The main concern of this paper is how to define proper measures of multipartite entanglement for mixed quantum states. Since the structure of partial separability/entanglement is getting complicated if the number of subsystems exceeds two, one cannot expect the existence of an ultimate scalar entanglement measure, which grasps even a small part of the rich hierarchical structure of multipartite entanglement, and some higher order structure characterizing that is needed. In this paper we make some steps towards this direction. First, we reveal the lattice-theoretic structure of the partial separability classification, introduced earlier [Sz. Szalay and Z. Kökényesi, Phys. Rev. A 86, 032341 (2012)]. It turns out that the structure of the entanglement classes is the up-set lattice of the structure of the different kinds of partial separability/entanglement, which is the down-set lattice of the lattice of the partitions of the subsystems. Second, we introduce the notion of multipartite monotonicity, expressing that a given set of entanglement monotones, while measuring the different kinds of entanglement, shows also the same hierarchical structure as the entanglement classes. Then we construct such hierarchies of entanglement measures, and propose a physically well-motivated one, being the direct multipartite generalization of the Entanglement of Formation based on the Entanglement Entropy, motivated by the notion of statistical distinguishability.

PACS numbers: 03.65.Ud, 03.67.Mn

| CONTENTS | |
| --- | --- |
| I. Introduction | 2 |
| II. Quantum states and entanglement: basics | 3 |
| A. Quantum states | 3 |
| B. Entanglement | 3 |
| III. Quantum states and entanglement for multipartite systems | 4 |
| A. Level 0: subsystems | 4 |
| B. Level I: partial separability hierarchy of the first kind | 4 |
| C. Examples | 6 |
| D. Level II: partial separability hierarchy of the second kind | 6 |
| E. Examples | 7 |
| F. Level III: partial separability classes | 8 |
| G. Examples | 9 |
| IV. Entanglement measures: basics | 9 |
| A. Mixedness of states | 9 |
| B. Distinguishability of states | 10 |
| C. LOCC monotonicity: entanglement measures | 11 |
| D. Discrimination: indicator functions | 11 |
| E. Local entropies: pure state measures | 12 |
| F. Convex roof extensions: mixed state measures | 12 |
| G. Examples | 12 |
| V. Sums and Means: a detour | 13 |
| A. The meaning of sums and means | 13 |
| B. Definitions and properties of power-sums and power-means | 14 |
| C. Definitions and properties of quasi-sums and quasi-arithmetic means | 15 |
| D. Means of measures | 15 |
| VI. Entanglement measures for multipartite systems | 15 |
| A. Level 0: bipartite entanglement | 16 |
| B. Level I: multipartite entanglement measures of the first kind | 16 |
| C. Examples | 17 |
| D. Level II: multipartite entanglement measures of the second kind | 18 |
| E. Examples | 20 |
| F. Entanglement measures for mixed states | 20 |
| G. Examples | 21 |
| H. Level III: detection of the classes | 21 |
| I. Examples | 21 |
| VII. Summary | 21 |

Acknowledgments | 22 |

A. On the lattice structure of the classification | 22 |
1. Posets and lattices: basics | 22 |
2. Construction by bipartitions | 23 |
3. Order isomorphisms of the first kind | 23 |
4. Meet-semilattice isomorphism of the first kind for pure states | 23 |
5. Order isomorphisms of the second kind | 24 |
I. INTRODUCTION

In the description of entanglement [1, 2], it is a hard problem unsolved yet how to step from the bipartite scenario to the multipartite one, and how to define proper measures of multipartite entanglement. The state of a bipartite quantum system can be either separable or entangled, while for more-than-two-partite systems the partial separability properties have a complicated structure [3], and a system of measures fitting to this structure, while being physically motivated, is not known.

The quantum entanglement in bipartite pure states can be described completely by the use of the singular value decomposition (SVD), mostly called Schmidt decomposition [4]. This leads to a local unitary canonical form, what allows of the separation of the nonlocal parameters of the state (relevant for the description of correlations) from the local (irrelevant) ones. The Schmidt coefficients contain then all nonlocal properties of the state, they have a simple structure, (which means that the pure bipartite entanglement itself has also a simple structure,) and, in principle, every measure of entanglement can be expressed by them. A well-known example is the Entanglement Entropy [5]. One can step from pure state entanglement measures to mixed ones by the use of convex roof extension [5], a well-known example is the Entanglement of Formation [6], which is the extension of the Entanglement Entropy.

For the case of multipartite pure states, local unitary canonical form is not known in general [7]. A higher order singular value decomposition (HOSVD, or Vidal decomposition) [8] can be formulated, which is a sequence of bipartite singular value decompositions. Although it is very important in numerical techniques of the quantum mechanics of strongly correlated systems [9], it does not give us so deep insight into the structure and quantitative description of multipartite entanglement as the SVD did in the bipartite case.

A very different approach is to build up the partial separability / entanglement structure from the grounds [3], and to define different entanglement measures for the different kinds of partial separability. The basic observation making this possible is that the whole construction can be formulated by the use of the notion of pure state entanglement with respect to a bipartite split, what is relatively well-understood. In the present paper, we carry out this program.

The first part of the paper is devoted to the classification of multipartite entanglement. After recalling the basic notions in the entanglement of bipartite systems in Section II, we build up those for the multipartite case in Section III, in a more clarified treatment than was done originally [3], making possible to achieve some new developments. Our results can be formulated naturally in the language of Lattice Theory [10]. We work out the hierarchical structure of different kinds of partial separability, which turns out to be the down-set lattice of the lattice of the partitions of the subsystems (see Section IIIID), and also the structure of the entanglement classes, which turns out to be also hierarchical, being the up-set lattice of the lattice above (see Section IIIIF).

The second part of the paper is devoted to the quantification of multipartite entanglement. After recalling the basic notions in the measures of bipartite entanglement in Section IV, we consider the power-sums and power-means (together with their generalizations) as useful tools for the construction of entanglement measures from entanglement measures in Section V, then we construct measures for the multipartite case in Section VI. Besides the usual entanglement monotonicity and discrimination, we introduce the multipartite monotonicity, which endows the set of multipartite entanglement measures with the same hierarchical structure as the partial separability has. We succeed in constructing a hierarchy of entanglement measures satisfying these requirements in Section VID, which are the direct generalizations of the Entanglement Entropy for pure states and the Entanglement of Formation for mixed states. These measures have information-geometrical meaning, related to the statistical distinguishability.

The summary is left to Section VII. The proofs of propositions in the main text are given in the Appendix. In the entanglement theory of mixed states, the central notion is the convexity: we deal mostly with convex or concave functions defined on convex sets. Apart from the results in the body of the text, this paper is intended to be a self-contained discourse on convexity [11, 12] and entanglement [1, 2, 11, 13–17]. So we also recall some known proofs of theorems about entanglement measures, and also some useful calculations about convexity to enlighten how this structure shows up [11, 12].
II. QUANTUM STATES AND ENTANGLEMENT: BASICS

Here we briefly recall the basic notions arising in the description of the states of singlepartite (Section II A) and bipartite (Section II B) quantum systems [2, 11, 13–18]. We fix some basic notational conventions for state vectors, pure and mixed states, separable and entangled states.

A. Quantum states

Entanglement theory deals with the states of quantum systems. A state vector is an element of a Hilbert space, $|\psi\rangle \in \mathcal{H}$, which is normalized with respect to the standard 2-norm of the Hilbert space, $\|\psi\| = \sqrt{\langle \psi | \psi \rangle} = 1$. In the paper, we consider the finite dimensional case only. The pure state of a quantum system is represented by a one-dimensional subspace (ray) in the Hilbert space (which is actually an element of the projective Hilbert space) which can be given by a state vector as the self-adjoint linear operator $\pi = |\psi\rangle \langle \psi|$, being the projector projecting to that one-dimensional subspace spanned by $|\psi\rangle$. (Note that the projectors are characterized by $\pi^2 = \pi = \pi^\dagger$. For projectors, having unit trace is equivalent to being of rank one.) The set of pure states over the Hilbert space $\mathcal{H}$ arises as

$$\mathcal{P}(\mathcal{H}) = \{ \pi \in \text{Lin}_{\mathbb{C}} \mathcal{H} \mid \pi^2 = \pi, \|\pi\|_{tr} = \text{tr} \pi = 1 \}. \tag{1a}$$

(If there is no ambiguity about the underlying Hilbert space, we use the notation $\mathcal{P} := \mathcal{P}(\mathcal{H})$.) A mixed state is represented by the convex combination (or mixture) of pure states, and it represents the state of an ensemble of quantum systems $\{|\psi_i\rangle \rangle 1, i = 1, \ldots, m\}$, described by the pure state $\pi_i$ with probability $p_i$. The convex body of mixed states over the Hilbert space $\mathcal{H}$ arises as

$$\mathcal{D}(\mathcal{H}) = \text{Conv} \mathcal{P}(\mathcal{H}) \equiv \left\{ \varrho \in \text{Lin}_{\mathbb{C}} \mathcal{H} \mid \exists \pi_i \in \mathcal{P}, p_i \geq 0, \sum_i p_i = 1 : \varrho = \sum_i p_i \pi_i \right\}. \tag{1b}$$

(If there is no ambiguity about the underlying Hilbert space, we use the notation $\mathcal{D} := \mathcal{D}(\mathcal{H})$.) This turns out to be equivalent to the positive semidefinite operators normalized with respect to the trace-norm,

$$\mathcal{D}(\mathcal{H}) \equiv \left\{ \varrho \in \text{Lin}_{\mathbb{C}} \mathcal{H} \mid \varrho \geq 0, \|\varrho\|_{tr} = \text{tr} \varrho = 1 \right\}. \tag{1c}$$

Geometrically, the pure states are the extremal points of the convex body of the mixed states [11],

$$\mathcal{P} = \text{Extr} \mathcal{D}, \tag{1d}$$

a pure state can not be mixed out from any states nontrivially.

Convexity is a central notion in quantum (and also in classical) probability theory [11]. State spaces are, in general, convex sets, which means that they are closed under convex combination, called also mixing. That is, for convex combination coefficients $0 \leq p_i \in \mathbb{R}$ with $\sum_i p_i = 1$, called also mixing weights, if $\varrho_i$ are states, then their convex combination $\sum_i p_i \varrho_i$ is also a state. Mixing is interpreted as forgetting some classical information about the state by which the system is described, so this is indeed a necessary property. The main difference between classical and quantum probability theory is that in the quantum case, because of the superposition principle, (linear structure in the Hilbert space), the pure state decomposition of a nonpure state is not unique, contrary to the classical case.

B. Entanglement

Entanglement theory, on the other hand, deals with the states of composite quantum systems. For example, for two subsystems, with state vectors being the normalized elements of the associated Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, the state vectors are the normalized elements of the tensor product Hilbert space, $\mathcal{H}_{12} := \mathcal{H}_1 \otimes \mathcal{H}_2$. We have again the pure states $\mathcal{P}_1 := \mathcal{P}(\mathcal{H}_1)$, $\mathcal{P}_2 := \mathcal{P}(\mathcal{H}_2)$ and $\mathcal{P}_{12} := \mathcal{P}(\mathcal{H}_{12})$, being the projectors onto one-dimensional subspaces in $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_{12}$, and the mixed states $\mathcal{D}_1 := \text{Conv} \mathcal{P}_1$, $\mathcal{D}_2 := \text{Conv} \mathcal{P}_2$ and $\mathcal{D}_{12} := \text{Conv} \mathcal{P}_{12}$, being the mixtures of pure states, for subsystem 1, 2 and the whole system 12. One can obtain the reduced (or marginal) states by the use of the partial trace operation, for example $\text{tr}_2 : \mathcal{D}_{12} \rightarrow \mathcal{D}_1$, which is linear, and $\text{tr}_2(X \otimes Y) = X(\text{tr} Y)$.

If the state vector $|\psi\rangle \in \mathcal{H}_{12}$ can be written as an elementary tensor $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ with suitable state vectors $|\psi_1\rangle \in \mathcal{H}_1$ and $|\psi_2\rangle \in \mathcal{H}_2$, then it is separable, else it is entangled, e.g., $|\psi\rangle = \frac{1}{\sqrt{2}} (|\psi_1\rangle \otimes |\psi_2\rangle + |\psi_1'\rangle \otimes |\psi_2'\rangle)$ with $\langle \psi_1 | \psi_1'\rangle = \langle \psi_2 | \psi_2'\rangle = 0$. The set of separable pure states is then

$$\mathcal{P}_{\text{sep}} = \left\{ \pi \in \mathcal{P}_{12} \mid \exists \pi_1 \in \mathcal{P}_1, \exists \pi_2 \in \mathcal{P}_2 : \pi = \pi_1 \otimes \pi_2 \right\}, \tag{2a}$$

that is, the rank-one projectors with separable images, the set of entangled pure states is its complement $\mathcal{P} \setminus \mathcal{P}_{\text{sep}}$, and the set of separable mixed states is

$$\mathcal{D}_{\text{sep}} = \text{Conv} \mathcal{P}_{\text{sep}}, \tag{2b}$$

the set of entangled mixed states is its complement $\mathcal{D} \setminus \mathcal{D}_{\text{sep}}$. This definition is motivated by that the separable mixed states can be created from uncorrelated (product) states by the use of local (quantum) operations and classical communication (LOCC) [19, 20], while entangled states cannot be. That is, the set of separable mixed states are closed under LOCC. Geometrically, the separable pure states are the extremal points of the convex body of the separable mixed states,

$$\mathcal{P}_{\text{sep}} = \text{Extr} \mathcal{D}_{\text{sep}}. \tag{2c}$$
The situation is summarized as
\begin{equation}
D_{\text{sep}} \subseteq D_{12} \\
\cup \quad \cup \\
P_{\text{sep}} \subseteq P_{12},
\end{equation}
what represents an important point of view in the sequel.
Thanks to the Schmidt decompositions for bipartite state vectors [4, 13], it is easy to decide whether a pure state is separable or not: for all \( \pi \in P_{12} 

\pi \in P_{\text{sep}} \iff \text{tr}_2 \pi \in P_1 \iff \text{tr}_1 \pi \in P_2.
\end{equation}

The mixed separability problem is, however, a hard optimization task [1, 2, 21].

From the classical point of view, one faces several counterintuitive consequences following from the existence of entangled states. These are based more-or-less on the fact that, as can be seen from (3), entangled pure states have mixed marginals, what is completely unimaginable for the classically thinking mind [22–24], since in the classical case the marginals of a pure joint probability distribution are pure ones.

III. QUANTUM STATES AND ENTANGLEMENT FOR MULTIPARTITE SYSTEMS

In this section, we reconstruct the Partial Separability Classification of multipartite mixed states in a more clarified way than was done originally [2, 3]. This classification is complete in the sense of partial separability, that is, it utilizes all the possible combinations of different kinds of partially separable pure states. We also reveal the lattice theoretic structure behind the class structure. For a quick summary on the basic elements of Lattice Theory we use, see Appendix A 1, based on [10].

The basic observation, on which the construction is built up, is that the whole construction can be formulated by the use of the notion of pure state entanglement with respect to a bipartite split. During the construction, we separate the abstract hierarchy of the labelling the partial separability properties from the concrete hierarchy of the state sets of pure and mixed states, what results in a very transparent building. This building is of four floors. The ground floor is the hierarchy of subsystems (Section III A), then the first and second floors are the hierarchic structures of state sets of different partial separability properties (Sections III B and III D), and the third floor is the hierarchic structure of classes of states showing different entanglement properties (Section III F).

A. Level 0: subsystems

First of all, let us introduce some convenient notations. For \( n \)-partite systems \( (n > 0) \), the set of the labels of the elementary subsystems is \( L = \{1, 2, \ldots, n\} \).

That is, for all \( a \in L \), we have a Hilbert space \( \mathcal{H}_a \), with \( \dim \mathcal{H}_a < \infty \), associated to the elementary subsystem of label \( a \). A subsystem (not elementary in general) is then labelled by a nonempty \( K \subseteq L \), and has the Hilbert space \( \mathcal{H}_K = \bigotimes_{a \in K} \mathcal{H}_a \) associated to it. For labelling the complementary subsystem, we have the notation \( \overline{K} = L \setminus K \).

We have the shorthand notation \( \mathcal{P} = \mathcal{P}_L \) and \( \mathcal{D} = \mathcal{D}_L \) for the pure and mixed states of the whole system, respectively.

In a formal sense, the label of a subsystem is a nonempty element of the power-set
\begin{equation}
P_0 := 2^L
\end{equation}
of the labels of the elementary subsystems \( L \), so we have the power-set lattice of subsystems [10],
\begin{equation}
(P_0, \subseteq, \cup, \cap, \setminus, \emptyset, L).
\end{equation}
The size of that is \( |P_0| = 2^{|L|} = 2^n \).

