum computation [58]. Here we show that a large class of graph states, including all cluster states, have a "universal" closest separable state that minimizes several inequivalent distance measures. This property helps to prove the previous Corollary 6.

Consider a general pure graph state $|G\rangle$ with underlying graph $G = (V,E)$, where V is the set of vertices and E is the set of edges. The *maximum independent set* α is the largest possible set of nonadjacent vertices, and the *minimum vertex cover* β is the complement of α, i.e., $\alpha + \beta = V$. The minimum vertex cover can be thought of as the minimal set of qubits that needs to be measured in the computational basis to completely disentangle the graph state.

As outlined in Ref. [43], the stabilizer $\mathcal{S}$ of $|G\rangle$ is generated by n generators $\{g_j\}_{j=1}^n$, and these generators stabilize a unique state, namely $|G\rangle$. If one or more of the generators from the generating set of $\mathcal{S}$ are discarded, the smaller set generates a new Abelian group $\mathcal{S}'$ which now stabilizes a set of states $\{|\psi_i\rangle\}$ rather than a unique $|G\rangle$. Depending on the structure of the generating set of $\mathcal{S}'$, the states $\{|\psi_i\rangle\}$ may or may not be entangled. In Ref. [43] it is shown that the optimal way of discarding generators, such that the stabilized states $\{|\psi_i\rangle\}$ are product states, is to discard generators corresponding to the vertices of the minimum vertex cover. So if we only keep the generators corresponding to the maximum independent set $\{g_j | j \in \alpha\}$, denoting the correspondingly generated Abelian group as $\mathcal{S}_\alpha$, the states it stabilizes $\{|\psi_i^\alpha\rangle\}$ are all product states. These states form the basis vectors used below.

Ignoring possible negative amplitudes that are not important here, $|G\rangle$ can be written as an equal superposition of the basis vectors,

$$|G\rangle = \frac{1}{\sqrt{D_\alpha}}\sum_{i=1}^{D_\alpha}|\psi_i^\alpha\rangle, \tag{38}$$

where D_α is the number of states $|\psi_i^\alpha\rangle$ and is related to the cardinality of the minimum vertex cover as $D_\alpha = 2^{|\beta|}$. In other words, for each generator discarded from the generating set, the size of the set of stabilized states doubles. The decomposition (38) is of minimal rank for graph states whose underlying graphs satisfy certain conditions [43]. This is the case for all bipartite (two-colorable) graphs, which includes all cluster states of arbitrary size and dimension. However, there also exist many non-two-colorable states that satisfy the conditions.

In the following we assume that $|G\rangle$ satisfies the conditions, i.e., the decomposition (38) is of minimal rank. Since the $|\psi_i^\alpha\rangle$ are product states, it immediately follows from (38) that any of the $|\psi_i^\alpha\rangle$ is a CPS, i.e., $\Lambda^2(|G\rangle) = |\langle G|\psi_i^\alpha\rangle|^2 = D_\alpha^{-1} = 2^{-|\beta|}$. Correspondingly, the separable state

$$\delta = \frac{1}{D_\alpha}\sum_{i=1}^{D_\alpha}|\psi_i^\alpha\rangle\langle\psi_i^\alpha| \tag{39}$$

was found to be a CSS for the REE [43]. Here we show that it is also the CSS in terms of the Bures distance (5) and the trace distance (3).

Theorem 28. Let $|G\rangle$ be a graph state of the form (38) with minimal rank. Then (39) is a closest separable state with respect to the quantum relative entropy, the Bures distance, and the trace distance.

Proof. For the quantum relative entropy this was shown in Ref. [59], and we also know that $\Lambda^2(|G\rangle) = D_\alpha^{-1}$. Hence, $\max_{\sigma\in\text{SEP}} F^2(|G\rangle,\sigma) = \max_{|\phi\rangle\in\text{PRO}} |\langle G|\phi\rangle|^2 = D_\alpha^{-1}$.

From $F^2(|G\rangle,\delta) = |\langle G|\delta|G\rangle| = D_\alpha^{-1}$ it then follows that δ minimizes the Bures distance (5).

Next consider the trace distance. The inequality $\min_{\sigma\in\text{SEP}} D_{\text{T}}(|G\rangle,\sigma) \geqslant 1 - \max_{\sigma\in\text{SEP}} F^2(|G\rangle,\sigma) = 1 - D_\alpha^{-1}$ follows from (8b). To show that δ minimizes the trace distance, it therefore suffices to show that $D_{\text{T}}(|G\rangle,\delta) = 1 - D_\alpha^{-1}$. From (3) we know that $D_{\text{T}}(|G\rangle,\delta) = \frac{1}{2}\sum_{i=1}^{D_\alpha}|\lambda_i|$, where the λ_i are the eigenvalues of $A := |G\rangle\langle G| - \delta = D_\alpha^{-1}\sum_{i\neq j}|\psi_i^\alpha\rangle\langle\psi_j^\alpha|$. Taking the states $\{|\psi_i^\alpha\rangle\}$ as the basis, elementary linear algebra yields the nonzero eigenvalues of A as $\lambda_1 = \frac{D_\alpha - 1}{D_\alpha}$ and $\lambda_2 = \cdots = \lambda_{D_\alpha} = -\frac{1}{D_\alpha}$, hence $D_{\text{T}}(|G\rangle,\delta) = \frac{D_\alpha-1}{D_\alpha} = 1 - D_\alpha^{-1}$. This completes the proof. ■

Theorem 28 applies to all cluster states. In particular, for n qubit cluster states $|C_n\rangle$ the minimum vertex cover has the size $|\beta| = \lfloor\frac{n}{2}\rfloor$, yielding the cardinality $D_\alpha = 2^{\lfloor\frac{n}{2}\rfloor}$. For even n, this yields $\Lambda^2(|C_n\rangle) = 2^{-\frac{n}{2}}$ and $D_{\text{T}}(|C_n\rangle,\delta) = 1 - 2^{-\frac{n}{2}}$.

For graph states that do not satisfy the minimal rank condition, the state (39) generally does not minimize the three distance measures, but it nevertheless yields upper bounds on the distances and on the corresponding entanglement measures, the REE, the BE, and TE.

VI. CONCLUSION

In this paper we reviewed and studied seven different definitions of GM for arbitrary multipartite systems. Five of these are known (G^{c}, G^{f}, G_{l}^{c}, G^{m}, and G_{l}^{m}), one has previously received only little interest (G_{l}^{f}), and one has not been studied before (G^{t}). The entanglement axioms of the measures were investigated and are summarized in Table II. A remaining open question is whether G_{l}^{c} satisfies weak monotonicity, something we showed to be true at least for two-qubit states and isotropic states. A complete quantitative hierarchy between the measures was derived (shown in Fig. 1), and it was found that this hierarchy can be employed to partition the state space into pairwise disjoint sets with clear physical properties (pure versus mixed and separable versus entangled). This is summarized in Table I and Fig. 2.

As a by-product of Corollary 6, we found that for pure input states $\rho = |\psi\rangle\langle\psi|$ the trace distance $D_{\text{T}}(|\psi\rangle,\cdot)$ has in general no pure CSS. This is in stark contrast to the Bures distance, for which (14) implies that $D_{\text{B}}(|\psi\rangle,\cdot)$ always has at least one pure CSS. It is therefore not trivial to find states for whom the Bures and trace distance have a common CSS, something we did for a large class of graph states in Theorem 28.

With regard to the convex roof-based measures G^{c} and G_{l}^{c}, it was found that—unlike G^{f} and G_{l}^{f} or G^{m} and G_{l}^{m}—these two measures are not simple functions of each other and, in fact, do not even have the same ordering. Nevertheless, some connections between the two measures and their optimal decompositions could be made (Lemma 3, Theorem 7, and Corollary 8). For this, the maximally correlated states were particularly helpful, because their optimal decompositions for G_{l}^{c} depend qualitatively on their parameters (Lemma 12, Proposition 13, and Proposition 14). This way it could be shown in Corollary 15 that for some states G^{c} and G_{l}^{c} do not share any common optimal decomposition.

For the linear GM it is known that the problem of finding the optimal convex roof decomposition is equivalent to finding the closest separable state for the fidelity ($G^{\text{f}} \equiv G^{\text{c}}$) [13]. Somewhat surprisingly, we found that this is not the case for the logarithmic GM. Already for bipartite systems the two problems are in general inequivalent ($G_{\text{l}}^{\text{f}} \not\equiv G_{\text{l}}^{\text{c}}$) and need to be solved separately. Nevertheless, for two-qubit systems and for some classes of states, such as all isotropic states and some maximally correlated states, the two problems coincide. While G_{l}^{f} could be verified to be weakly monotonous, the weak monotonicity of G_{l}^{c} remains an open problem. At least for two-qubit systems this question can be answered in the affirmative, because $G_{\text{l}}^{\text{f}} \equiv G_{\text{l}}^{\text{c}}$ then holds. Another open question is whether there exist states ρ with $G_{\text{l}}^{\text{c}}(\rho) = G_{\text{l}}^{\text{f}}(\rho)$ for which the set of optimal decompositions of $G^{\text{c}}(\rho)$ is strictly larger than that of $G_{\text{l}}^{\text{c}}(\rho)$, cf. Theorem 7.