B. Level I: partial separability hierarchy of the first kind

We would like to form mixtures from a given kind of partially separable pure states. To this end, let \( \alpha = \{K_1, K_2, \ldots, K_{|\alpha|}\} = K_1|K_2|\ldots|K_{|\alpha|} \) be a splitting of the system, that is a partition of the labels \( L \) into disjoint nonempty sets \( K_i \subseteq L \), which together amount to \( L \). We have the set of all the possible partitions
\begin{equation}
P_1 = \left\{ \alpha = K_1|K_2|\ldots|K_k \mid \forall K \in \alpha : K \in P_0 \setminus \emptyset, \forall K, K' \in \alpha : K \neq K' \Rightarrow K \cap K' = \emptyset, \bigcup_{K \in \alpha} K = L \right\}.
\end{equation}

We call the partitions labels of the first kind, and we use them for the labelling of such states. The number of them for all \( n \) is given by the \( |P_1| = B_n \) Bell numbers [25], given by the recursive formula \( B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k \), with \( B_0 = B_1 = 1 \).

There is a natural (partial) order on the set of the partitions. For two partitions \( \beta, \alpha \in P_1 \), \( \beta \) is a refinement of \( \alpha \), ("\( \beta \) is finer than \( \alpha \), or, "\( \alpha \) is coarser than \( \beta \)"")
denoted as \( \beta \preceq \alpha \), if \( \alpha \) can be obtained from \( \beta \) by joining some (maybe none) of the parts of \( \beta \), that is,

\[
\beta \preceq \alpha \quad \text{def.} \quad \forall K' \in \beta, \exists K \in \alpha : K' \subseteq K. \tag{8}
\]

This defines a partial order on the set of partitions \([10]\), and \((P_1, \preceq)\) is a poset (partially ordered set). For example, for the tripartite case \(1|2|3 \leq 1|23 \leq 123\). Moreover, there is a top and a bottom element, which are the full \( n \)-partite split and the trivial partition without split, respectively, \( \bot = 1|2\ldots|n \preceq \alpha \preceq \top = 12\ldots n \). (In the writing, we omit the braces and also the comma, if this does not cause confusion.)

For the poset of partitions, one can define the greatest lower bound or meet, \( \alpha \land \beta \), and the least upper bound or join, \( \alpha \lor \beta \) as

\[
\alpha \land \beta = \big\{ K \cap K' \neq \emptyset \mid K \in \alpha, K' \in \beta \big\}, \tag{9a}
\]

\[
\alpha \lor \beta = \bigwedge \{\alpha, \beta \}, \tag{9b}
\]

so the set of partitions forms a lattice,

\[
(P_1, \preceq, \lor, \land, 1|2\ldots|n, 12\ldots n). \tag{10}
\]

(For a summary on the basic constructions in order theory, see Appendix A 1.)

It is important that the bipartitions \( K|\overline{K} \in P_1 \) can be used for the generation of all partitions,

\[
\alpha \supseteq \bigwedge_{K \in \alpha} K|\overline{K}. \tag{11}
\]

(For the proof, see Appendix A 2.)

This turns out to be crucial later, when the multipartite entanglement measures are built upon bipartite ones.

For a partition \( \alpha \in P_1 \), we have the set of \( \alpha \)-separable pure states

\[
P_\alpha := \big\{ \pi \in \text{Lin } H \mid \forall K \in \alpha, \exists \pi_K \in P_K : \pi = \bigotimes_{K \in \alpha} \pi_K \big\}, \tag{12a}
\]

and the set of \( \alpha \)-separable mixed states

\[
D_\alpha := \text{Conv } P_\alpha, \tag{12b}
\]

(that is, \( \varrho \) is \( \alpha \)-separable if and only if it can be mixed by the use of \( \alpha \)-separable pure states). It also holds by construction that

\[
P_\alpha := \text{Extr } D_\alpha, \tag{12c}
\]

there are no other extremal \( \alpha \)-separable states than the pure ones. For the 1-partite trivial split \( \alpha = \{K_1\} = \{L\} \), we have that the \( \{L\}\)-separable pure and mixed states \( P_{\{L\}} = P_L = P \) and \( D_{\{L\}} = D_L = D \) are obviously all the pure and mixed states of the system. Note that for all \( \alpha \in P_1 \), the state sets \( D_\alpha \) are closed under LOCC.

Note that these definitions only demand the separability with respect to a given split, independently whether the separability with respect to a finer split also holds. That is, the \( P_\alpha \) and \( D_\alpha \) sets are containing (and also closed), and the sets

\[
P_{1,P} = \{ P_\alpha \mid \alpha \in P_1 \}, \quad \tag{13a}
\]

\[
P_{1,D} = \{ D_\alpha \mid \alpha \in P_1 \} \quad \tag{13b}
\]

are posets with respect to the inclusion, \((P_{1,P}, \subseteq)\), \((P_{1,D}, \subseteq)\). Moreover, the set-theoretical inclusion is perfectly resembles the ordering of the respective partitions,

\[
\beta \preceq \alpha \iff P_\beta \subseteq P_\alpha, \tag{14a}
\]

and

\[
\beta \preceq \alpha \iff D_\beta \subseteq D_\alpha, \tag{14b}
\]

(that is, separability with respect to a finer split implies that with respect to a coarser one) so the posets \((P_1, \preceq)\), \((P_{1,P}, \subseteq)\) and \((P_{1,D}, \subseteq)\) are isomorphic. (We recall the proof in Appendix A 3 from [2, 3] in a slightly modified form adjusted to the present construction.)

Do other structures meet and join \( \land, \lor \) resemble the set-theoretical intersection and union \( \cap, \cup \) for the state sets \((13a)\) and \((13b)\)? We know from \((8)\) that \( \alpha \land \beta \preceq \alpha, \beta \preceq \alpha \lor \beta \), this leads to \( P_{\alpha \land \beta} \subseteq P_\alpha, P_\beta \subseteq P_{\alpha \lor \beta} \)

and \( D_{\alpha \land \beta} \subseteq D_\alpha, D_\beta \subseteq D_{\alpha \lor \beta} \) due to \((14a)\) and \((14b)\).

From these we have

\[
P_{\alpha \land \beta} \subseteq P_{\alpha} \cap P_{\beta}, \quad P_{\alpha} \cup P_{\beta} \subseteq P_{\alpha \lor \beta}, \tag{15a}
\]

\[
D_{\alpha \land \beta} \subseteq D_{\alpha} \cap D_{\beta}, \quad D_{\alpha} \cup D_{\beta} \subseteq D_{\alpha \lor \beta}. \tag{15b}
\]

These are what we have by the use of only the \((14a)\)-(14b) isomorphisms of the orderings. However, there is more to be known for pure states if one takes into consideration the \((12a)\) definition of the \( P_\alpha \) sets of \( \alpha \)-separable pure states. In this case it can be proven that

\[
P_{\alpha} \cap P_{\alpha'} = P_{\alpha \land \alpha'}, \tag{16}
\]

that is, a pure state is separable under the splits \( \alpha \) and \( \alpha' \) if and only if it is separable under their meet \( \alpha \land \alpha' \) \((9a)\).

(For the proof, see Appendix A 4.)

This means that \( P_{1,P} \) is closed under intersection, and

\[
(P_{1,P}, \subseteq, \cap, P_{1|2\ldots|n}, P_{12\ldots n}) \tag{17}
\]

is a meet-semilattice, and due to \((14a)\) and \((16)\), this structure is isomorphic to that of \( P_1 \) given in \((10)\),

\[
(P_{1,P}, \subseteq, \cap, P_{1|2\ldots|n}, P_{12\ldots n}) \equiv (P_1, \preceq, \land, 1|2\ldots|n, P_{12\ldots n}). \tag{18}
\]

A corollary of \((16)\) and \((11)\) is that

\[
P_\alpha = \bigcap_{K \in \alpha} P_{K|\overline{K}}, \tag{19}
\]

that is, a pure state is separable under a split \( \alpha \) if and only if it is separable under all bipartitions \( K|\overline{K} \), where
states, we have only the poset $K$ less here. The lattice for the inclusion hierarchy of the sets of as was mentioned before.

and, due to (14b), this structure is isomorphic to that of $P_1$ given in (10),

\[ (P_1, \subseteq, D_1, \ldots, n, D_{12} \ldots n) \equiv (P_1, \subseteq, 1, 2 \ldots n, 12 \ldots n), \]

as was mentioned before.

C. Examples

Writing out some examples explicitly might not be useless here. The lattice $P_1$ for the cases $n = 2$ and $3$ can be seen in the upper-left parts of Figures 1 and 1. As we have learned in (14a) and (14b), we need to draw only this lattice for the inclusion hierarchy of the sets of $\alpha$-separable pure and mixed states $P_1 \equiv \mathcal{A}_a$ and $D_a$ (12b).

For the bipartite case we have $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2,$ and we get back the content of Section II B. The $\alpha$-separable pure states are

\[ P_{12} \equiv \mathcal{P}(\mathcal{H}_{12}), \]

\[ P_{1/2} = \{ \pi \in P_{12} \mid \pi = \pi_1 \otimes \pi_2 \} = \mathcal{P}_{\text{sep}}. \]

Note that $P_{1/2} \subseteq P_{12}.$ The $\alpha$-separable mixed states are

\[ D_{12} = \text{Conv} \ P_{12} \equiv \mathcal{D}(\mathcal{H}_{12}), \]

\[ D_{1/2} = \text{Conv} \ P_{1/2} = \mathcal{D}_{\text{sep}}. \]

Note that, again, $D_{1/2} \subseteq D_{12}.$

For the tripartite case we have $\mathcal{H}_{123} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3.$ The $\alpha$-separable pure states are

\[ P_{123} \equiv \mathcal{P}(\mathcal{H}_{123}), \]

\[ P_{a|bc} = \{ \pi \in P_{123} \mid \pi = \pi_a \otimes \pi_{bc} \} , \]

\[ P_{1/23} = \{ \pi \in P_{123} \mid \pi = \pi_1 \otimes \pi_2 \otimes \pi_3 \} , \]

with all bipartition $a|bc$ of $\{1, 2, 3\}$. Note that $P_{1/23} \subseteq P_{a|bc} \subseteq P_{123}.$ The $\alpha$-separable mixed states are

\[ D_{123} = \text{Conv} \ P_{123} \equiv \mathcal{D}(\mathcal{H}_{123}), \]

\[ D_{a|bc} = \text{Conv} \ P_{a|bc}, \]

\[ D_{1/23} = \text{Conv} \ P_{1/23}. \]

Note that, again, $D_{1/23} \subseteq D_{a|bc} \subseteq D_{123}.$

D. Level II: partial separability hierarchy of the second kind

An important observation is in the entanglement of multipartite mixed states [26] that there are mixed states which cannot be mixed by the use of any given nontrivial $\alpha$-separable pure states, while can be mixed by the use of pure states of different $\alpha$-separability. For example, for the tripartite case in Section III C, there are states $\varrho \notin D_{a|bc},$ which can be mixed by the use of bipartite entanglement in subsystems $12,$ $13$ and $23,$ that is, $\varrho \in \text{Conv}(P_{123} \cup P_{2|13} \cup P_{3|12}).$ Such states should not be considered fully tripartite-entangled, and these kinds of situations have to be handled.

So, we would also like to form mixtures from different kinds of partially separable pure states. To this end, let $\alpha$ be a down-set in $P_1,$ that is, a nonempty $\alpha = \{ \alpha_1, \alpha_2, \ldots, \alpha_{|\alpha|} \} \subseteq P_1$ set, which contains every partition which is finer than its maximal elements (see Appendix A 1). We have the set of all the possible nonempty down-sets

\[ P_1 \equiv \mathcal{O}_\downarrow(P_1) \setminus \emptyset \]

\[ = \{ \alpha \in 2^{P_1} \setminus \emptyset \mid \forall \alpha \in \alpha : \beta \subseteq \alpha \Rightarrow \beta \in \alpha \}. \]

We call the nonempty down-sets of partitions labels of the second kind, and we use them for the labelling of such states. (For a nonempty down-set $\alpha,$ the set of its maximal elements, $\max \alpha,$ was called proper label, and denoted in the same way previously [2, 3]. Since the set of maximal elements of a nonempty down-set $\alpha$ determines $\alpha = \downarrow \max \alpha$ uniquely, and vice versa, both $\max \alpha$ and $\alpha$ are equally suitable for the labelling of the sets of states with the given partial separability properties. The former one is more ink-saving, and perhaps more expressive in some sense, while the latter one leads to simpler and more transparent mathematical construction.)

The $P_{II}$ set of nonempty down-sets of the lattice $P_1$ forms a lattice with respect to the inclusion, intersection and union [10], so we have

\[ (P_{II}, \subseteq, \wedge, \vee, \{1, 2 \ldots n\}, \downarrow\{12 \ldots n\}) \]

\[ = \{ \mathcal{O}_\downarrow(P_1) \setminus \emptyset \subseteq, \subseteq, \cup, \{1, 2 \ldots n\}, \downarrow\{12 \ldots n\} \}. \]

For example, for the tripartite case $\downarrow\{123\} = \{123\} \subseteq \downarrow\{1|23\} \leq \downarrow\{1|23|2|13\} \leq \downarrow\{1|23|2|13|3|12\} \leq \downarrow\{123\} = P_1.$

For a down-set $\alpha \in P_{II},$ we have the set of $\alpha$-separable pure states

\[ \mathcal{P}_\alpha := \bigcup_{\alpha \in \alpha} \mathcal{P}_\alpha = \bigcup_{\alpha \in \max \alpha} \mathcal{P}_\alpha, \]

(28a)

(that is, $\pi$ is $\alpha$-separable if and only if it is $\alpha$-separable for at least one $\alpha \in \alpha,$ on the other hand, because of (14a), $\max \alpha$ is also suitable for the labelling) and the set of $\alpha$-separable mixed states

\[ \mathcal{D}_\alpha := \text{Conv} \mathcal{P}_\alpha, \]

(28b)

(that is, $\varrho$ is $\alpha$-separable if and only if it can be mixed by the use of the $\alpha$-separable pure states for which $\alpha \in \alpha$).

It also holds by construction that

\[ \mathcal{P}_\alpha := \text{Extr} \mathcal{D}_\alpha, \]

(28c)
so there are no other extremal $\alpha$-separable states than the pure ones. For the $\alpha$ containing only the 1-partite trivial split $\alpha = \{\alpha\} = \{L\}$, we have that the $\{L\}$-separable pure and mixed states $P_{\{L\}} = P_L \equiv P$ and $D_{\{L\}} = D_L \equiv D$ are obviously all the pure and mixed states of the system. Note that for all $\alpha \in \Pi_I$, the state sets $\mathcal{D}_\alpha$ are closed under LOCC.

Note that these definitions only demand the separability with respect to any of the given splits, independently whether the separability with respect to finer splits also holds. That is, the $\mathcal{P}_\alpha$ and $\mathcal{D}_\alpha$ sets are containing (and also closed), and the sets

\[
P_{\Pi_I} = \{P_{\alpha} \mid \alpha \in \Pi_I\}, \quad (29a)
\]

\[
P_{\Pi_D} = \{D_{\alpha} \mid \alpha \in \Pi_I\} \quad (29b)
\]

are posets with respect to the inclusion, $(P_{\Pi_I}, \subseteq)$, $(P_{\Pi_D}, \subseteq)$. Moreover, the set-theoretical inclusion is perfectly resembles the ordering (inclusion) (27) of the respective labels of the second kind,

\[
\beta \preceq \alpha \iff \mathcal{P}_\beta \subseteq \mathcal{P}_\alpha, \quad (30a)
\]

and

\[
\beta \preceq \alpha \iff \mathcal{D}_\beta \subseteq \mathcal{D}_\alpha, \quad (30b)
\]

(is, a separability lower in the hierarchy implies a higher one), so the posets $(P_{\Pi_I}, \preceq)$, $(P_{\Pi_D}, \preceq)$ and $(P_{\Pi_I}, \subseteq)$ are isomorphic. (We recall the proof in Appendix A5 from [2, 3] in a slightly modified form adjusted to the present construction.)