As G^{m} and G_{l}^{m} assess the entanglement as well as the mixedness of states [26], they are not entanglement measures. On the other hand, $G^{\text{f/c}}$ is the only known definition of GM that yields a strong entanglement measure. Because of its strong monotonicity and convexity, $G^{\text{f/c}}$ never increases as states become more mixed. Recalling that $G^{\text{f/c}}(|\psi\rangle) = G^{\text{m}}(|\psi\rangle)$ for pure states, the inequality $G^{\text{f/c}}(\rho) \leqslant G^{\text{m}}(\rho)$ becomes intuitively clear. What is more, Theorem 25 states that this inequality turns into an equality only if ρ is pure. Therefore, $\Delta G(\rho) := G^{\text{m}}(\rho) - G^{\text{f/c}}(\rho)$ can be considered an entropic quantity depending on the mixedness of ρ, akin to the linear entropy. Equivalently, $\Delta G_{\text{l}}(\rho) := G_{\text{l}}^{\text{m}}(\rho) - G_{\text{l}}^{\text{f}}(\rho)$ can be considered an alternative to the von Neumann entropy. In contrast to this, $G_{\text{l}}^{\text{m}}(\rho) - G_{\text{l}}^{\text{c}}(\rho)$ cannot be expected to be a meaningful entropy quantifier, because this quantity can be zero for genuinely mixed states (e.g., maximally correlated states belonging to set D_3, cf. Theorem 26). For such states the relationship between their closest product state and optimal decomposition was derived in Theorem 20. It should also be noted that G^{m} and G_{l}^{m} are readily accessible upper bounds for all linear and logarithmic GM definitions, respectively. These bounds can be easily computed from $\Lambda_{\text{m}}^2(\rho)$, but they become weaker as ρ becomes more mixed.

The newly introduced extension of the linear GM by means of the trace distance, G^{t}, is not an entanglement measure, and it is yet unclear what its benefits or operational implications for the study of multipartite entanglement are. In contrast to this, the little-known quantity G_{l}^{f}, which is closely related through Λ_{f}^2 to the well-known definitions G^{f} and G^{c}, has many desirable properties: It satisfies normalization, convexity, and weak monotonicity and is zero for separable states. Most importantly, G_{l}^{f} is so far the only known definition of GM that yields a normalized entanglement measure (for G_{l}^{c} the question of weak monotonicity remains open). We therefore propose G_{l}^{f} as the preferred definition of GM for studies where normalization is a desirable feature, such as quantitative entanglement characterization, entanglement scaling, or comparison with other multipartite entanglement measures.

Even if G_{l}^{c} should be verified as a weak entanglement measure, G_{l}^{f} has some other benefits over the logarithmic convex roof: The value of $G_{\text{l}}^{\text{f}}(\rho)$ is immediately known if $\Lambda_{\text{f}}^2(\rho)$ is known, which is the case if either $G^{\text{f}}(\rho)$ or $G^{\text{c}}(\rho)$ have been computed. Furthermore, G_{l}^{f} is based on the fidelity, a widely studied and physically meaningful distance in quantum information theory, whereas G_{l}^{c} is based on an abstract mathematical definition.

In conclusion, for most situations either of the two fidelity-based definitions $G^{\mathrm{f/c}}$ or G^{f}_1 could be regarded as the best choice of GM. Since the two definitions have the same ordering and are closely related via Λ^2_{f}, the choice depends merely on whether normalization or strong monotonicity is more desirable. The entanglement of a state ρ can then be found either by computing the maximal fidelity to separable states, $\Lambda^2_{\mathrm{f}}(\rho)$, or by finding the optimal decomposition for the linear convex roof. Furthermore, the Bures entanglement (10b) and the Groverian entanglement $E_{\mathrm{Gr}}(\rho) = G^{\mathrm{f}}(\rho)^{1/2}$ are closely related to Λ^2_{f}, and therefore to G^{f} and G^{f}_1 themselves, with all measures having the same ordering. Knowing the value of either of these provides lower bounds to $G^{\mathrm{m}}(\rho)$, $G^{\mathrm{c}}_1(\rho)$ as well as $G^{\mathrm{m}}_1(\rho)$. Finally, defining GM though the maximal fidelity is also the most straightforward definition from a historical viewpoint, because GM was originally defined by the maximal fidelity between pure states.