Do the other structures meet and join $\wedge, \vee$ (27) resemble the set-theoretical intersection and union $\cap, \cup$ for the state sets (29a) and (29b)? We know from (A8) that $\alpha \wedge \beta \preceq \alpha, \beta \preceq \alpha \vee \beta$, this leads to $P_{\alpha \wedge \beta} \subseteq P_{\alpha}, P_{\beta} \subseteq P_{\alpha \vee \beta}$ and $D_{\alpha \wedge \beta} \subseteq D_{\alpha}, D_{\beta} \subseteq D_{\alpha \vee \beta}$ due to (30a) and (30b). From these we have

\[
P_{\alpha \wedge \beta} \subseteq P_{\alpha} \cap P_{\beta}, \quad P_{\alpha} \cup P_{\beta} \subseteq P_{\alpha \vee \beta}, \quad (31a)
\]

\[
D_{\alpha \wedge \beta} \subseteq D_{\alpha} \cap D_{\beta}, \quad D_{\alpha} \cup D_{\beta} \subseteq D_{\alpha \vee \beta}. \quad (31b)
\]

These are what we have by the use of only the (30a)-(30b) isomorphisms of the orderings. However, there is to be known for pure states if one takes into consideration the (28a) definition of the $\mathcal{P}_\alpha$ sets of $\alpha$-separable pure states. In this case it can be proven that

\[
P_{\alpha} \cap P_{\alpha^\prime} = P_{\alpha \wedge \alpha^\prime}, \quad P_{\alpha} \cup P_{\alpha^\prime} = P_{\alpha \vee \alpha^\prime}. \quad (32)
\]

(For the proof, see Appendix A6.) This means that $P_{\Pi_I}$ is closed under intersection and union, and

\[
(P_{\Pi_I}, \subseteq, \cap, \cup, P_{\{1\} \cup \ldots \cup \{n\}}, P_{\{1\} \cup \ldots \cup \{n\}}) \quad (33)
\]

is a lattice, and due to (30a) and (32), this structure is isomorphic to that of $P_{\Pi}$ given in (27),

\[
(P_{\Pi_I}, \subseteq, \cap, \cup, P_{\{1\} \cup \ldots \cup \{n\}}, P_{\{1\} \cup \ldots \cup \{n\}}) \equiv (P_{\Pi}, \subseteq, \cap, \cup, \{1\} \cup \ldots \cup \{n\}, \{1\} \cup \ldots \cup \{n\}) \quad (34)
\]

On the other hand, because of the convex hull construction (28b), there is no such result for mixed states, we have only the poset

\[
(P_{\Pi_I}, \subseteq, D_{\{1\} \cup \ldots \cup \{n\}}, D_{\{1\} \cup \ldots \cup \{n\}}), \quad (35)
\]

and, due to (30b), this structure is isomorphic to that of $P_{\Pi}$ given in (27),

\[
(P_{\Pi_I}, \subseteq, D_{\{1\} \cup \ldots \cup \{n\}}, D_{\{1\} \cup \ldots \cup \{n\}}) \equiv (P_I, \subseteq, \{1\} \cup \ldots \cup \{n\}, \{1\} \cup \ldots \cup \{n\}), \quad (36)
\]

as was mentioned before.

\section{E. Examples}

Writing out some examples explicitly might not be useless here. The lattices $P_{\Pi}$ for the cases $n = 2$ and 3 can be seen in the upper-right part of Figures 1 and 2. As we have learned in (30a) and (30b), we need to draw only this lattice for the inclusion hierarchy of the sets of $\alpha$-separable pure and mixed states $P_{\alpha}$ (28a) and $D_{\alpha}$ (28b).

For the bipartite case, we do not have additional structure over that of the first kind (22)-(23), and we get back the content of Section II B. The $\alpha$-separable pure states are

\[
P_{\{1\} \cup \{2\}} = P_{12} \equiv P(H_{12}). \quad (37a)
\]

\[
P_{\{1\} \cup \{2\}} = P_{12} \equiv P_{\text{sep}}. \quad (37b)
\]

Note that $P_{\{1\} \cup \{2\}} \subseteq P_{\{1\} \cup \{2\}}$. The $\alpha$-separable mixed states are

\[
D_{\{1\} \cup \{2\}} = D_{12} \equiv D(H_{12}), \quad (38a)
\]
Note that, again, $\mathcal{D}_{1[2]} \subseteq \mathcal{D}_{1[12]}$.

For the tripartite case we do have additional structure over that of the first kind (24)-(25). The $\alpha$-separable pure states are

$$
\mathcal{P}_{\downarrow(123)} = \mathcal{P}_{123} \equiv \mathcal{P}(\mathcal{H}_{123}),
$$

$$
\mathcal{P}_{\downarrow(123,2,13,3|12)} = \mathcal{P}_{123} \cup \mathcal{P}_{213} \cup \mathcal{P}_{312},
$$

$$
\mathcal{P}_{\downarrow(123,2,13,3|12)} = \mathcal{P}_{\mathcal{H}_{123}},
$$

$$
\mathcal{P}_{\downarrow(a|bc|ab)} = \mathcal{P}_{a|bc} \cup \mathcal{P}_{b|ab},
$$

$$
\mathcal{P}_{\downarrow(a|bc)} = \mathcal{P}_{a|bc},
$$

$$
\mathcal{P}_{\downarrow(12|3)} = \mathcal{P}_{123},
$$

with all bipartitions $a|bc$ of $\{1,2,3\}$. Note that $\mathcal{P}_{\downarrow(123)} \subseteq \mathcal{P}_{\downarrow(a|bc)} \subseteq \mathcal{P}_{\downarrow(a|bc,b|ac)} \subseteq \mathcal{P}_{\downarrow(a|bc,b|ac,c|ab)} \subseteq \mathcal{P}_{\downarrow(123)}$. The $\alpha$-separable mixed states are

$$
\mathcal{D}_{\downarrow(123)} = \text{Conv} \mathcal{P}_{\downarrow(123)} \equiv \mathcal{D}(\mathcal{H}_{123}),
$$

$$
\mathcal{D}_{\downarrow(123,2,13,3|12)} = \text{Conv} \mathcal{P}_{\downarrow(123,2,13,3|12)},
$$

$$
\mathcal{D}_{\downarrow(a|bc|ab)} = \text{Conv} \mathcal{P}_{\downarrow(a|bc|ab)},
$$

$$
\mathcal{D}_{\downarrow(a|bc)} = \text{Conv} \mathcal{P}_{\downarrow(a|bc)},
$$

$$
\mathcal{D}_{\downarrow(12|3)} = \text{Conv} \mathcal{P}_{\downarrow(12|3)}.
$$

Note that, again, $\mathcal{D}_{\downarrow(123)} \subseteq \mathcal{D}_{\downarrow(a|bc)} \subseteq \mathcal{D}_{\downarrow(a|bc,b|ac)} \subseteq \mathcal{D}_{\downarrow(a|bc,b|ac,c|ab)} \subseteq \mathcal{D}_{\downarrow(123)}$.

### F. Level III: partial separability classes

The state sets $\mathcal{D}_\alpha$ of given $\alpha$-separability are containing (30b), that is, if a state is $\alpha$-separable, $\varrho \in \mathcal{D}_\alpha$, it can happen that it is also $\beta$-separable for a $\beta$ lower in the hierarchy ($P_{11}, \preceq$). Now we construct the partial separability classes, which are the sets of states being $\alpha$-separable but not separable under any $\beta \preceq \alpha$.

The partial separability classes are defined as the intersections of the $\mathcal{D}_\alpha$ sets of states of different partial separability. First we select a sublattice of $P_{11}$,

$$
(P_{11}, \preceq, \land, \lor) \subseteq (P_{11}, \preceq, \land, \lor)
$$

by the use of which we can tune how fine/coarse the arising classification is, and what kinds of entanglement are taken into account. The elements of this (sub)lattice give rise to a (sub)hierarchy, based on which the classification is carried out. If the whole lattice is taken, $P_{11} = P_{11}$, then we get the complete classification in the sense of partial separability, which utilizes all the possible combinations of different kinds of partially separable states [2, 3]. If only the principal elements of $P_{11}$ are taken, that is, $P_{11} = \{\downarrow\{\alpha\} \mid \alpha \in P_{1}\}$, then we get an incomplete classification introduced in [27, 28]. If $P_{11} = \{\varrho \in P_{11} \mid 3k \leq n : |\alpha| = k \varrho \alpha \in \alpha\}$, then we get an intermediate classification introduced in [26].

![FIG. 2. Lattices of the labels of the first and second kinds and class labels, $P_1$, $P_{II}$ and $P_{II}$ are illustrated for $n = 3$. The partitions $\alpha \in P_1$ are denoted by small pictograms, the labels of the second kind $\alpha \in P_{II}$ are down-sets of partitions, in which case the different elements are drawn with different colors. The class labels $\alpha \in P_{II}$ are up-sets of labels of the second kind, these are written side by side. The order relation is denoted with arrow: $\beta \rightarrow \alpha$ means $\beta \preceq \alpha$ and $\beta \rightarrow \alpha$ means $\beta \preceq \alpha$. By means of (14a), (14b) and (30a), (30b), the $P_1$ and $P_{II}$ resemble the inclusion of the sets of $\alpha$-separable pure ($P_\alpha$), $\alpha$-separable mixed ($P_\alpha$), and $\alpha$-separable pure ($P_\alpha$), $\alpha$-separable mixed states ($P_\alpha$), see in Sections III B and III D. The lattice $P_{II}$ is the class hierarchy, see Section III F.](image)
\[ \alpha \subseteq P_{11^*} \text{ as} \]
\[
C_\alpha = \bigcap_{\beta \notin C_\alpha} \overline{D_\alpha} \cap \bigcap_{\alpha \in C_\alpha} D_\alpha. \tag{42}
\]

However, because of the inclusions \((30b)\), some of the intersections are empty by construction,
\[
\exists \alpha \in \alpha, \ \exists \beta \notin \alpha \colon \alpha \preceq \beta \implies C_\alpha = \emptyset. \tag{43}
\]
(This comes from elementary set-algebra: if \(A \subseteq B\) then \(B \cap A = A \setminus B = \emptyset\).) It turns out that if a class is not empty by construction, then its label \(\alpha\) is a nonempty element of the up-set lattice of \(P_{11^*}\) (see Appendix A 1), which is now denoted with
\[
P_{11} = \mathcal{O}_1(P_{11^*}) \setminus \emptyset
\]
\[
= \left\{ \alpha \in 2^{P_{11^*}} \setminus \emptyset \mid \forall \alpha \in \alpha \colon \alpha \preceq \beta \Rightarrow \beta \in \alpha \right\}. \tag{44}
\]
Again, the \(P_{11}\) set of nonempty up-sets of the lattice \(P_{11^*}\) forms a lattice with respect to the inclusion, intersection and union \([10]\), so we have
\[
(P_{11}, \preceq, \land, \lor) = (\mathcal{O}_1(P_{11^*}) \setminus \emptyset, \subseteq, \cap, \cup). \tag{45}
\]
Based on this hierarchy, it is meaningful to say that states of a given class \(C_\alpha\) are more entangled than states of \(C_\beta\), if \(\beta \supsetneq \alpha\).

With the above definitions in hand, we can prove that \(P_{11}\) is sufficient for the labelling of the classes in the above sense, that is,
\[
C_\alpha \neq \emptyset \implies \alpha \in P_{11}. \tag{46}
\]
(For the proof, see Appendix A 7.) A conjecture is that the reverse implication also holds even for the most detailed case when \(P_{11^*} = P_{11}^+ \). An advantage of the formulation by the labelling constructions is, roughly speaking, that by the use of that “we have separated the algebraic and the geometric part” of the problem of nonemptiness of the classes. At this point, it seems that we have tackled all the algebraic issues of the problem, and this conjecture can not be proven without the investigation of the geometry of \(D\), more precisely, the geometry of the different kinds of \(P_\alpha\) extremal points.

Note that, since the set of minimal elements of a nonempty up-set \(\alpha\) determines \(\alpha = \uparrow \min \alpha\), uniquely, and vice versa, both \(\min \\alpha\) and \(\alpha\) are equally suitable for the labelling of the classes.

G. Examples

Writing out some examples explicitly might not be useless here. The lattices \(P_{11}\) for \(P_{11^*} = P_{11}^+\) for the cases \(n = 2\) and \(3\) can be seen in the lower-left part of Figures 1 and 2. The classes have this hierarchical structure, however, this does not manifest itself in inclusion hierarchy, since the classes are disjoint.

For the bipartite case, we get back the content of Section II B,
\[
C_{\{1\{1\}2\}} = \overline{D_{\{1\{1\}2\}}} \cup \overline{D_{\{1\}2}} = D_{\{1\}2} = C_{\text{sep}}, \tag{47a}
\]
\[
C_{\{1\{2\}1\}} = \overline{D_{\{1\{2\}1\}}} \cup \overline{D_{\{1\}2}} = D_{\{1\}2} \setminus D_{\{1\}2} = C_{\text{ent}}, \tag{47b}
\]
being the separable and entangled state classes.

For the tripartite case, we have \(1 + 18 + 1 = 20\) classes, shown in Table III G. The meaning of the classes is discussed in [2, 3].

IV. ENTANGLEMENT MEASURES: BASICS

A very basic question of entanglement theory is how to quantify entanglement. There are many different measures of entanglement obtained by the use of two main approaches, the operational and the axiomatic ones [29]. Here we follow more-or-less the axiomatic way, because, on the one hand, it clearly distinguishes between relevant and irrelevant properties of quantities, and, on the other hand, it allows of experimenting.

Starting with this section, we mainly deal with real valued functions on state spaces, which are convex sets. On convex sets it is meaningful to define convex
\[
f\left(\sum p_i \varrho_i\right) \leq \sum p_i f(\varrho_i) \tag{48a}
\]
and concave
\[
g\left(\sum p_i \varrho_i\right) \geq \sum p_i g(\varrho_i) \tag{48b}
\]
functions. The meaning of these properties is being decreasing or increasing, respectively, for forgetting information. On convexity, see sections 2 and 3 of [12].

A. Mixedness of states

Before turning to measuring entanglement of multipartite states, in this and the next subsections, we recall some important notions in the characterization of states considered as a whole, without respect to the existence of subsystems (tensor product structure in the Hilbert space).

The mixedness of two quantum states can be characterized by real-valued functions called (generalized) entropies. The most widely used of them is the von Neumann entropy \([30, 31]\)
\[
S(\varrho) = - \text{tr}(\varrho \ln \varrho). \tag{49a}
\]

Other notable entropies are the one-parameter families of quantum Tsallis entropies \([\|]\)
\[
S_q^{\text{Tr}}(\varrho) = \frac{1}{1 - q} (\text{tr} \varrho^q - 1), \quad q > 0, \tag{49b}
\]
and quantum Rényi entropies \[32\]

\[ S_q^R(\varrho) = \frac{1}{1 - q} \ln \text{tr} \varrho^q, \quad q > 0. \tag{49c} \]

The \textit{concurrence-squared} is a qubit-normalized version of the \( q = 2 \) Tsallis entropy,

\[ C^2(\varrho) = 2S_{Ts}^2(\varrho) = 2(1 - \text{tr} \varrho^2), \tag{50} \]

for qubits, it obeys \( 0 \leq C^2(\varrho) \leq 1 \). (The same holds if \( \log_2 \) is used in the definitions of von Neumann and Rényi entropies.)

All of these are nonnegative, vanishing exactly for pure states,

\[ S(\varrho), S_q^{Ts}(\varrho), S_q^R(\varrho) \geq 0, \]

\[ S(\varrho), S_q^{Ts}(\varrho), S_q^R(\varrho) = 0 \iff \varrho \in \mathcal{P} \subset \mathcal{D}. \tag{51} \]

It is also important to know that not all Rényi entropies are concave (48b),

\[ S\left(\sum_i p_i \varrho_i\right) \geq \sum_i p_i S(\varrho_i), \tag{52a} \]

\[ S_q^{Ts}\left(\sum_i p_i \varrho_i\right) \geq \sum_i p_i S_q^{Ts}(\varrho_i) \quad \text{for all } q > 0, \tag{52b} \]

\[ S_q^R\left(\sum_i p_i \varrho_i\right) \geq \sum_i p_i S_q^R(\varrho_i) \quad \text{if } q \leq 1. \tag{52c} \]

(For some useful tools in matrix analysis, see Appendix B1, and [15, 31, 33].)

A common property of these functions is that they are monotonically increasing in \textit{bistochastic quantum channels} \( \Phi \),

\[ S(\Phi(\varrho)) \geq S(\varrho), \tag{53a} \]

\[ S_q^{Ts}(\Phi(\varrho)) \geq S_q^{Ts}(\varrho), \tag{53b} \]

\[ S_q^R(\Phi(\varrho)) \geq S_q^R(\varrho). \tag{53c} \]

(For the theory of quantum channels, see, for example, [11, 13, 15–17].)

In quantum probability theory, contrary to the classical one, the entropy is not monotonically decreasing for the restriction to subsystems (partial trace), e.g.,

\[ S(\varrho_{KK'}) \not\leq S(\varrho_K), \]

for the disjoint subsystems \( K \) and \( K' \). (For pure states, this is entanglement itself, see (3), (51).) However, the \textit{subadditivity} holds in some cases [34, 35],

\[ S(\varrho_{KK'}) \leq S(\varrho_K) + S(\varrho_{K'}), \tag{54a} \]

\[ S_q^{Ts}(\varrho_{KK'}) \leq S_q^{Ts}(\varrho_K) + S_q^{Ts}(\varrho_{K'}) \quad \text{for } q > 1 \tag{54b} \]

Unfortunately, the Rényi entropies are not subadditive [36].

\section*{B. Distinguishability of states}

There are several quantities measuring the distinguishability of two quantum states, here we consider only the \textit{Unegaki relative entropy} or quantum Kullback-Leibler divergence [31, 37]. For the density matrices \( \varrho, \omega \in \mathcal{D} \), it is given as

\[ D_{KL}(\varrho\|\omega) = \text{tr} \varrho (\ln \varrho - \ln \omega). \tag{55} \]

This expresses the \textit{statistical distinguishability} of the state \( \varrho \) from the state \( \omega \) [11? ]. It is nonnegative, and vanishes if and only if the two states are equal,

\[ D_{KL}(\varrho\|\omega) \geq 0, \quad D_{KL}(\varrho\|\omega) = 0 \iff \varrho = \omega. \tag{56} \]

It is not a distance, but only a divergence, since it is not symmetric, and only a weak version of the triangle inequality holds [11]. It is also jointly convex,

\[ D_{KL}\left(\sum_i p_i \varrho_i \bigg\| \sum_i p_i \omega_i\right) \leq \sum_i p_i D_{KL}(\varrho_i\|\omega_i), \tag{57} \]
from what the convexity (48a) holds in both arguments separately. An important property of the relative entropy is that it is monotonically decreasing in quantum channels \( \Phi \),

\[
D_{\text{KL}}(\Phi(\varrho)\|\Phi(\omega)) \leq D_{\text{KL}}(\varrho\|\omega).
\]

(58)

For nice summaries on the properties and meaning of the relative entropy, see, for example, [11, 31, 38]. There are also Rényi and Tsallis versions [\ref{39}]

\[
\text{C. LOCC monotonicity: entanglement measures}
\]

The most fundamental property of entanglement measures [39] is the monotonicity under LOCC (local operation and classical communication, [6, 20, 29, 40]). An \( f : \mathcal{D} \rightarrow \mathbb{R} \) is (nonincreasing) monotonic under LOCC, if

\[
f(\Lambda(\varrho)) \leq f(\varrho)
\]

(59a)

for any LOCC transformation \( \Lambda \), which expresses that, (i) as any correlation, entanglement does not increase locally, and, (ii) while classical correlation does, entanglement does not increase by classical communication (“classical interaction”) either. An \( f : \mathcal{D} \rightarrow \mathbb{R} \) is non-increasing on average under LOCC if

\[
\sum_i p_i f(\varrho'_i) \leq f(\varrho).
\]

(59b)

for all ensembles \( \varrho \mapsto \{ (p_i, \varrho_i) \} \) resulted from LOCC transformation \( \Lambda \), where the LOCC is constituted as \( \Lambda = \sum_i \Lambda_i \), where the \( \Lambda_i \)'s are the suboperations of the LOCC realizing the outcomes of selective measurements, and \( \varrho'_i = \frac{1}{p_i} \Lambda_i(\varrho) \) with \( p_i = \text{tr} \Lambda_i(\varrho) \). This latter condition is stronger than the former one if the function is convex (48a),

\[
f(\sum_i p_i \varrho_i) \leq \sum_i p_i f(\varrho_i)
\]

(59c)

for all ensembles \( \{ (p_i, \varrho_i) \} \), which expresses that entanglement can not increase for mixing. This is also a fundamental, and also plausible property, since mixing is interpreted as forgetting some classical information concerning the identity of a \( \varrho \) member of an ensemble \( \{ (p_i, \varrho_i) \} \), what can be done locally [40]. An \( f : \mathcal{D} \rightarrow \mathbb{R} \) is called an entanglement monotone if (59b) and (59c) hold [40]. There is common agreement that LOCC-monotonicity (59a) is the only necessary postulate for a function to be an entanglement measure [1], however, the stronger condition (59b) is often satisfied too, and it is often easier to prove.

If \( f \) is defined only for pure states, \( f : \mathcal{P} \rightarrow \mathbb{R} \), then only (59b) makes sense, whose restriction is that a pure function is nonincreasing on average under pure LOCC or entanglement monotone,

\[
\sum_i p_i f(\varrho'_i) \leq f(\varrho).
\]

(60)

Here \( \pi \mapsto \{ (p_i, \pi'_i) \} \) is the ensemble of pure states arising from the pure LOCC suboperations \( \Lambda_i \). That is, mathematically, one can decompose the LOCC \( \Lambda \) into pure suboperations having only one Kraus operator each, leading to much simpler constructions. Note that not all \( \pi'_i \) results of these operations may be accessible physically, only the outcomes of the LOCC, which are formed by partial mixtures of this ensemble [29].

Clearly, functions obeying any particular one of the requirements in (59) and (60) form a cone, that is, their sums and multiples by nonnegative real numbers are also obey the particular requirement.

Since fully separable states can be reversibly converted into each other by means of LOCC, it follows that if a function obey (59a), then it takes the same value for all fully separable states [40].

\[
\text{D. Discriminance: indicator functions}
\]

We will extensively use another property of functions \( f : \mathcal{D} \rightarrow \mathbb{R} \) on state spaces, which is the discriminance with respect to a convex set \( \mathcal{D}_* \subseteq \mathcal{D} \), that is,

\[
\varrho \in \mathcal{D}_* \iff f(\varrho) = 0.
\]

(61a)

So the vanishing of the function gives a necessary and sufficient criterion for that subset. In this paper we deal only with functions having this property, which are often called indicator functions with respect to a kind of states.

If \( f \) is defined only for pure states, \( f : \mathcal{P} \rightarrow \mathbb{R} \), then the discriminance for the closed set \( \mathcal{P}_* \subseteq \mathcal{P} \) is

\[
\pi \in \mathcal{P}_* \iff f(\pi) = 0.
\]

(61b)

Discriminance with respect to \( \mathcal{D}_{\text{sep}} \) is an important property for functions measuring bipartite entanglement (Section II B).

\[
\text{E. Local entropies: pure state measures}
\]

First, we can get pure state measures. It is proven by Vidal [29, 40] that any properly choosen function applied for one of the reduced density matrices of a pure state leads to a measure of pure state entanglement in the sense of (60).

\[
\text{Theorem 1 Let } F : \mathcal{D}(\mathcal{H}_K) \rightarrow \mathbb{R} \text{ be (i) symmetric and extensible function of the eigenvalues, and (ii) concave (48b),}
\]

\[
F(\sum_i p_i \varrho_i) \geq \sum_i p_i F(\varrho_i)
\]

(62)

then \( f : \mathcal{P} \rightarrow \mathbb{R} \) defined as

\[
f_K(\pi) := F(\text{tr}_\mathcal{P} \pi)
\]

(63)

is an entanglement monotone (60).
We recall the simpler proof of Horodecki [29] in Appendix B 2. It turns out that, roughly speaking, the entanglement monotonicity (60) is the concavity on the subsystem.

This construction characterizes the entanglement of the subsystem $K$ with the rest of the system $\overline{K}$, that is, bipartite entanglement with respect to the split $K|\overline{K}$.

**F. Convex roof extensions: mixed state measures**

The pure state entanglement measures can be extended to mixed states by the so called convex roof extension [5, 6, 41–43]. It is motivated by the practical approach of the optimal mixing of the mixed state from pure states, that is, using as little amount of pure state entanglement as possible. For a continuous function $f : \mathcal{P} \to \mathbb{R}$, its convex roof extension $f^\uparrow : \text{Conv} \mathcal{P} = D \to \mathbb{R}$ is defined as

$$f^\uparrow(\rho) = \min_{\sum_i p_i \pi_i = \rho} \sum_i p_i f(\pi_i),$$

where the minimization takes place over all $\{(p_i, \pi_i)\}$ pure state decompositions of $\rho$. It follows from Schrödinger’s mixture theorem [44], also called Gisin-Hughston-Jozsa-Wootters lemma [45, 46], that the decompositions of a mixed state into an ensemble of $m$ pure states are labelled by the elements of a Stiefel manifold, which is a compact complex manifold. On the other hand, the Carathéodory theorem ensures that we need only finite $m$, or, to be more precise, $m \leq (\text{rk} \rho)^2 \leq (\dim \mathcal{H})^2$, shown by Uhlmann [47]. These observations guarantee the existence of the minimum in (64).

Obviously, for pure states the convex roof extension is trivial,

$$\forall \pi \in \mathcal{P} : f^\uparrow(\pi) = f(\pi)$$

The convex roof extension of a function is convex (59c),

$$f^\uparrow \left( \sum_i p_i \rho_i \right) \leq \sum_i p_i f^\uparrow(\rho_i),$$

moreover, it is the largest convex function taking the same values for pure states as the original function [47].

It is proven by Vidal [29, 40] that if a function $f : \mathcal{P} \to \mathbb{R}$ is nonincreasing on average for pure states (60), then its convex roof extension is also nonincreasing on average for mixed states (59b). That is, we have the following theorem.

**Theorem 2** For a continuous $f : \mathcal{P} \to \mathbb{R}$,

$$\sum_i p_i f(\pi_i') \leq f(\pi) \implies \sum_i p_i f^\uparrow(\rho_i') \leq f^\uparrow(\rho),$$

for all $\pi \mapsto \{(p_i, \pi_i')\}$ and $\rho \mapsto \{(p_i, \rho_i')\}$ ensembles resulting from LOCC.

We recall the simpler proof of Horodecki [29] in Appendix B 3. Because of the latter two properties, $f^\uparrow(\rho)$ is also an entanglement-monotone (59b)-(59c).

It is remarkable that in Theorem 1 a reverse implication hold in the bipartite case: all bipartite mixed entanglement monotones (satisfying (59b) and (59c)) can be obtained in the above way, that is, by the use of the convex roof extension of an $f$ given by an $F$ satisfying (i) and (ii) of Theorem 1 applied to the reduced density matrix.

The convex roof extension preserves the discrimination property (61b) if we additionally assume that $f \geq 0$,

$$\left( \pi \in \mathcal{P}_* \iff f(\pi) = 0 \right) \implies \left( \varrho \in \mathcal{D}_* \iff f^\uparrow(\varrho) = 0 \right),$$

what can be used for the detection of mixed state entanglement. (For the proof, see Appendix B 4.) Note that these properties are based more-or-less only on that $\mathcal{P}_* = \text{Extr} \mathcal{D}_*$ and $\mathcal{D}_* = \text{Conv} \mathcal{P}_*$. Another important property of the convex roof construction is the mononicity. For functions $f, g : \mathcal{P} \to \mathbb{R}$,

$$\forall \pi \in \mathcal{P} : f(\pi) \leq g(\pi) \iff \forall \varrho \in \mathcal{D} : f^\uparrow(\varrho) \leq g^\uparrow(\varrho)$$

(For the proof, see Appendix B 5.)

**G. Examples**

For recalling some well-known examples, let us consider the bipartite case, with the notations of Section II B. Particular choices for functions fulfilling the requirements in Theorem 1 are some entropies given in Section IV A. Since the entangled pure states are the ones which have mixed marginals (3), it is expressive to say that “the more mixed the marginals is the more entangled the state is”. In particular, using the $F = S : \mathcal{D} \to \mathbb{R}$ von Neumann entropy (49a) by construction (63) leads to the “Entanglement Entropy”,

$$f(\psi) := E(\psi) := S(\text{tr}_2 \pi)$$

which is also called simply “Entanglement”. [Note that the spectra of the marginals of a bipartite pure state are the same, apart from the multiplicity of the zero eigenvalues.] Apart the von Neumann entropy, the Tsallis entropies (49b) for all $0 < q$, and the Rényi entropies (49c) for all $0 < q < 1$ [40] are known to be concave (52), and all of them are symmetric and extensible function of the eigenvalues. They lead to the “Tsallis or Rényi Entropy of Entanglement”,

$$f(\psi) := E^q(\psi) := S^q_\text{r}(\text{tr}_2 \pi),$$

A particular choice is the “concurrence (of entanglement)”, with the concurrence (50),

$$f(\psi) := C(\psi) := C(\text{tr}_2 \pi).$$
All of the above functions measure the pure bipartite entanglement in the sense that they satisfy (60) by Theorem 1, and they are indicators of pure separability, that is, discriminant (61b) with respect to $P_{\text{sep}}$ of (2a) by (3) and (51).

Having these pure measures in hand, thanks to Theorem 2, we can extend them to mixed states by the use of convex roof extension (64). The resulting measures are called “Entanglement of Formation” [6], “Tsallis or Rényi Entanglement of Formation” [40] and “Concurrence of Formation” [48].

\[
E_{\text{of}}(\rho) := E^J(\rho),
\]

\[
E_{\text{of},q}(\rho) := E^T_{q,J}(\rho),
\]

\[
E_{\text{of},q}^R(\rho) := E^R_{q,J}(\rho),
\]

\[
C_{\text{of}}(\rho) := C^J(\rho).
\]

All of these functions measure the mixed bipartite entanglement in the sense that they satisfy (59b) and (59c) by Theorem 2, and they are indicators of mixed separability, that is, discriminant (61a) with respect to $P_{\text{sep}}$ by (68). A remarkable result of Wooters is a closed formula for the minimization in the convex roof extension in the Entanglement of Formation (through that for the Concurrence of Formation) for the case when $\dim \mathcal{H}_1 = \dim \mathcal{H}_2 = 2$, that is, for two qubits [48, 49].

V. SUMS AND MEANS: A DETOUR

In the sequel, we will need to construct entanglement measures as functions of more basic ones in a systematic way. For these, we need some properties being valid, such as monotonicity, homogeneity, concavity, permutation invariance in many cases. The power-sums and power-means, or the more general quasi-sums and quasi-arithmetic means turn out to be suitable tools in this situation. Power-means equate things, what is sometimes undesirable for our investigations, so power-sums turn out to be more suitable in these cases. Apart from this, they share the properties most important for us, such as monotonicity, convexity/concavity and vanishing properties. Moreover, the power-sums and power-means of homogeneous functions of a given degree is of the same degree, what is a property which seems to be of great importance in the topic of entanglement of pure states.

A. The meaning of sums and means

Let us suppose that we have a nonegative quantity $X$, which can characterize $m$ different entities as $X_1, \ldots, X_m$, and which can also characterize these entities “together”, as a total value $X_{\text{tot}}$. Suppose moreover that we have a “law” saying that the total value of this quantity and the values for the individual entities are connected by a summation for their $q$-th powers, as

\[
Y = X_1^q + \ldots + X_m^q.
\]

Then, on the one hand, the total value is

\[
X_{\text{tot}} = (X_1^q + \ldots + X_m^q)^{1/q},
\]

which is called power-sum (or $q$-sum). On the other hand, a natural question is what is the “mean” value of this quantity in this situation, that is, what is the uniform value for all $X_j$ which leads to the same $Y$ under the same law,

\[
Y = X_1^q + \ldots + X_m^q = X_1^q + \ldots + X_m^q = mX_{\text{mean}}^q.
\]

This leads to

\[
X_{\text{mean}} = \left[\frac{1}{m} (X_1^q + \ldots + X_m^q)\right]^{1/q},
\]

which is called power-mean (or $q$-mean, or Hölder mean).

For $q = 1$, we get back the sum and the arithmetic mean. Well-known examples are the total and the mean resistance of $m$ resistors connected in series (or total/mean conductance in parallel) or the total and the mean capacity of $m$ capacitors connected in parallel. For $q = -1$, we get back the harmonic sum and the harmonic mean. Well-known examples are the total and the mean resistance of $m$ resistors connected in parallel (or total/mean conductance in series) or the total and the mean capacity of $m$ capacitors connected in series. For $q = 2$, we get back the quadratic sum and the quadratic mean. If we consider an $m$ dimensional hypercuboid of edges of length $X_j$, then the quadratic mean of the length of the edges is the length of edges of an $m$-dimensional hypercube having diagonal of the same length as the original hypercuboid. In this case, the meaning of $X_{\text{tot}}$ is the length of the diagonal.

A conceptually (but mathematically not too much) different situation is when the “law” is about products,

\[
Y = X_1 \cdot \ldots \cdot X_m.
\]

This leads to the

\[
X_{\text{mean}} = (X_1 \cdot \ldots \cdot X_m)^{1/m}
\]

geometric mean. We will see that this can be obtained as the power mean for $q = 0$.

If we consider an $m$ dimensional hypercuboid of edges of length $X_j$ again, then the geometric mean of the length of the edges is the length of the edge of a hypercube of the same volume. In this case, the meaning of $Y$ is the volume.

A more general, but still relevant situation is when the “law” involves summation of more distorted values, as

\[
Y = h(X_{\text{tot}}) = h(X_1) + \ldots + h(X_m)
\]

for some invertible $h$. Then, on the one hand, the total value is

\[
X_{\text{tot}} = h^{-1}(h(X_1) + \ldots + h(X_m)),
\]
which is called quasi-sum. On the other hand, for the uniform value this leads to the quasi-arithmetic mean (or Kolmogorov mean)

\[ X_{\text{mean}} = h^{-1}\left( \frac{1}{m} (h(X_1) + \ldots + h(X_m)) \right). \]

The \( h(x) = x^q \) gives back the \( q \)-mean for \( q \neq 0 \), while \( h(x) = \ln(x) \) gives back the geometric mean.

**B. Definitions and properties of power-sums and power-means**

Let \( x = (x_1, \ldots, x_m) \in \mathbb{R}^m, x > 0 \). (The latter is meant elementwise, \( x_j > 0, j = 1, \ldots, m \).)