ACKNOWLEDGMENTS

We thank Mark Wilde for pointing out an operational meaning of (22). L.C. was mainly supported by MITACS and NSERC. The CQT is funded by the Singapore MoE and the NRF as part of the Research Centres of Excellence programme. M.H. acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) through Grants No. 23-01770, No. 23540463, and No. 23240001. L.C. and M.H. were also funded by the Singapore National Research Foundation under NRF Grant No. NRF-NRFF2013-01.

APPENDIX A: WEAK MONOTONICITY OF G^{c}_1 FOR TWO QUBITS

It is known that Wootters's concurrence $C(\rho)$ is a weak entanglement monotone for two-qubit states, i.e., $C(\rho) \geqslant C(\sigma)$ holds for any trace-preserving quantum operation $\rho \mapsto \sigma$ [6,52]. Using the monotonically increasing function $f(x) = -\log \frac{1+\sqrt{1-x^2}}{2}$ and Lemma 29 below, it follows that $G^{\mathrm{c}}_1(\rho) = f(C(\rho)) \geqslant f(C(\sigma)) = G^{\mathrm{c}}_1(\sigma)$. So G^{c}_1 is weakly monotonous for two-qubit systems.

Lemma 29. Let ρ be an arbitrary two-qubit state. Then $G^{\mathrm{c}}_1(\rho) = -\log \frac{1+\sqrt{1-C(\rho)^2}}{2}$, where $C(\rho)$ is the concurrence.

Proof. The proof is similar to the derivation of $G^{\mathrm{c}}(\rho)$ in Ref. [6]. First, it is easy to verify that the claim holds for pure states. The function $f(x) = -\log \frac{1+\sqrt{1-x^2}}{2}$ is monotonically increasing and convex for $x \in [0,1]$. Suppose $\rho = \sum_i P_i |\Psi_i\rangle\langle\Psi_i|$ is an optimal decomposition for $G^{\mathrm{c}}_1(\rho)$. Then

$$G^{\mathrm{c}}_1(\rho) = \sum_i P_i G_1(|\Psi_i\rangle) = \sum_i P_i f(C(|\Psi_i\rangle)) \geqslant f\left(\sum_i P_i C(|\Psi_i\rangle)\right) \geqslant f(C(\rho)), \tag{A1}$$

where the inequalities follow from the convexity of $f(x)$ and $C(\rho)$ [52], respectively. On the other hand, Wootters found an optimal decomposition $\rho = \sum_i S_i |\Phi_i\rangle\langle\Phi_i|$ for the entanglement of formation, such that each $|\Phi_i\rangle$ has the same concurrence as ρ. With this decomposition we obtain

$$G^{\mathrm{c}}_1(\rho) \leqslant \sum_i S_i G_1(|\Phi_i\rangle) = \sum_i S_i f(C(|\Phi_i\rangle)) = f(C(\rho)). \tag{A2}$$

From (A2) and (A1) it follows that $G^{\mathrm{c}}_1(\rho) = f(C(\rho))$. ■

APPENDIX B: COUNTEREXAMPLE FOR CONCAVITY OF G^{t}

Consider the isotropic state $\rho_{\mathrm{iso}} = p\frac{\mathbb{1}}{d^2} + (1-p)|\Psi\rangle\langle\Psi|$, where $|\Psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|ii\rangle$. A counterexample for the concavity of G^{t} is found if

$$G^{\mathrm{t}}(\rho_{\mathrm{iso}}) < pG^{\mathrm{t}}(\mathbb{1}/d^2) + (1-p)G^{\mathrm{t}}(|\Psi\rangle\langle\Psi|) \tag{B1}$$

holds for some $p \in (0,1)$ and $d \geqslant 2$. Obviously, $G^{\mathrm{t}}(|\Psi\rangle\langle\Psi|) = G(|\Psi\rangle) = 1 - \frac{1}{d}$, and because of the isotropic nature of the maximally mixed state, for any $|\phi\rangle \in \mathcal{H}$ we can represent $\frac{\mathbb{1}}{d^2} - |\phi\rangle\langle\phi|$ in matrix form as $\mathrm{diag}(\frac{1}{d^2}-1, \frac{1}{d^2}, \ldots, \frac{1}{d^2})$ by choosing a basis with $|\phi\rangle$ as the first basis vector. Therefore,

$$\begin{aligned} G^{\mathrm{t}}\left(\frac{\mathbb{1}}{d^2}\right) &= \min_{|\varphi\rangle\in\mathrm{PRO}} \frac{1}{4}\left(\mathrm{Tr}\left|\frac{\mathbb{1}}{d^2} - |\varphi\rangle\langle\varphi|\right|\right)^2 \\ &= \frac{1}{4}\left[\left|\frac{1}{d^2}-1\right| + (d^2-1)\frac{1}{d^2}\right]^2 = \left(1-\frac{1}{d^2}\right)^2. \end{aligned}$$