The \( q \)-sums of \( x \) is defined for nonzero \( q \in \mathbb{R} \) parameters as

\[ N_q(x) = \left[ \sum_{j=1}^m x_j^q \right]^{1/q} \quad \text{for } q \neq 0. \quad (72a) \]

It can be defined for positive and negative infinities by its limits, leading to

\[ N_{\pm \infty}(x) := \lim_{q \to \pm \infty} N_q(x) = \max(x_j), \quad (72b) \]
\[ N_{-\infty}(x) := \lim_{q \to -\infty} N_q(x) = \min(x_j). \quad (72c) \]

Since \( \lim_{q \to q^+} N_q(x) = \infty \) and \( \lim_{q \to q^-} N_q(x) = 0 \), \( N_q \) is not continuous in \( q = 0 \). We have the usual sum \( N_1 \), the harmonic sum \( N_{-1} \), and the quadratic sum \( N_2 \).

The power-means or \( q \)-means of \( x \) is defined for nonzero \( q \in \mathbb{R} \) parameters as

\[ M_q(x) = \left[ \frac{1}{m} \sum_{j=1}^m x_j^q \right]^{1/q} \quad \text{for } q \neq 0. \quad (73a) \]

It can be defined for parameter zero and for positive and negative infinities by its limits, leading to

\[ M_0(x) := \lim_{q \to 0} M_q(x) = \left[ \prod_j x_j \right]^{1/m}, \quad (73b) \]
\[ M_{+\infty}(x) := \lim_{q \to +\infty} M_q(x) = \max(x_j), \quad (73c) \]
\[ M_{-\infty}(x) := \lim_{q \to -\infty} M_q(x) = \min(x_j). \quad (73d) \]

We have the geometric mean \( M_0 \), the usual arithmetic mean \( M_1 \), the harmonic mean \( M_{-1} \), and the quadratic mean \( M_2 \).

The power-sums and power-means above are defined for strictly positive \( x_j \)'s, however, we would like to use them for nonnegative values too. For \( q > 0 \), the definitions (72a) and (73a) work well for \( x_j = 0 \) values. For \( q < 0 \), notice that

\[ M_q(x) = M_{-|q|}(x) = \left[ \frac{1}{x_1^{|q|}} + \ldots + \frac{1}{x_m^{|q|}} \right]^{-1/|q|}, \]

showing that \( \lim_{x_j \to 0^+} M_q(x) = 0 \), and the same holds for \( N_q \). So defining for \( q \leq 0 \) and for any \( x_j = 0 \) the \( q \)-sum and \( q \)-mean by their limit, \( N_q(x) = M_q(x) = 0 \), allows us to use \( q \)-sum and \( q \)-mean of \( x \geq 0 \) nonnegative numbers.

Let us see the most important properties of \( q \)-sums and \( q \)-means. \( N_q \) is continuous for \( 0 \neq q \in \mathbb{R}, N_q \geq 0, M_q \) is continuous for \( q \in \mathbb{R}, M_q \geq 0 \), and have the vanishing properties

if \( q > 0 \):
\[ N_q(x) = 0 \iff \forall j : x_j = 0, \quad (74a) \]
\[ M_q(x) = 0 \iff \exists j : x_j = 0. \quad (74b) \]

and

if \( q > 0 \):
\[ M_q(x) = 0 \iff \forall j : x_j = 0, \quad (75a) \]
\[ M_q(x) = 0 \iff \exists j : x_j = 0. \quad (75b) \]

They are homogeneous functions, that is,

for \( c \geq 0 \):
\[ N_q(cx) = cN_q(x), \quad M_q(cx) = cM_q(x). \quad (76) \]

For all \( 0 \neq q \in \mathbb{R}, N_q(x) \) and for all \( q \in \mathbb{R}, M_q(x) \) are monotonically increasing for all arguments \( x_j \) (see in Appendix C1), and their convexity/concavity

\[ N_q\left( \sum_i p_i x_i \right) \leq \sum_i p_i N_q(x_i) \iff q \geq 1, \quad (77a) \]
\[ N_q\left( \sum_i p_i x_i \right) \geq \sum_i p_i N_q(x_i) \iff 0 \neq q \leq 1, \quad (77b) \]

and

\[ M_q\left( \sum_i p_i x_i \right) \leq \sum_i p_i M_q(x_i) \iff q \geq 1, \quad (78a) \]
\[ M_q\left( \sum_i p_i x_i \right) \geq \sum_i p_i M_q(x_i) \iff q \leq 1 \quad (78b) \]

(see in Appendix C2). On the other hand, for all \( x \geq 0 \), \( N_q(x) \) is monotonically decreasing for the parameter \( q \), however, having a discontinuity in \( q = 0 \), as

\[ 0 < q < q' \implies M_q(x) \geq M_{q'}(x), \quad (79a) \]
\[ q < q' < 0 \implies M_q(x) \geq M_{q'}(x), \quad (79b) \]
\[ 0 < q < q' \implies M_q(x) \leq M_{q'}(x). \quad (79c) \]

On the other hand, for all \( x \geq 0 \), \( M_q(x) \) is monotonically increasing for the parameter \( q \), which is the power-mean inequality:

\[ q < q' \implies M_q(x) \leq M_{q'}(x), \quad (80) \]

with equality if and only if \( x_1 = \cdots = x_m \). From this, it immediately follows that the \( q \)-mean of numbers is between the minimal (73d) and the maximal (73c) ones. On the other hand, the inequality between the arithmetic and geometric means \( M_0(x) \leq M_1(x) \) is a particular case of this.
C. Definitions and properties of quasi-sums and quasi-arithmetic means

Let \( \mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m \), \( \mathbf{x} > 0 \). For a continuous strictly monotonic function \( h : \mathbb{R} \to \mathbb{R} \), the quasi-sum of \( \mathbf{x} \) is defined as

\[
N_h(\mathbf{x}) = h^{-1}\left(\sum_{j=1}^{m} h(x_j)\right),
\]

and the quasi-arithmetic mean of \( \mathbf{x} \) is defined as \([50, 51]\)

\[
M_h(\mathbf{x}) = h^{-1}\left(\frac{1}{m} \sum_{j=1}^{m} h(x_j)\right).
\]

Note that \( h \) and \( ah + b \) with \( a, b \in \mathbb{R}, a \neq 0 \), leads to the same function, \( N_h = N_{ah+b}, M_h = M_{ah+b} \). The choice \( h(x) = x^2 \) gives back the q-sum and q-mean for \( q \neq 0 \), while \( h(x) = \ln(x) \) gives back the geometric mean. For these quite general function classes, only few of the properties can be known in general.

D. Means of measures

Some power-sums and power-means preserve entanglement monotonicity and discrimination for pure states. We have in general the following lemma about entanglement monotonicity.

**Lemma 3** Let \( f_j : \mathcal{P} \to \mathbb{R} \) be nonnegative functions for \( j = 1, \ldots, m \), which are pure entanglement monotones \((60)\). If \( G : \mathbb{R}^m \to \mathbb{R} \) is monotonic increasing in all arguments and concave, then \( G(f_1, \ldots, f_m) : \mathcal{P} \to \mathbb{R} \) is an entanglement monotone \((60)\), that is,

\[
\sum_i p_i G(f_1, \ldots, f_m)(\pi_i) \leq G(f_1, \ldots, f_m)(\pi)
\]

for all \( \pi \mapsto \{(p_i, \pi_i)\} \) ensembles resulting from LOCC.

This is a simple consequence of the concavity (see in Appendix D1). Note that similar results can be proven for mixed states for the properties \((59a)\) and \((59b)\), however, the convexity \((59c)\) would, of course, fail. Because of the monotonicity and \((77b), (78b)\), we have

**Corollary 4** The q-sum \((72)\) and q-mean \((73)\) of entanglement monotones \((60)\) for \( q \leq 1 \), that is,

\[
\sum_i p_i N_q(f_1, \ldots, f_m)(\pi_i) \leq N_q(f_1, \ldots, f_m)(\pi)
\]

for \( 0 \neq q \leq 1 \),

\[
\sum_i p_i M_q(f_1, \ldots, f_m)(\pi_i) \leq M_q(f_1, \ldots, f_m)(\pi)
\]

for \( q \leq 1 \),

for all \( \pi \mapsto \{(p_i, \pi_i)\} \) ensemble resulting from LOCC.

We have in general the following lemma about entanglement discrimination.

**Lemma 5** Let \( f_j : \mathcal{P} \to \mathbb{R} \) be nonnegative functions for \( j = 1, \ldots, m \), which are discriminant with respect to a \( \mathcal{P}_1 \subseteq \mathcal{P} \) set \((61b)\). If \( G : \mathbb{R}^m \to \mathbb{R} \) obeys the vanishing properties

\[
G(\mathbf{x}) = 0 \iff \exists j : x_j = 0,
\]

then \( G(f_1, \ldots, f_m) : \mathcal{P} \to \mathbb{R} \) is also discriminant with respect to the same set \((61b)\).

(This is obvious.) Because of \((74)\) and \((75)\), we have

**Corollary 6** The q-sum \((72)\) and q-mean \((73)\) of indicators with respect to a \( \mathcal{P}_1 \subseteq \mathcal{P} \) set \((61b)\) are also indicators with respect to the same set \((61b)\).

A more interesting situation arises when the different \( f_j \) functions are discriminant with respect to different sets, as we will see in the sequel.

VI. ENTANGLEMENT MEASURES FOR MULTIPARTITE SYSTEMS

In Section III, we have introduced the different meaningful kinds of partial separability, and built up a hierarchy of those. Now we construct a hierarchy of entanglement measures which resembles this hierarchical structure.

In Section IV, we had two main requirements for quantities measuring bipartite entanglement: the LOCC monotonicity in the sense of \((59a)\) and the discrimination \((61a)\) with respect to the given separability. For the hierarchy of quantities measuring multipartite entanglement we introduce a third requirement: the multipartite monotonicity, which reflects a natural relation among these quantities, and connects them to the hierarchy of entanglement. By this property we can grasp the hierarchy of multipartite entanglement by the measures. These three requirements seem to be mandatory. Also, a fourth one should be satisfied by these measures: being meaningful in some sense. This last one is quite hard to fulfill, but not impossible.

The construction of measures here reflects the construction of the partial separability hierarchy in Section III. It is based on the measures of pure bipartite entanglement (Section VI A), from which the first and second kind hierarchy of measures of pure multipartite entanglement are built up (Sections VI B and VI D). We turn to mixed states only in the final step (Section VI F), by the use of convex roof extension, as in the bipartite construction in Section IV G, then follows the detection of the classes (Section VIII).
A. Level 0: bipartite entanglement

Following Section IV.E, let \( F \) satisfy (i) and (ii) of Theorem 1, and

\[
F(\varrho) \geq 0, \\
F(\varrho) = 0 \iff \varrho \in \mathcal{P}.
\]  
(87)

With this, in the sense of Theorem 1, the function

\[
f_K := F \circ \text{tr}_K : \mathcal{P} \to \mathbb{R}
\]

is an entanglement monotone (60) indicator function (61b) with respect to \( \mathcal{P}_{K|\overline{K}} \), that is, for the pure bipartite entanglement with respect to the bipartite split \( K|\overline{K} \). The latter is

\[
f_K(\pi) = 0 \iff \pi \in \mathcal{P}_{K|\overline{K}},
\]

which is the consequence of (87) and (3).

B. Level I: multipartite entanglement measures of the first kind

In Section III.B, we have the \( (P_1, \unlhd) \) partial separability hierarchy of the first kind (10). Now we consider the \( f_\alpha : \mathcal{P} \to \mathbb{R} \) functions for all \( \alpha \in P_1 \) labels of the first kind (partitions),

\[
P_{1,f} = \{ f_\alpha : \mathcal{P} \to \mathbb{R} \mid \alpha \in P_1 \},
\]

and formulate their important properties expected for the measuring of the pure \( \alpha \)-entanglement. Entanglement monotonicity (60) is, of course, mandatory for all pure state measures. The others are as follows.

For the \( \alpha \) label of the first kind (partition), the function \( f_\alpha : \mathcal{P} \to \mathbb{R} \) is called pure \( \alpha \)-indicator function (or indicator function of the first kind with respect to \( \mathcal{P}_\alpha \)), if it is discriminant with respect to \( \mathcal{P}_\alpha \), (61b), that is, if it vanishes exactly for \( \alpha \)-separable pure states (12a),

\[
f_\alpha(\pi) = 0 \iff \pi \in \mathcal{P}_\alpha.
\]

(91)

Using (19), one can formulate the vanishing of the \( \alpha \)-indicator function \( f_\alpha \) by the vanishing (89) of the functions \( f_K \) of (88) as

\[
f_\alpha = 0 \iff \forall K \in \alpha : f_K = 0.
\]

(92)

From this and the inclusion hierarchy (14a), we immediately have that the indicator functions (91) obey

\[
\alpha \unlhd \beta \implies (f_\alpha = 0 \implies f_\beta = 0).
\]

(93)

That is, separability with respect to a finer split implies that with respect to a coarser one, as it has to be. This connects the \( P_{1,f} \) set of functions to the partial separability hierarchy of the first kind \( P_1 \) (10) and \( P_{1,P} \) (13a) in a sense. In addition to this, one can formulate a stronger property reflecting a stronger connection between these two structures, what provides \( P_{1,f} \) with a hierarchical structure. For the \( P_1 \) labels of the first kind (partition), the set of functions \( P_{1,P} \) is called multipartite-monotonic of the first kind, if

\[
\alpha \preceq \beta \implies f_\alpha \preceq f_\beta.
\]

(94)

(The map \( \alpha \mapsto f_\alpha \) is monotonically decreasing with respect to the labels of the first kind.) That is, entanglement with respect to a coarser partition cannot be higher than entanglement with respect to a finer one. The multipartite monotonicity (94) is indeed stronger than the vanishing implications (93), since the latter one follows from the former one. The multipartite monotonicity, on the other hand, gives the same hierarchic structure to \( P_{1,f} \) as \( P_1 \) and \( P_{1,P} \) to

\[
(P_{1,f}, \preceq) \cong (P_1, \preceq).
\]

(95)

With the above definitions in hand, we construct multipartite monotonic (94) hierarchies of entanglement measures for pure states for the hierarchy of the first kind, consisting of entanglement monotonic (60) \( \alpha \)-indicator functions (91). Let us start with the construction of \( \alpha \)-indicators, then check the monotonicity properties.

There are several ways of constructing \( \alpha \)-indicator functions (91), based on the \( K|\overline{K} \)-indicators (88) as in (92). Perhaps the simplest one is a sum,

\[
f_\alpha := \sum_{K \in \alpha} f_K.
\]

(96a)

It clearly obeys \( \alpha \)-discriminance (91) through (92), and entanglement monotonicity (60). (For convenience, one can also use the definition \( f_\alpha := \frac{1}{|\alpha|} \sum_{K \in \alpha} f_K \), leading to \( f_K|\overline{K} = f_K \).) Another candidate is the arithmetic mean,

\[
f_\alpha := \frac{1}{|\alpha|} \sum_{K \in \alpha} f_K = M_1(f_{K_1}, \ldots, f_{K_{|\alpha|}}),
\]

(96b)

which is just a sum, multiplied by a factor \( \frac{1}{|\alpha|} \), what does not ruin the entanglement monotonicity and \( \alpha \)-discriminance. One can notice that we can use power means (73) with general parameters \( q \),

\[
f_\alpha := N_q(f_{K_1}, \ldots, f_{K_{|\alpha|}}), \quad 0 < q \leq 1,
\]

(96c)

\[
f_\alpha := M_q(f_{K_1}, \ldots, f_{K_{|\alpha|}}), \quad 0 < q \leq 1.
\]

(96d)

Indeed, power sums and power means are concave for \( q \leq 1 \), see (77b) and (78b), what is needed for the entanglement monotonicity (60), (see Corollary 4), while the proper vanishing properties (74a) and (75a) are satisfied for \( 0 < q \), what is needed for the \( \alpha \)-discriminance (91) through (92).