Now consider $G^{\mathrm{t}}(\rho_{\mathrm{iso}})$. From a geometric viewpoint the relationship between $p\frac{\mathbb{1}}{d^2} + (1-p)|\Psi\rangle\langle\Psi|$ and $|\varphi\rangle \in \mathrm{PRO}$ is entirely determined by the angle between $|\Psi\rangle$ and $|\varphi\rangle$. We therefore parametrize $|\varphi\rangle = \sqrt{\alpha}|\Psi\rangle + \sqrt{1-\alpha}|\Psi^{\perp}\rangle$, where $|\Psi^{\perp}\rangle$ is some state orthogonal to $|\Psi\rangle$. From $\langle\Psi|\varphi\rangle = \sqrt{\alpha}$, $|\varphi\rangle \in \mathrm{PRO}$ and $\Lambda^2(|\Psi\rangle) = \frac{1}{d}$, it follows that $\alpha \leqslant \frac{1}{d}$. To show that this bound can be reached, we construct an example: The states $\{|\Psi_j\rangle\}_{j=0}^{d-1}$ with $|\Psi_j\rangle = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} e^{i\frac{2\pi jk}{d}}|kk\rangle$ form an orthonormalized basis of MES. Obviously $|\Psi\rangle = |\Psi_0\rangle$, and defining $|\Psi^{\perp}\rangle := \frac{1}{\sqrt{d-1}}\sum_{i=1}^{d-1}|\Psi_i\rangle$, we obtain $|\varphi\rangle = \sqrt{1/d}|\Psi\rangle + \sqrt{(d-1)/d}|\Psi^{\perp}\rangle = \frac{1}{d}\sum_{j=0}^{d-1}\sum_{k=0}^{d-1} e^{i\frac{2\pi jk}{d}}|kk\rangle = |00\rangle \in \mathrm{PRO}$.

We now write

$$\begin{aligned} \rho_{\mathrm{iso}} - |\varphi\rangle\langle\varphi| = {} & p\frac{\mathbb{1}}{d^2} + (1-p)|\Psi\rangle\langle\Psi| - \alpha|\Psi\rangle\langle\Psi| \\ & - \sqrt{\alpha(1-\alpha)}(|\Psi\rangle\langle\Psi^{\perp}| + |\Psi^{\perp}\rangle\langle\Psi|) \\ & - (1-\alpha)|\Psi^{\perp}\rangle\langle\Psi^{\perp}|, \end{aligned} \tag{B2}$$

and using $|\Psi\rangle$ and $|\Psi^{\perp}\rangle$ as the first two basis vectors for the matrix representation of (B2), we obtain

$$\rho_{\mathrm{iso}} - |\varphi\rangle\langle\varphi| = \begin{pmatrix} \frac{p}{d^2}+1-p-\alpha & -\sqrt{\alpha(1-\alpha)} & 0 & \cdots & 0 \\ -\sqrt{\alpha(1-\alpha)} & \frac{p}{d^2}-1+\alpha & 0 & \cdots & 0 \\ 0 & 0 & \frac{p}{d^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \frac{p}{d^2} \end{pmatrix}.$$

The eigenvaules of this matrix are $\{\frac{p}{d^2} - \frac{p}{2} \pm \frac{1}{2}\sqrt{(2-p)^2 - 4\alpha(1-p)}, \frac{p}{d^2}, \ldots, \frac{p}{d^2}\}$. The radicand in the first two eigenvalues is positive for $p \in (0,1)$ and

$\alpha \in [0,\frac{1}{d}]$. The first and second eigenvalue are positive and negative, respectively, for all $p \in (0,1)$, $\alpha \in [0,\frac{1}{d}]$, and $d \geqslant 2$. Hence,

$$G^{\mathrm{t}}(\rho_{\mathrm{iso}}) = \min_{|\varphi\rangle\in\mathrm{PRO}} \frac{1}{4}\left[\mathrm{Tr}\left|p\frac{\mathbb{1}}{d^2} + (1-p)|\Psi\rangle\langle\Psi| - |\varphi\rangle\langle\varphi|\right|\right]^2 = \min_{\alpha\in[0,\frac{1}{d}]} \frac{1}{4}\left[\sqrt{(2-p)^2 - 4\alpha(1-p)} + (d^2-2)\frac{p}{d^2}\right]^2,$$

and the minimum is obviously reached for $\alpha = \frac{1}{d}$. We can now rewrite (B1) as

$$\frac{1}{4}\left[\sqrt{(2-p)^2 - \frac{4}{d}(1-p)} + (d^2-2)\frac{p}{d^2}\right]^2 < p\left(1-\frac{1}{d^2}\right)^2 + (1-p)\left(1-\frac{1}{d}\right),$$

and it can be easily verified numerically that this inequality is satisfied for all $p \in (0,1)$ and all $d \geqslant 2$. Hence, G^{t} is not concave.