Now we would like to argue that, from the constructions above, the simplest choice is the best motivated:
the sum (96a). First of all, as we have learned in Section V A, using means would infer an underlying “law”, saying that the sum of the \(q\)-th power of the functions \(f_K\) is meaningful. This seems to be true for \(q = 1\) only. But in this case, a sum may have more meaning than the arithmetic \((q = 1)\) mean. Let us see, why. Taking \(F = S\) with the von Neumann entropy (49a), we simply get
\[
f_\alpha(\pi) = S(\pi_{K_1}) + S(\pi_{K_2}) + \cdots + S(\pi_{K_{|\alpha|}}),
\]
the sum of the entropies of disjoint subsystems given by the split \(\alpha = K_1|K_2|\ldots|K_{|\alpha|}\). And this has a well-defined meaning. The divergence (55) of \(\varrho \in \mathcal{D}\) from its uncorrelated version (with respect to \(\alpha\)),
\[
D^{\text{KL}}(\varrho \| \varrho_{K_1} \otimes \varrho_{K_2} \otimes \cdots \otimes \varrho_{K_{|\alpha|}}) = S(\varrho_{K_1}) + S(\varrho_{K_2}) + \cdots + S(\varrho_{K_{|\alpha|}}) - S(\varrho)
\]
is a possible multipartite generalization of the mutual information [52]. This characterizes the all the correlations (classical and quantum, [53]) contained in the state \(\varrho\) with respect to the split \(\alpha = K_1|K_2|\ldots|K_{|\alpha|}\), in the sense that it expresses how easy to distinguish the state from its uncorrelated “part”. Moreover, it can be proven [53] that
\[
\arg\min_{\varrho_K \in \mathcal{D}_K \forall K \in \alpha} \{ D^{\text{KL}}(\varrho \| \bigotimes_{K \in \alpha} \varrho_K) \} = \bigotimes_{K \in \alpha} \varrho_K,
\]
that is, the least indistinguishable state which is uncorrelated with respect to \(\alpha\) is formed by the marginals of the state, so the mutual information is actually a geometric measure of correlation,
\[
\min_{\varrho_K \in \mathcal{D}_K \forall K \in \alpha} \{ D^{\text{KL}}(\varrho \| \bigotimes_{K \in \alpha} \varrho_K) \} = D^{\text{KL}}(\varrho \| \bigotimes_{K \in \alpha} \varrho_K) = I_\alpha(\varrho).
\]

Now, applying this to a pure state \(\pi \in \mathcal{P} \subseteq \mathcal{D}\), since \(S(\pi) = 0\), we have that
\[
f_\alpha(\pi) = \sum_{K \in \alpha} f_K(\pi) \equiv \sum_{K \in \alpha} S(\pi_{K}) - \sum_{K \in \alpha} S(\pi_{0}) = D^{\text{KL}}(\pi \| \bigotimes_{K \in \alpha} \pi_K) = I_\alpha(\pi).
\]

That is, for pure states, the sum of the von Neumann entropies of disjoint subsystems given by the split \(\alpha = K_1|K_2|\ldots|K_{|\alpha|}\) is a meaningful quantity, and it characterizes the whole amount of correlation contained in the state \(\pi\) with respect to that split, being the mutual information above.

This reasoning enlightens also the meaning of entanglement itself, (what holds also in the bipartite case (70a)).

In classical probability theory, pure states are always uncorrelated, so if in the quantum case a pure state shows correlation, then this correlation is considered to be of quantum origin, and this correlation is defined to be the entanglement. From this point of view, the quantum versions of classical correlation measures applied to pure quantum states are pure entanglement measures, both in the bipartite and the multipartite scenario.

By this reasoning, let us define the \(\alpha\)-Entanglement Entropy, or simply \(\alpha\)-Entanglement \(E_\alpha : \mathcal{P} \to \mathbb{R}\) as
\[
E_\alpha(\pi) := \frac{1}{2} I_\alpha(\pi) = \frac{1}{2} \sum_{K \in \alpha} S(\pi_{K}),
\]
being the direct Level I. multipartite generalization of the Entanglement Entropy (70a). (Note that while the mutual information \(I_\alpha\) is defined for the whole state space \(\mathcal{D}\), \(E_\alpha\) is defined only for the pure states \(\mathcal{P}\), in accordance with (70a).)

As we can see from this reasoning, one can find other ways for forming the first kind hierarchy of entanglement measures, different from the one based on the \(K|\mathcal{K}\) measures: one can directly have \(f_\alpha(\pi) = D(\pi, \pi_{K_1} \otimes \pi_{K_2} \otimes \cdots \otimes \pi_{K_{|\alpha|}})\) with some distance or divergence \(D\). Here one has several choices again, one is the trace-distance, which has also a statistical meaning. Rényi, Tsallis generalizations of the Kullback-Leibler divergence could also be suitable.

Until this point, we have taken into consideration only the entanglement monotonicity (60) and the \(\alpha\)-discriminance (91). The multipartite monotonicity (94) is an additional concern, has to be satisfied too. For the measures (102) based on the von Neumann entropy (49a), the multipartite monotonicity (94) is a simple consequence of the subadditivity (54a) of the von Neumann entropy. The \(q > 1\) Tsallis entropies (49b) are also suitable (54b), however, Rényi entropies (49c) are not. Note that, since means equate things, using arithmetic mean (96b) instead of sum (96a) ruins the multipartite monotonicity for these cases.

C. Examples

Writing out some examples explicitly might not be useless here. Here we consider the \(\alpha\)-Entanglement Entropy (102), arising from the construction (101) using the von Neumann entropy (49a). Since the resulting functions are multipartite monotonic (94) indicator functions (91), we can read off these relations from the lattice \(P_1\), which can be seen for the cases \(n = 2\) and \(3\) in the upper-left part of Figures 1 and 2.

For the bipartite case, we get back the content of Section IV G,
\[
f_{1|2}(\pi) := E_{1|2}(\pi) = \frac{1}{2} (S(\pi_1) + S(\pi_2)) = S(\pi_{a}),
\]

(103a)
\[
f_{12}(\pi) := E_{12}(\pi) = \frac{1}{2} S(\pi_{12}) = 0. \quad (103b)
\]

Note that the multipartite monotonicity (94) holds, \(E_{1|2}(\pi) \geq E_{12}(\pi)\). We have also the discrimination (91),
\[
\pi \in P_{1|2} \iff E_{1|2}(\pi) = 0, \quad (104a)
\]
\[
\pi \in P_{12} \iff E_{12}(\pi) = 0. \quad (104b)
\]

For the \textit{tripartite} case,
\[
f_{1|2|3}(\pi) := E_{1|2|3}(\pi) = \frac{1}{2} (S(\pi_1) + S(\pi_2) + S(\pi_3)), \quad (105a)
\]
\[
f_{a|bc}(\pi) := E_{a|bc}(\pi) = \frac{1}{2} (S(\pi_a) + S(\pi_{bc})) = S(\pi_a), \quad (105b)
\]
\[
f_{123}(\pi) := E_{123}(\pi) = \frac{1}{2} S(\pi_{123}) = 0, \quad (105c)
\]

with all bipartitions \(a|bc\) of \(\{1, 2, 3\}\). Note that the multipartite monotonicity (94) holds, \(E_{1|2|3}(\pi) \geq E_{a|bc}(\pi) \geq E_{123}(\pi)\). We have also the discrimination (91),
\[
\pi \in P_{1|2|3} \iff E_{1|2|3}(\pi) = 0, \quad (106a)
\]
\[
\pi \in P_{a|bc} \iff E_{a|bc}(\pi) = 0, \quad (106b)
\]
\[
\pi \in P_{123} \iff E_{123}(\pi) = 0. \quad (106c)
\]

D. Level II: multipartite entanglement measures of the second kind

In Section III D, we have the \((P_{II}, \preceq)\) partial separability hierarchy of the second kind (27). Now, similarly to Section VI B, we consider the \(f_\alpha : P \to \mathbb{R}\) functions for all \(\alpha \in P_{II}\) labels of the second kind,
\[
P_{II, \ell} = \{ f_\alpha : P \to \mathbb{R} \mid \alpha \in P_{II} \}, \quad (107)
\]
and formulate their important properties expected for the measuring of the pure \(\alpha\)-entanglement. Entanglement monotonicity (60) is, of course, mandatory for all pure state measures. The others are as follows.

For the \(\alpha\) label of the second kind, the function \(f_\alpha : P \to \mathbb{R}\) is called \textit{pure \(\alpha\)-indicator function (or indicator function of the second kind with respect to \(P_\alpha\))}, if it is \textit{discriminant} with respect to \(P_\alpha\), (61b), that is, if it vanishes exactly for \(\alpha\)-separable pure states (28a),
\[
f_\alpha(\pi) = 0 \iff \pi \in P_\alpha. \quad (108)
\]

Using (28a), one can formulate the vanishing of the \(\alpha\)-indicator function \(f_\alpha\) by the vanishing of the \(\alpha\)-indicator functions \(f_\alpha\) of (91) as
\[
f_\alpha = 0 \iff \exists \alpha \in \alpha : f_\alpha = 0. \quad (109)
\]

From this and the inclusion hierarchy (30a), we immediately have that the indicator functions (108) obey
\[
\alpha \preceq \beta \implies (f_\alpha = 0 \implies f_\beta = 0). \quad (110)
\]

That is, a separability lower in the hierarchy implies a higher one, as it has to be. This connects the \(P_{II, \ell}\) set of functions to the partial separability hierarchy of the second kind \(P_{II}\) (27) and \(P_{II, P}\) (29a) in a sense. In addition to this, one can formulate a stronger property reflecting a stronger connection between these two structures, what provides \(P_{II, \ell}\) with a hierarchic structure. For the \(P_{II}\) labels of the second kind, the set of functions \(P_{II, \ell}\) is called \textit{multipartite-monotonic of the second kind, if}
\[
\alpha \preceq \beta \implies f_\alpha \geq f_\beta. \quad (111)
\]

(The map \(\alpha \mapsto f_\alpha\) is monotonically decreasing with respect to the labels of the second kind.) That is, entanglement higher in the hierarchy cannot be higher than entanglement lower in there. The multipartite monotonicity (111) is indeed stronger than the vanishing implications (110), since the latter one follows from the former one. The multipartite monotonicity, on the other hand, gives the same hierarchic structure to \(P_{II, \ell}\) as \(P_{II}\) and \(P_{II, P}\),
\[
(P_{II, \ell, \preceq}) \cong (P_{II, \succeq}). \quad (112)
\]

With the above definitions in hand, we construct a multipartite monotonic (111) hierarchy of entanglement measures for pure states for the hierarchy of the second kind, consisting of entanglement monotonic (60) \(\alpha\)-indicator functions (108). Let us start with the construction of \(\alpha\)-indicators, then check the monotonicity properties.

There are several ways of constructing \(\alpha\)-indicator functions (108), based on the \(\alpha\)-indicators (91). Perhaps the simplest one is a \textit{product},
\[
f_\alpha := \prod_{\alpha \in \alpha} f_\alpha. \quad (113a)
\]

Unfortunately, while it clearly obeys \(\alpha\)-discriminance (108) through (109), it lacks for entanglement monotonicity (60). This is because the set of functions obeying (60) is not closed under multiplication, what is related to the fact that the product of two concave functions is not concave in general. Moreover, a recent result of Eltschka et al. [54] suggests that homogeneous functions obeying (60) can not be of arbitrary high degree. (See Theorem I in [54], concerning a special class of functions.) This is an indication for using some power sums (72) or power means (73), since they do not change the degree. The geometric mean (73b),
\[
f_\alpha := \left[\prod_{\alpha \in \alpha} f_\alpha\right]^{1/|\alpha|} = M_0(f_{a_1}, \ldots, f_{a_{|\alpha|}}) \quad (113b)
\]
obeys \(\alpha\)-discriminance as the product (113a) does, and it turns out to be entanglement monotonic (60) [2, 3].

One can notice that we can use power sums (72) and power means (73) with general parameters \(q\),
\[
f_\alpha := N_q(f_{a_1}, \ldots, f_{a_{|\alpha|}}), \quad q < 0, \quad (113c)
\]
\[ f_\alpha := M_q(f_{\alpha_1}, \ldots, f_{\alpha_{|\alpha|}}), \quad q \leq 0. \quad (113d) \]

Indeed, power sums and power means are concave for \( q \leq 1 \), see (77b) and (78b), what is needed for the entanglement monotonicity (60), (see Corollary 4), while the proper vanishing properties (74b) and (75b) are satisfied for \( q < 0 \) and \( q \leq 0 \), what is needed for the \( \alpha \)-discriminance (108) through (109).

However, geometric means, or all \( q \neq 1 \) power-means of indicator functions of the first kind constructed from entropies in the way of Section VI B do not seem to make any sense in this situation. As we have learned in Section V A, using the power mean of entropies would infer an underlying “law” saying that the sum is the \( q \)-th power of the functions \( f_\alpha \) is meaningful, what seems to be true only for \( q = 1 \). We have two ways for getting out from this deadlock. The first one is using some transformed quantities for the indicator functions of the first kind, the second one is using the \(-\infty\)-mean (73d), that is, the minimum, what can make some sense.

To follow the first way, let us start with the \( \alpha \)-Entanglement (102), \( E_\alpha(\pi) = \frac{1}{2} \sum_{K \in \alpha} S(\pi_K) \), which is an “entropy-type” quantity. Only the sum of entropy-type quantities seems to be meaningful, however, a sum does not fulfill the \( \alpha \)-discriminance (109). Although, a product does fulfill the \( \alpha \)-discriminance, the product of entropy-type quantities is meaningful. A product which is meaningful is the product of “probability-type” quantities. And, indeed, in information theory (both classical [55], and quantum [13, 16]) entropy-type quantities appear often as arguments of \( e^{-x} \), leading to probability-type quantities. So, following this way, for the indicator functions of the first kind, we use the \( f_\alpha = g \circ E_\alpha \) transformed version of the mutual information, with the continuous invertible function \( g : \mathbb{R} \to \mathbb{R} \). Then we take the geometric mean of these indicator functions, preserving entanglement monotonicity and discriminance, and then do the transformation back, in order to get an entropy-type quantity again. This function \( g \) (i) should be the same for all \( \alpha \) (for simplicity), (ii) should map from non-negative to non-negative values (for being meaningful), (iii) should take zero at zero (for the \( \alpha \)-discriminance (109)), (iv) should be invertible, (v) should be monotonically increasing, (vi) should be concave (this is necessary but not sufficient for the entanglement monotonicity (60)). A particular function obeying these requirements is

\[ g(x) := 1 - e^{-x}, \quad (114) \]

being a perfect candidate for the conversion from entropy- to probability-type quantities. With this, let

\[
\begin{align*}
f_\alpha &:= g^{-1}\left(M_0(g(E_{\alpha_1}), \ldots, g(E_{\alpha_{|\alpha|}}))\right) \\
&= -\ln\left(1 - \prod_{\alpha \in \alpha} (1 - e^{-E_\alpha})^{1/|\alpha|}\right) \\
&= M_{\ln \circ g}(E_{\alpha_1}, \ldots, E_{\alpha_{|\alpha|}}), \quad (115)
\end{align*}
\]

also formulated by the quasi-arithmetic mean (82) for \( h = \ln q \). It is far from obvious that this function is an entanglement monotone (60). (For the proof, see Appendix D 2.) On the other hand, it is an \( \alpha \)-indicator (109). Its only drawback it that it is not hierarchy monotonic (111). Using product instead of the geometric mean in the construction would give hierarchy monotonicity, however, that would ruin entanglement monotonicity.

To follow the second way, which is the simpler one, take (113d) with \( q \to -\infty \) with the indicator functions

\[ f_\alpha(\pi) = I_\alpha(\pi) = \sum_{K \in \alpha} S(\pi_K) \text{ as in } (101), \text{ that is,} \]

\[ f_\alpha := M_{-\infty}(I_{\alpha_1}, \ldots, I_{\alpha_{|\alpha|}}) = \min(I_{\alpha_1}, \ldots, I_{\alpha_{|\alpha|}}). \quad (116) \]

And this also has a well-defined meaning. To clarify this, recall that the mutual information \( I_\alpha \) in (98) characterizes all the correlations in the sense of statistical distinguishability, that is, the distinguishability of the state from the closest (least distinguishable) uncorrelated state with respect to \( \alpha \), see in (100). Now the quantity \( \min(I_{\alpha_1}, \ldots, I_{\alpha_{|\alpha|}}) \) is distinguishability of the state from the closest (least distinguishable) uncorrelated state with respect to any \( \alpha \in \alpha \), and we define the Level II. version of the mutual information

\[
\begin{align*}
\min_{\alpha \in \alpha} \{I_\alpha(\varrho)\} &= \min_{\omega_k \in K \forall \omega_k \in \alpha} \left\{D^{K\alpha}_{\text{KL}}(\varrho \bigotimes \omega_K)\right\} \\
&= \min_{\alpha \in \alpha} \left\{\sum_{K \in \alpha} S(\varrho_K) - S(\varrho)\right\} =: I(\varrho),
\end{align*}
\]

which is also a geometric measure of correlation. Now, applying this to a pure state \( \pi \in \mathcal{P} \subseteq \mathcal{D} \), since \( S(\pi) = 0 \), we have that

\[
\begin{align*}
f_\alpha(\pi) &= \min_{\alpha \in \alpha} \{I_\alpha(\pi)\} = \min_{\omega_k \in K \forall \omega_k \in \alpha} \left\{D^{K\alpha}_{\text{KL}}(\varrho \bigotimes \omega_K)\right\} \\
&= \min_{\alpha \in \alpha} \left\{\sum_{K \in \alpha} S(\varrho_K)\right\} = I_\alpha(\pi).
\end{align*}
\]

That is, the minimal among the sums of the von Neumann entropies of disjoint subsystems given by the different splits \( \alpha \in \alpha \) is a meaningful quantity, characterizing the distinguishability of the state from the closest (least distinguishable) uncorrelated state with respect to any \( \alpha \in \alpha \).

By this reasoning, let us define the \( \alpha \)-Entanglement Entropy, or simply \( \alpha \)-Entanglement \( E_\alpha : \mathcal{P} \to \mathbb{R} \) as

\[ E_\alpha(\pi) := \min_{\alpha \in \alpha} \{E(\alpha, \pi)\}, \quad (119) \]

by the use of the \( \alpha \)-Entanglement (102). This is the Level II. multipartite generalization of the Entanglement Entropy (70a).

This function is an entanglement monotone (60) \( \alpha \)-indicator (109), moreover, it can easily be checked that it is also hierarchy monotonic (111). Note that, because the
\[ f_a = E_\alpha \] Level I. functions are multipartite monotonic (94), in the minimization during the calculation of the \( f_a = E_\alpha \) Level II. functions, it is enough to consider only the functions labelled by \( \max \alpha \),

\[ E_\alpha = \min_{\alpha \in \alpha} \{ E_\alpha \} = \min_{\alpha \in \max \alpha} \{ E_\alpha \}. \quad (120) \]

\section*{E. Examples}

Writing out some examples explicitly might not be useless here. Here we consider the \( \alpha \)-Entanglement Entropy (119), arising from the construction (118) based on the measures of the first kind (101) using the von Neumann entropy (49a). Since the resulting functions are multipartite monotonic (111) indicator functions (108), we can read off these relations from the lattice \( \mathcal{P}_1 \), which can be seen for the cases \( n = 2 \) and \( 3 \) in the upper-right part of Figures 1 and 2.

For the bipartite case, based on (103), we get back the content of Section IV G,

\[ f_{12}^1(\pi) := E_{12}^1(\pi) = \min \{ E_{1|2}(\pi), E_{12}(\pi) \} = E_{12}(\pi) = 0, \quad (121a) \]

\[ f_{12}^1(\pi) := E_{12}^1(\pi) = \min \{ E_{1|2}(\pi) \} = E_{12}(\pi) = S(\pi_2). \quad (121b) \]

Note that the multipartite monotonicity (111) holds, \( E_{12}^1(\pi) \geq E_{12}^1(\pi) \). We have also the discrimination (108),

\[ \pi \in \mathcal{P}_{12} \iff E_{12}(\pi) = 0, \quad (122a) \]
\[ \pi \in \mathcal{P}_{12} \iff E_{12}(\pi) = 0. \quad (122b) \]

For the tripartite case, based on (105),

\[ f_{123}(\pi) := E_{123}(\pi) = \min \{ E_{1|2|3}(\pi), E_{123}(\pi) \} = E_{123}(\pi) = 0, \quad (123a) \]

\[ f_{123}^{123}(\pi) := E_{123}^{123}(\pi) = \min \{ E_{1|2|3}(\pi), E_{1|2|3}(\pi), E_{1|2|3}(\pi) \} = E_{123}(\pi) = 0. \quad (123b) \]

\[ f_{123}^{123}(\pi) := E_{123}^{123}(\pi) = \min \{ E_{1|2|3}(\pi), E_{1|2|3}(\pi), E_{1|2|3}(\pi) \} = E_{123}(\pi) = 0. \quad (123c) \]

\[ f_{123}^{123}(\pi) := E_{123}^{123}(\pi) = \min \{ E_{1|2|3}(\pi), E_{1|2|3}(\pi), E_{1|2|3}(\pi) \} = E_{123}(\pi) = 0. \quad (123d) \]

\[ f_{1}^{123}(\pi) := E_{1|2|3}(\pi) = \min \{ E_{1|2|3}(\pi) \} = E_{123}(\pi) = 0. \quad (123e) \]

Note that the multipartite monotonicity (111) holds, \( E_{1|2|3}(\pi) \geq E_{1|2|3}(\pi) \geq E_{1|2|3}(\pi) \). We have also the discrimination (108),

\[ \pi \in \mathcal{P}_{123} \iff E_{123}(\pi) = 0, \quad (124a) \]
\[ \pi \in \mathcal{P}_{123} \iff E_{123}(\pi) = 0. \quad (124b) \]
\[ \pi \in \mathcal{P}_{123} \iff E_{123}(\pi) = 0. \quad (124c) \]
\[ \pi \in \mathcal{P}_{123} \iff E_{123}(\pi) = 0. \quad (124d) \]
\[ \pi \in \mathcal{P}_{123} \iff E_{123}(\pi) = 0. \quad (124e) \]

\section*{F. Multipartite entanglement measures for mixed states}

Now, it is easy to step from the pure states to mixed ones, thanks to the useful properties of the convex roof extension, listed in Section IV F.

If the function \( f_a \) is an entanglement monotone, that is, nonincreasing on average for pure states (60), (for example the \( \alpha \)-Entanglement Entropy in (119)), then, thanks to Theorem 2, its convex roof extension (64)

\[ f_a^\alpha(\rho) = \min_{\sum_n p_n \pi_n = \rho} \sum_n p_n f_a(\pi_n) \quad (125) \]

is also nonincreasing on average (59b) and also convex (59c) so it is an entanglement monotone.

If the function \( f_a \) is a pure \( \alpha \)-indicator (108), (for example the \( \alpha \)-Entanglement Entropy in (119)), then, thanks to (68), its convex roof extension (125) is a mixed \( \alpha \)-indicator,

\[ \rho \in \mathcal{D}_\alpha \iff f_a^\alpha(\rho) = 0. \quad (126) \]

If the set of functions \( \mathcal{P}_{11} \) is multipartite monotonic of the second kind (111), (for example the \( \alpha \)-Entanglement Entropy in (119)), then, thanks to (69), their convex roof extension (125) is also multipartite monotonic of the second kind (111).

By this reasoning, let us define the \( \alpha \)-Entanglement of Formation, as the convex roof extension of the \( \alpha \)-Entanglement Entropy (119) as

\[ E_{\alpha,OF} := E_{\alpha}^\cup. \quad (127) \]

This is the multipartite generalization of the Entanglement of Formation (71a). Note that in this case,

\[ E_{\alpha}^\cup = \left( \min_{\alpha \in \alpha} \{ E_\alpha \} \right)_{\cup} = \min_{\alpha \in \alpha} \{ E_\alpha^\cup \} \quad (128) \]
what is a consequence of (69).

G. Examples

Writing out some examples explicitly might not be useless here. Here we consider $\alpha$-Entanglement of Formation (127), which is the convex roof extension of the $\alpha$-Entanglement Entropy (119), arising from the construction (118) based on the measures of the first kind (101) using the von Neumann entropy (49a). Since the resulting functions are multipartite monotonicity (111) indicator functions (126), we can read off these relations from the lattice $P_1$, which can be seen for the cases $n = 2$ and 3 in the lower-left part of Figures 1 and 2.

For the bipartite case, based on (121), we get back the content of Section IV G,

\[
f^{\uparrow}(12) \langle \varrho \rangle := E^{\uparrow}(12) \langle \varrho \rangle = E^{\uparrow}(12) \langle \varrho \rangle = 0, \quad (129a)
\]

\[
f^{\uparrow}(1|2) \langle \varrho \rangle := E^{\uparrow}(1|2) \langle \varrho \rangle = E^{\uparrow}(1|2) \langle \varrho \rangle = E_{\alpha F}(\varrho). \quad (129b)
\]

Note that the multipartite monotonicity (111) holds, $E^{\uparrow}(12) \langle \varrho \rangle \leq E^{\uparrow}(1|2) \langle \varrho \rangle$. We have also the discrimination (126),

\[
\varrho \in D^{\uparrow}(12) \iff E^{\uparrow}(12) \langle \varrho \rangle = 0, \quad (130a)
\]

\[
\varrho \in D^{\uparrow}(1|2) \iff E^{\uparrow}(1|2) \langle \varrho \rangle = 0. \quad (130b)
\]

For the tripartite case, based on (123),

\[
f^{\uparrow}(123) \langle \varrho \rangle := E^{\uparrow}(123) \langle \varrho \rangle = E^{\uparrow}(123) \langle \varrho \rangle = 0, \quad (131a)
\]

\[
f^{\uparrow}(123) \langle \varrho \rangle := E^{\uparrow}(123) \langle \varrho \rangle = E^{\uparrow}(123) \langle \varrho \rangle = 0, \quad (131b)
\]

\[
f^{\uparrow}(1|2|3) \langle \varrho \rangle := I^{\uparrow}(1|2|3) \langle \varrho \rangle = E^{\uparrow}(1|2|3) \langle \varrho \rangle = 0. \quad (131c)
\]

Note that the multipartite monotonicity (111) holds, $E^{\uparrow}(123) \langle \pi \rangle \leq E^{\uparrow}(1|2|3) \langle \pi \rangle \leq E^{\uparrow}(1|2|3) \langle \pi \rangle \leq E^{\uparrow}(1|2|3) \langle \pi \rangle \leq E^{\uparrow}(1|2|3) \langle \pi \rangle$. We have also the discrimination (126),

\[
\varrho \in D^{\uparrow}(123) \iff E^{\uparrow}(123) \langle \varrho \rangle = 0, \quad (132a)
\]

\[
\varrho \in D^{\uparrow}(1|2|3) \iff E^{\uparrow}(1|2|3) \langle \varrho \rangle = 0. \quad (132b)
\]

\[
\varrho \in D^{\uparrow}(a|b|c) \iff E^{\uparrow}(a|b|c) \langle \varrho \rangle = 0. \quad (132c)
\]

\[
\varrho \in D^{\uparrow}(a|b|c) \iff E^{\uparrow}(a|b|c) \langle \varrho \rangle = 0. \quad (132d)
\]

\[
\varrho \in D^{\uparrow}(1|2|3) \iff E^{\uparrow}(1|2|3) \langle \varrho \rangle = 0. \quad (132e)
\]

H. Level III: detection of the classes

By the use of the mixed $\alpha$-indicators (126), one can detect also the classes (42),

\[
\varrho \in \mathcal{C}_{\{1|2\}} \iff f_{\alpha} \neq 0 \quad \forall \alpha \not\in \mathbf{\alpha}, \quad (133)
\]

which is a simple consequence of (126).

I. Examples

Writing out some examples explicitly might not be useless here. For the detection of the classes, here we consider $\alpha$-Entanglement of Formation (127), which is the convex roof extension of the $\alpha$-Entanglement Entropy (119), arising from the construction (118) based on the measures of the first kind (101) using the von Neumann entropy (49a).

For the bipartite case, we get back the content of Section IV G,

\[
\varrho \in \mathcal{C}_{\{1|2\}} = \mathcal{C}_{\text{sep}} \iff E^{\uparrow}(1|2) \langle \varrho \rangle = 0 \quad \text{and} \quad E^{\uparrow}(1|2) \langle \varrho \rangle = 0. \quad (134a)
\]

\[
\varrho \in \mathcal{C}_{\{1|2\}} = \mathcal{C}_{\text{ent}} \iff E^{\uparrow}(1|2) \langle \varrho \rangle \neq 0 \quad \text{and} \quad E^{\uparrow}(1|2) \langle \varrho \rangle = 0. \quad (134b)
\]

for the detection of the separable and entangled state classes.

For the tripartite case, the detection of the classes are shown in Table VII.

VII. SUMMARY

In this work, we have considered the entanglement classification and quantification problem for multipartite mixed states. We have worked out the hierarchical structure of different kinds of partial separability (Section III), which has turned out to be the down-set lattice of the partition of the subsystems (Section III D), and also the structure of the entanglement classes, which has turned out to be also hierarchical, being the up-set lattice of the lattice above (Section III F). Then we have constructed entanglement measures for multipartite quantum systems (Section VI). Besides the usual entanglement monotonicity and discrimination, we have introduced the multipartite monotonicity, as a plausible property, which endows the set of multipartite entanglement measures with the same hierarchical structure as the partial separability has. We have succeeded in constructing a hierarchy of entanglement measures satisfying these requirements (Section VII D), which are the direct generalizations of the Entanglement Entropy for pure states and the Entanglement of Formation for mixed states. These
measures have information-geometrical meaning, related to the statistical distinguishability.

There are wide-ranging possibilities for the generalization of the results, from what we can conclude that the entanglement monotonicity together with the discriminance property does not yield a condition too strong. The multipartite monotonicity, however, is more demanding.

ACKNOWLEDGMENTS

Discussions with Péter Vrana and István Kovács are gratefully acknowledged. This project was supported financially by the New Széchenyi Plan of Hungary (project ID: TÁMOP-4.2.2.B-10/12010-9), the Hungarian Scientific Research Fund (project ID: OTKA-K100908) and the “Lendület” program of the Hungarian Academy of Sciences. Last but not least, I thank my wife, Márta, for her everlasting love, patience and support :).

Appendix A: On the lattice structure of the classification

1. Posets and lattices: basics

Here we list some definitions and notations in Order Theory, following [10] and [56].

A partially ordered set, or simply poset, \((P, \preceq)\) is a set \(P\) endowed with a partial order \(\preceq\), being reflexive, antisymmetric and transitive, that is, for all \(x, y, z \in P\),

\[
\begin{align*}
x \preceq x, \\
x \preceq y \text{ and } y \preceq x & \implies x = y, \\
x \preceq y \text{ and } y \preceq z & \implies x \preceq z.
\end{align*}
\]

(A1a) \hspace{2cm} (A1b) \hspace{2cm} (A1c)

For the posets \((P, \preceq)\) and \((Q, \preceq)\), a map \(\phi : P \rightarrow Q\) is an order isomorphism when

\[
x \preceq y \iff \phi(x) \preceq \phi(y).
\]

(A2)

It follows easily from the reflexivity and antisymmetry of the partial order that such a map is bijective,

\[
\phi(x) = \phi(y) \iff \phi(x) \preceq \phi(y) \text{ and } \phi(y) \preceq \phi(x)
\]

\[
\iff x \preceq y \text{ and } y \preceq x
\]

\[
\iff x = y.
\]

(A3)

A poset \(P\) may have a bottom and a top element, denoted with \(\bot, \top \in P\), if

\[
\bot \preceq x \preceq \top \quad \forall x \in P.
\]

(A4)

(If the bottom and top exist, then they are unique ones, what is the consequence of the antisymmetry of the ordering.)

For a subset \(Q \subseteq P\) one can define its minimal and maximal elements as

\[
\min Q = \{x \in Q \mid (y \in Q \text{ and } y \preceq x) \implies y = x\},
\]

(A5a)

\[
\max Q = \{x \in Q \mid (y \in Q \text{ and } x \preceq y) \implies y = x\}.
\]

(A5b)

A subset \(Q \subseteq P\) is a down-set, or order ideal, if

\[
(x \in Q \text{ and } y \preceq x) \implies y \in Q
\]

(A6a)

(it is “closed downwards”). The set of all down-sets of \(P\) is denoted with \(C_\downarrow(P)\). Similarly, a subset \(Q \subseteq P\) is an up-set, or order filter, if

\[
(x \in Q \text{ and } x \preceq y) \implies y \in Q
\]

(A6b)
(it is “closed upwards”). The set of all up-sets of \( P \) is denoted with \( \mathcal{O}_1(P) \). For a subset \( Q \subseteq P \) one can define

\[
\updownarrow Q = \{ x \in P \mid \exists y \in Q : x \preceq y \}, \quad (A7a)
\]
\[
\uparrow Q = \{ x \in P \mid \exists y \in Q : y \preceq x \}, \quad (A7b)
\]

which are a down-set and an up-set, respectively.

The greatest lower bound or meet, \( a \land b \in P \), and the least upper bound or join, \( a \lor b \in P \) of the elements \( a \) and \( b \) in a poset are defined as

\[
(a \land b \preceq a, b) \quad \text{and} \quad (c \preceq a \land b \Rightarrow c \preceq a \lor b), \quad (A8a)
\]
\[
(a \lor b \preceq a, b) \quad \text{and} \quad (a, b \preceq c \Rightarrow a \land b \preceq c). \quad (A8b)
\]

A poset \( P \) is called a lattice if for all \( x, y \in P \) pairs, \( a \land b \) and \( a \lor b \) exist. A poset \( P \) is called a complete lattice if for all \( Q \subseteq P \) subsets, \( \bigwedge Q \) and \( \bigvee Q \) exist. Every finite lattice is complete. A finite lattice always have bottom and top elements. If only the meet or only the join can be defined, then the poset is called meet-semilattice or join-semilattice, respectively. A finite meet-semilattice always has a bottom element, a finite join-semilattice always has a top element.

Let \( P \) be a finite meet-semilattice having a top element \( \top \). Then the join can be defined as

\[
x \lor y = \bigvee \uparrow \{ x, y \}, \quad (A9a)
\]

so \( P \) is a lattice. (Proposition 3.3.1. in [56].) Dually, let \( P \) be a finite join-semilattice having a bottom element \( \bot \). Then the meet can be defined as

\[
x \land y = \bigwedge \downarrow \{ x, y \}, \quad (A9b)
\]

so \( P \) is a lattice.

Closing this section, we recall some examples [10, 56].

For a set \( X \), its power set \( 2^X = \{ A \subseteq X \} \) is a complete lattice with the ordering \( \subseteq \) (inclusion), the meet \( \cap \) (intersection) and the join \( \cup \) (union).

The \( \mathcal{O}_1(P) \) set of all down-sets (ideals) of a poset \( P \) is a lattice, called down-set lattice, with the ordering \( \subseteq \) (inclusion), the meet \( \cap \) (intersection) and the join \( \cup \) (union). Dually, the \( \mathcal{O}_1(P) \) set of all up-sets (filters) of a poset \( P \) is a lattice, called up-set lattice, with the ordering \( \subseteq \) (inclusion), the meet \( \cap \) (intersection) and the join \( \cup \) (union).