APPENDIX C: AUXILIARY RESULTS FOR CALCULATION OF G_1^{c} FOR MAXIMALLY CORRELATED STATES

Lemma 30. Let $k \in \mathbb{N}$, $n > 0$, $X > 0$, and $Y \geqslant 0$ be constants, and let $x_1,\ldots,x_k > 0$, $y_1,\ldots,y_k \geqslant 0$ be variables with the restrictions $\sum_{i=1}^k x_i = X$ and $\sum_{i=1}^k y_i = Y$. Then

$$\min_{\substack{\{x_1,\ldots,x_k\}\\ \{y_1,\ldots,y_k\}}} \sum_{i=1}^k (x_i+y_i)\log\left[n\left(1+\frac{y_i}{x_i}\right)\right] = (X+Y)\log\left[n\left(1+\frac{Y}{X}\right)\right]. \tag{C1}$$

Proof. Because of $\log[n(1+\frac{y_i}{x_i})] = \log n + \log(1+\frac{y_i}{x_i})$, it suffices to consider $n=1$. Equation (C1) clearly holds for $k=1$, so we assume $k \geqslant 2$. First, we consider the x_i to be, fixed with their values denoted as $\{x'_1,\ldots,x'_k\}$. Then the function

$$f(y_1,\ldots,y_k) := \sum_{i=1}^k (x'_i+y_i)\log\left(1+\frac{y_i}{x'_i}\right) \tag{C2}$$

is a function of k variables. We use this function to define the k functions of $k-1$ variables,

$$g_j(y_1,\ldots,y_{j-1},y_{j+1},\ldots,y_k) := f(y_1,\ldots,y_{j-1},Y_j,y_{j+1},\ldots,y_k), \tag{C3}$$

with $j \in [1,k]$, and where Y_j is shorthand for $Y_j := Y - \sum_{i\neq j} y_i$. Obviously, if (C2) is minimized at $(y'_1,\ldots,y'_k)$ under the condition $\sum_{i=1}^k y_i = Y$, then (C3) is minimized at $(y'_1,\ldots,y'_{j-1},y'_{j+1},\ldots,y'_k)$. In particular, all partial derivatives must vanish at this point, and from (C2) and (C3) it follows that

$$\frac{\partial g_j}{\partial y_l} = \log\frac{x'_j x'_l + x'_j y_l}{x'_j x'_l + Y_j x'_l}, \quad \text{for all } j,l \in [1,k] \quad \text{with} \quad j \neq l. \tag{C4}$$

To rule out boundary points as solutions, we calculate the second derivatives,

$$\frac{\partial^2 g_j}{\partial y_l^2} = \frac{x'_l + y_l + x'_j + Y_j}{(x'_l+y_l)(x'_j+Y_j)} > 0,$$

and find that they are strictly positive. From this it follows that (C4) can be zero at only one point and that the unique minimum is reached there. Since $x'_j y_l = Y_j x'_l$ results in $\frac{\partial g_j}{\partial y_l} = 0$, we obtain $x'_j y'_l = y'_j x'_l$ for all $j,l \in [1,k]$ with $j \neq l$. From this it follows that $x'_j(\sum_i y'_i) = y'_j(\sum_i x'_i)$, and thus $\frac{y'_j}{x'_j} = \frac{Y}{X}$ for all $j \in [1,k]$. Inserting this into the left-hand side of (C1) yields the right-hand side of (C1), regardless of the initial choice of the fixed $\{x'_1,\ldots,x'_k\}$. This completes the proof. ■

Lemma 31. For constants $m,n \in \mathbb{N}$ with $\frac{m}{n} \leqslant 1$ and $q \in (0,1)$, the minimum of the function

$$f(h,s) = \frac{q - s(1-q)}{1-hs}(1+h)\log[m(1+h)] + \frac{(1-q)-hq}{1-hs}(1+s)\log[n(1+s)]$$

with the domain given by $h \in [0,\frac{1-q}{q}]$ and $s \in [0,\frac{q}{1-q}]$, excluding the point $(\frac{1-q}{q},\frac{q}{1-q})$, is

$$\min f(h,s) = \begin{cases} q\log m + (1-q)\log n & \text{for } \frac{m}{n} \geqslant \frac{1}{e} \\ \log\left(\frac{m}{q}\right) & \text{for } \frac{m}{n} < \frac{1}{e} \text{ and } q \geqslant \frac{em}{n}. \\ \log n - q\frac{n\log e}{me} & \text{for } \frac{m}{n} < \frac{1}{e} \text{ and } q < \frac{em}{n} \end{cases}$$

Proof. The function $f(h,s)$ is continuous and differentiable in its entire domain, with the partial derivatives

$$\frac{\partial f(h,s)}{\partial h} = \frac{q-(1-q)s}{(1-hs)^2\ln 2}\left[1 - hs + (1+s)\ln\left(\frac{m(1+h)}{n(1+s)}\right)\right],$$

$$\frac{\partial f(h,s)}{\partial s} = \frac{(1-q)-hq}{(1-hs)^2\ln 2}\left[1 - hs + (1+h)\ln\left(\frac{n(1+s)}{m(1+h)}\right)\right].$$

For $f(h,s)$ to attain a minimum at an interior point, both partial derivatives must vanish, i.e.,

$$1 - hs + (1+s)\ln\left[\frac{m(1+h)}{n(1+s)}\right] = 1 - hs + (1+h)\ln\left[\frac{n(1+s)}{m(1+h)}\right] = 0,$$

from which it follows that $\frac{1+s}{1+h} = -1$. Since this cannot be true for any h and s, the minimum of $f(h,s)$ must be attained at a boundary point. The boundaries of $f(h,s)$ are

$$\begin{aligned} f\left(h,\frac{q}{1-q}\right) &= \log\left(\frac{n}{1-q}\right) && \text{for}\quad h\in\left[0,\frac{1-q}{q}\right)\\ f\left(\frac{1-q}{q},s\right) &= \log\left(\frac{m}{q}\right) && \text{for}\quad s\in\left[0,\frac{q}{1-q}\right)\\ f_h(h) := f(h,0) &= \log n + q(1+h)\log\left[\frac{m(1+h)}{n}\right] && \text{for}\quad h\in\left[0,\frac{1-q}{q}\right]\\ f_s(s) := f(0,s) &= \log m + (1-q)(1+s)\log\left[\frac{n(1+s)}{m}\right] && \text{for}\quad s\in\left[0,\frac{q}{1-q}\right]. \end{aligned}$$

Because f is constant on the first two boundaries, and because $f_h(\frac{1-q}{q}) = \log(\frac{m}{q})$ and $f_s(\frac{q}{1-q}) = \log(\frac{n}{1-q})$, it suffices to find the minimum of f_h and f_s. From $\frac{n}{m} \geqslant 1$ it is clear that $f_s(s)$ is monotonically increasing in s, hence $\min_s f_s(s) = f_s(0) = f(0,0) = q\log m + (1-q)\log n \geqslant \min_h f(h,0) = \min_h f_h(h)$. Therefore, it suffices to find the minimum of f_h. The derivative of f_h is

$$f_h'(h) = \frac{\partial f(h,0)}{\partial h} = q\log\frac{em(1+h)}{n}.$$

For $\frac{m}{n} \geqslant \frac{1}{e}$, we have $f_h'(h) \geqslant 0$, hence f_h attains its minimum at $f_h(0) = f(0,0)$. For $\frac{m}{n} < \frac{1}{e}$ and $q \geqslant \frac{em}{n}$, we have $f_h'(h) \leqslant 0$, hence the minimum is $f_h(\frac{1-q}{q}) = \log(\frac{m}{q})$. On the other hand, for $\frac{m}{n} < \frac{1}{e}$ and $q < \frac{em}{n}$, then $f_h'(\frac{n}{em} - 1) = 0$, so the minimum of f_h is $f_h(\frac{n}{em} - 1) = \log n - q\frac{n\log e}{me}$. ■

[1] M. B. Plenio and S. Virmani, Quant. Inf. Comp. **7**, 1 (2007).
[2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[3] A. Shimony, Ann. N. Y. Acad. Sci. **755**, 675 (1995).
[4] D. C. Brody and L. P. Hughston, J. Geom. Phys. **38**, 19 (2001).
[5] H. Barnum and N. Linden, J. Phys. A: Math. Gen. **34**, 6787 (2001).
[6] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).
[7] D. Shapira, Y. Shimoni, and O. Biham, Phys. Rev. A **73**, 044301 (2006).
[8] H. Zhu, L. Chen, and M. Hayashi, New J. Phys. **12**, 083002 (2010).
[9] A. Streltsov, H. Kampermann, and D. Bruß, Phys. Rev. A **84**, 022323 (2011).
[10] D. Markham, A. Miyake, and S. Virmani, New J. Phys. **9**, 194 (2007).
[11] T.-C. Wei, M. Ericsson, P. M. Goldbart, and W. J. Munro, Quant. Inf. Comp. **4**, 252 (2004).
[12] M. Hayashi, D. Markham, M. Murao, M. Owari, and S. Virmani, Phys. Rev. Lett. **96**, 040501 (2006).
[13] A. Streltsov, H. Kampermann, and D. Bruß, New J. Phys. **12**, 123004 (2010).
[14] T.-C. Wei, Phys. Rev. A **78**, 012327 (2008).
[15] T.-C. Wei, J. B. Altepeter, P. M. Goldbart, and W. J. Munro, Phys. Rev. A **70**, 022322 (2004).
[16] L. Chen, A. Xu, and H. Zhu, Phys. Rev. A **82**, 032301 (2010).
[17] M. Aulbach, D. Markham, and M. Murao, New J. Phys. **12**, 073025 (2010).
[18] M. Aulbach, Int. J. Quantum Inf. **10**, 1230004 (2012).
[19] J. Martin, O. Giraud, P. A. Braun, D. Braun, and T. Bastin, Phys. Rev. A **81**, 062347 (2010).
[20] O. Biham, M. A. Nielsen, and T. J. Osborne, Phys. Rev. A **65**, 062312 (2002).
[21] Y. Shimoni, D. Shapira, and O. Biham, Phys. Rev. A **69**, 062303 (2004).
[22] R. F. Werner and A. S. Holevo, J. Math. Phys. **43**, 4353 (2002).
[23] D. Gross, S. T. Flammia, and J. Eisert, Phys. Rev. Lett. **102**, 190501 (2009).
[24] M. Van den Nest, W. Dür, A. Miyake, and H. J. Briegel, New J. Phys. **9**, 204 (2007).
[25] C. E. Mora, M. Piani, A. Miyake, M. Van den Nest, W. Dür, and H. J. Briegel, Phys. Rev. A **81**, 042315 (2010).
[26] M. Hayashi, D. Markham, M. Murao, M. Owari, and S. Virmani, Phys. Rev. A **77**, 012104 (2008).
[27] H. A. Carteret, A. Higuchi, and A. Sudbery, J. Math. Phys. **41**, 7932 (2000).
[28] R. Orús, S. Dusuel, and J. Vidal, Phys. Rev. Lett. **101**, 025701 (2008).
[29] R. Orús, Phys. Rev. Lett. **100**, 130502 (2008).
[30] R. Orús and T.-C. Wei, Phys. Rev. B **82**, 155120 (2010).
[31] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[32] I. Bengtsson and K. Ż yczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
[33] G. Vidal, J. Mod. Opt. **47**, 355 (2000).
[34] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).

[35] M. B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005).

[36] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).

[37] M. B. Plenio and V. Vedral, J. Phys. A: Math. Gen. **34**, 6997 (2001).

[38] E. M. Rains, Phys. Rev. A **60**, 179 (1999).

[39] E. M. Rains, Phys. Rev. A **63**, 019902(E) (2000).

[40] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor, Phys. Rev. Lett. **87**, 217902 (2001).

[41] K. G. H. Vollbrecht and R. F. Werner, Phys. Rev. A **64**, 062307 (2001).

[42] E. M. Rains, IEEE Trans. Inform. Theory **47**, 2921 (2001).

[43] M. Hajdušek and M. Murao, New J. Phys. **15**, 013039 (2013).

[44] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).

[45] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[46] Y. Cao and A. M. Wang, J. Phys. A: Math. Theor. **40**, 3507 (2007).

[47] M. Horodecki, Open Syst. Inform. Dynam. **12**, 231 (2005).

[48] M. Ozawa, Phys. Lett. A **268**, 158 (2000).

[49] R. Hübener, M. Kleinmann, T.-C. Wei, C. González-Guillén, and O. Gühne, Phys. Rev. A **80**, 032324 (2009).

[50] E. Jung, M.-R. Hwang, H. Kim, M.-S. Kim, D. K. Park, J.-W. Son, and S. Tamaryan, Phys. Rev. A **77**, 062317 (2008).

[51] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).

[52] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[53] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).

[54] P. Hayden, K. Milner, and M. M. Wilde, in *Proceedings of the 28th IEEE Conference on Computational Complexity* (IEEE, Palo Alto, California, 2013), pp. 156–167.

[55] I. Devetak and A. Winter, Proc. R. Soc. A **461**, 207 (2005).

[56] Barbara M. Terhal and Karl Gerd H. Vollbrecht, Phys. Rev. Lett. **85**, 2625 (2000).

[57] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).

[58] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).

[59] M. Hajdušek and V. Vedral, New J. Phys. **12**, 053015 (2010).