### 2. Construction by bipartitions

For the proof of (11), we need by (A8a) that (i) \( \alpha \leq K | K \) for all \( K \in \alpha \), and (ii) if \( \beta \leq K | K \) for all \( K \in \alpha \), then \( \beta \leq \alpha \). For (i), we have that \( \alpha = K_1 \cdot K_2 \cdots \cdot K_{|\alpha|} \leq K_1 \cdot (\bigcup_{j \neq i} K_i) = K_i \), since every \( K_j \) is contained either in \( K_i \) or in \( K_1 \) (the \( K_i \) sets are disjoint ones) and (8) holds. For (ii), let \( \beta = K'_1 \cdot K'_2 \cdots \cdot K'|_{|\beta|} \leq K | K \) for all \( K \in \alpha \), then, by definition (8), for all \( K \in \alpha \), we have

\[
\pi_K = \bigotimes_{K \in \alpha} \pi_K \quad \text{and} \quad \pi_{K'} = \bigotimes_{K' \in \alpha'} \pi_{K'}.
\]

Thus, we can rewrite (11) as

\[
\pi_K = \bigotimes_{K \in \alpha} \pi_K \quad \text{and} \quad \pi_{K'} = \bigotimes_{K' \in \alpha'} \pi_{K'}.
\]

### 3. Order isomorphisms of the first kind

Here we prove (14a) and (14b).

For the proof of (14a), using the definition (12a), we can reformulate \( \mathcal{P}_\beta \subseteq \mathcal{P}_\alpha \) as follows

\[
\pi \in \mathcal{P}_\beta \quad \overset{(12a)}{\Rightarrow} \quad \forall K' \in \beta, \exists \pi_K' \in \mathcal{P}_{K'} : \pi = \bigotimes_{K' \in \beta} \pi_{K'} \quad \overset{(12a)}{\Rightarrow} \quad \forall K \in \alpha, \exists \pi_K \in \mathcal{P}_K : \pi = \bigotimes_{K \in \alpha} \pi_K
\]

To see the \( \Rightarrow \) implication in (14a), note that if \( \beta \preceq \alpha \), then, by definition (8), one can collect every \( K' \in \beta \) for which \( K' \subseteq K \), and construct \( \pi_K = \bigotimes_{K' \subseteq K \in \beta} \pi_{K'} \). This can be done for all \( K \in \alpha \), leading to the implication above. To see the \( \Leftarrow \) implication in (14a), note that for the contrapositive statement, we have

\[
\beta \npreceq \alpha \quad \overset{\text{def.}}{\iff} \quad \exists K' \in \beta, \forall K \in \alpha : K' \nsubseteq K \quad (A10) \quad \overset{(8)}{\Rightarrow} \quad \forall K \in \alpha, \exists \pi_K \in \mathcal{P}_K : \pi = \bigotimes_{K' \subseteq K \in \beta} \pi_{K'}
\]

in general, by (A10), so \( \pi \notin \mathcal{P}_\alpha \), so \( \mathcal{P}_\beta \nsubseteq \mathcal{P}_\alpha \).

For (14b), we have (14a), and we claim

\[
\mathcal{P}_\beta \subseteq \mathcal{P}_\alpha \iff \mathcal{D}_\beta \subseteq \mathcal{D}_\alpha,
\]

what comes from the geometry of quantum states. The \( \Rightarrow \) implication is obvious from (12b), while for the \( \Leftarrow \) implication one has \( \mathcal{P}_\beta = \text{Extr} \mathcal{D}_\beta \subseteq \mathcal{D}_\beta \subseteq \mathcal{D}_\alpha \), so any \( \pi \in \mathcal{P}_\beta \) is an element of \( \mathcal{D}_\alpha \) moreover, it is a pure state, so it is an element also of \( \text{Extr} \mathcal{D}_\alpha = \mathcal{P}_\alpha \) (12c), what is exactly what we need.

It is general for posets that if \( \alpha \mapsto \mathcal{P}_\alpha \) is an order isomorphism as in (14a) then the two posets are isomorphic, that is, in our case, the map \( \alpha \mapsto \mathcal{P}_\alpha \) is bijective, see in Appendix A 1. The same holds for \( \alpha \mapsto \mathcal{D}_\alpha \) based on (14b).

### 4. Meet-semilattice isomorphism of the first kind for pure states

For the proof of (16), let \( \pi \in \mathcal{P}_\alpha \cap \mathcal{P}_{\alpha'} \), where \( \alpha = K_1 | K_2 | \ldots | K_{|\alpha|} \) and \( \alpha' = K'_1 | K'_2 | \ldots | K'_{|\alpha'|} \), then from definition (12a) we have

\[
\pi = \bigotimes_{K \in \alpha} \pi_K = \bigotimes_{K' \in \alpha'} \pi_{K'}
\]

with \( \pi_K \in \mathcal{P}_K, \pi_{K'} \in \mathcal{P}_{K'} \). For all \( K \in \alpha \), we have

\[
\pi_K = \bigotimes_{K' \in \alpha'} \pi_{K'}
\]

for all \( K' \in \alpha \).
leading to the decomposition

\[ \pi = \bigotimes_{K \in \alpha} \pi_K = \bigotimes_{K \in \alpha} \bigotimes_{K' \in \alpha'} \pi'_{K \cap K'}, \]

that is, \( \pi \) is separable with respect to the split \{ \( K \cap K' \neq \emptyset \mid K \in \alpha, K' \in \alpha' \} \), what is just \( \alpha \wedge \alpha' \) by (9a), so we have \( \mathcal{P}_\alpha \cap \mathcal{P}_{\alpha'} \subseteq \mathcal{P}_{\alpha \wedge \alpha'} \). The reverse inclusion is the first one in (15a), what completes the proof.

5. Order isomorphisms of the second kind

Here we prove (30a) and (30b).

For (30a), using the definitions (27) and (28a), we have

\[ \beta \preceq \alpha \quad \Leftrightarrow \quad \{ \mathcal{P}_\beta \mid \beta \in \beta \} \subseteq \{ \mathcal{P}_\alpha \mid \alpha \in \alpha \} \]
\[ \Leftrightarrow \quad \bigcup_{\beta \in \beta} \mathcal{P}_\beta \subseteq \bigcup_{\alpha \in \alpha} \mathcal{P}_\alpha \]
\[ \overset{(28a)}{\Leftrightarrow} \quad \mathcal{P}_\beta \subseteq \mathcal{P}_\alpha, \]

where the second implication is because \( \alpha \mapsto \mathcal{P}_\alpha \) is bijective due to (14a), while the third one is obvious.

For (30b), we have (30a), and we claim

\[ \mathcal{P}_\beta \subseteq \mathcal{P}_\alpha \quad \Leftrightarrow \quad \mathcal{D}_\beta \subseteq \mathcal{D}_\alpha, \]

what can be proven in the same way as the parallel result of the first kind in Appendix A 3.

Again, it is general for posets that if \( \alpha \mapsto \mathcal{P}_\alpha \) is an order isomorphism as in (30a) then the two posets are isomorphic, that is, in our case, the map \( \alpha \mapsto \mathcal{P}_\alpha \) is bijective, see in Appendix A 1. The same holds for \( \alpha \mapsto \mathcal{D}_\alpha \) based on (30b).

6. Lattice isomorphism of the second kind for pure states

Here we prove (32). For the first part,

\[ \mathcal{P}_\alpha \cap \mathcal{P}_{\alpha'} = \bigcup_{\alpha \in \alpha} \mathcal{P}_\alpha \cap \bigcup_{\alpha' \in \alpha'} \mathcal{P}_{\alpha'} = \bigcup_{\alpha \in \alpha} \bigcup_{\alpha' \in \alpha'} \mathcal{P}_{\alpha \wedge \alpha'} \]
\[ \subseteq \bigcup_{\beta \in \alpha \wedge \alpha'} \mathcal{P}_\beta = \bigcup_{\beta \in \alpha \wedge \alpha'} \mathcal{P}_\beta = \mathcal{P}_{\alpha \wedge \alpha'}, \]

where the first and last equations are by definition (28a), the last but one equation is by definition (27), the second equation is the distributivity of \( \cap \) over \( \cup \), the third equation is (16) from Level I., The inclusion is from that for all \( \alpha \in \alpha \) and \( \alpha' \in \alpha' \), \( \alpha \wedge \alpha' \leq \alpha \in \alpha \) and \( \alpha \wedge \alpha' \leq \alpha' \in \alpha' \) hold, because \( \alpha \) and \( \alpha' \) are down-sets (26), so \( \alpha \wedge \alpha' \in \alpha \wedge \alpha' \), so all \( \mathcal{P}_{\alpha \wedge \alpha'} \) sets in the lefthand side appear in the righthand side as a \( \mathcal{P}_\beta \). The reverse inclusion is the first one in (31a), what completes the proof. For the second part, we have

\[ \mathcal{P}_\alpha \cup \mathcal{P}_{\alpha'} = \bigcup_{\alpha \in \alpha} \mathcal{P}_\alpha \cup \bigcup_{\alpha' \in \alpha'} \mathcal{P}_{\alpha'} = \bigcup_{\alpha \in \alpha} \mathcal{P}_\alpha = \mathcal{P}_{\alpha \wedge \alpha'}, \]

where the first and third equations are by definition (28a), the second one is obvious using elementary set algebra.

7. Lattice of the labels of the classes

For the proof of (46), we start with the contrapositive form of (43),

\[ C_{\alpha} \neq \emptyset \quad \Rightarrow \quad \forall \alpha \in \alpha, \forall \beta \notin \alpha : \alpha \nmid \beta \]
\[ \Leftrightarrow \quad \forall \alpha \in \alpha : (\beta \notin \alpha \Rightarrow \alpha \nmid \beta) \]
\[ \Leftrightarrow \quad \forall \alpha \in \alpha : (\alpha \leq \beta \Rightarrow \beta \in \alpha), \]

where the last implication is the contrapositive reformulation of the parenthesis, leading just to the definition of the up-set lattice \( P_{\Pi^1} \) in (44).

Appendix B: On entanglement measures

1. Convexity/concavity of operator functions

Here we recall a useful result in the theory of trace functions from Section 2.2 of [33].

Let \( f : \mathbb{R} \to \mathbb{R} \) be continuous, and \( F = \text{tr} \circ f : \text{Lin}_{\mathbb{R}} \mathcal{H} \to \mathbb{R} \) the associated trace function, then if \( f \) is monotonically increasing (decreasing) then \( F \) is monotonically increasing (decreasing) and if \( f \) is convex (concave) then \( F \) is convex (concave).

Based on these, the (52a) and (52b) concavity of the von Neumann (49a) and Tsallis (49b) entropies can be proven using the functions \( f(x) = -x \ln x \) and \( f(x) = \frac{1}{1-q}(x^q - x) \). For the (52c) concavity of the Rényi entropy (49c), we have that \( f(x) = x^q \) is concave if and only if \( q \leq 1 \), then so is \( F(g) = \text{tr} g^q \), while \( \ln(x) \) is concave and monotonically increasing, so \( \ln \text{tr} g^q \) is also concave. However, for \( q \geq 1 \), \( F(g) = \text{tr} g^q \) is convex, while \( -\ln(x) \) is convex and monotonically decreasing, so the concavity/concavity cannot be decided by this.

2. Pure entanglement measures

Here we recall Horodecki’s proof [29] for Theorem 1 given in Section IV.

A function is symmetric in its arguments if it does not change its value for the permutation of its arguments; and it is expansible if it takes the same values for arguments
\( (x_1, \ldots, x_d) \) and \( (x_1, \ldots, x_d, 0) \). So, first of all, from (i) it follows that
\[
F(\text{tr}_K \pi) = F(\text{tr}_K \pi), \tag{B1}
\]
since the two arguments have the same spectrum, apart from the multiplicity of the zero eigenvalue. Let us decompose the LOCC \( \Lambda \) into the pure operations \( \Lambda_i \) consisting of single Kraus operators each, \( \Lambda_i(\cdot) = A_i(\cdot)A_i^\dagger \). These operations are separable ones, that is, \( A_i = A_{1,i} \otimes A_{2,i} \otimes \cdots \otimes A_{n,i} \), and they can further be decomposed into the composition of \( \Lambda_{a,i}(\cdot) \) acting nontrivially on the \( a \)th subsystem only. Applying these to the initial (pure) state \( \pi \) results in the ensemble of pure states \( \pi_i' = \frac{1}{p_i} \Lambda_{a,i}(\pi) \) (with probabilities \( p_i = \text{tr} \Lambda_{a,i}(\pi) \)). The resulting mixed states are then \( \pi' = \sum_i p_i \pi_i' \).

Take the \( \Lambda_{a,i} \)'s acting on subsystems \( a \) not contained in \( K \). These operations leave subsystems \( K \) invariant, \( \text{tr}_K \pi' = \text{tr}_K \pi \), which leads to
\[
\text{tr}_K \pi = \text{tr}_K \pi' = \sum_i p_i \text{tr}_K \pi_i' \quad \text{for} \ a \notin K. \tag{B2}
\]

Now we can write
\[
f_K(\pi) = F(\text{tr}_K \pi) = F\left( \sum_i p_i \text{tr}_K \pi_i' \right) \\
\geq \sum_i p_i F(\text{tr}_K \pi_i') = \sum_i p_i f_K(\pi_i'),
\]
where the first and last equalities are by construction (63), the second one is (B2), and the inequality is the concavity (ii) of Theorem 1.

For handling the \( \Lambda_{a,i} \)'s acting on subsystems \( a \) contained in \( K \), we use that these operations leave subsystem \( K \) invariant, \( \text{tr}_K \pi = \text{tr}_K \pi' \), leading to
\[
\text{tr}_K \pi = \text{tr}_K \pi' = \sum_i p_i \text{tr}_K \pi_i' \quad \text{for} \ a \in K. \tag{B3}
\]

Now, we can repeat the previous reasoning for subsystem \( K \), taking advantage of (B1)
\[
f_K(\pi) = F(\text{tr}_K \pi) = F(\text{tr}_K \pi') = F\left( \sum_i p_i \text{tr}_K \pi_i' \right) \\
\geq \sum_i p_i F(\text{tr}_K \pi_i') = \sum_i p_i F(\text{tr}_K \pi_i') = \sum_i p_i f_K(\pi_i').
\]
This completes the proof.

3. Convex roofs of pure entanglement monotones

Here we recall Horodecki’s proof [29] for Theorem 2 given in Section IV.

Let \( f : \mathcal{P} \to \mathbb{R} \) a function satisfying (60), that is,
\[
\sum_i p_i f(\pi_i') \leq f(\pi) \tag{B4}
\]
for all \( \pi \mapsto \{ (p_i, \pi_i') \} \) ensembles resulting from a LOCC \( \Lambda \) consisting of the \( \Lambda_i \) pure operations. Take an \( f \)-optimal pure decomposition \( \{(q_j, \pi_j)\} \) of \( \rho \), that is, \( \rho = \sum_j q_j \pi_j \) for which the argument of the minimization in the righthandside of (64) takes its minimum, so
\[
f^\ast(\rho) = \sum_j q_j f(\pi_j). \tag{B5}
\]
Applying the \( \Lambda_i \) pure operators to the pure states \( \pi_j \) of this ensemble results in the ensembles of pure states \( \pi_{ji} = \frac{1}{p_{ji}} \Lambda_i(\pi_j) \) (with probabilities \( p_{ji} = \text{tr} \Lambda_i(\pi_j) \)). Applying the \( \Lambda_i \) pure operators to the mixed state \( \rho \) results in
\[
\rho_i' = \frac{1}{p_i} \Lambda_i(\rho) = \frac{1}{p_i} \sum_j q_j \Lambda_i(\pi_j) = \frac{1}{p_i} \sum_j q_j p_{ji} \pi_{ji}' \tag{B6}
\]
(with probability \( p_i = \text{tr} \Lambda_i(\rho) = \sum_j q_j \text{tr} \Lambda_i(\pi_j) = \sum_j q_j p_{ji} \)). With these, we can write
\[
f^\ast(\rho) = \sum_j q_j f(\pi_j) \\
\geq \sum_j q_j \sum_i p_{ji} f(\pi_{ji}') \\
= \sum_j q_j \sum_i p_{ji} f^\ast(\pi_{ji}') \\
= \sum_i p_i \left( \frac{1}{p_i} \sum_j q_j p_{ji} f^\ast(\pi_{ji}') \right) \\
\geq \sum_i p_i f^\ast \left( \sum_j q_j p_{ji} \pi_{ji}' \right) = \sum_i p_i f^\ast(\rho_i'),
\]
where the first equality is the optimality (B5), the first inequality is due to the entanglement monotonicity of \( f \) on pure states (B4), the second equality is (65), and the second inequality is the convexity of the convex roof extension (66), and the last equality is (B6).

Note that this reasoning does not depend on wether \( \{(q_j, \pi_j)\} \) is a result of a LOCC or some other class of operations, as far as (B4) holds.

4. Convex roof extension preserves discrimination

For the proof of (68), let \( \mathcal{P}_\ast \subseteq \mathcal{P} \), and \( \mathcal{D}_\ast = \text{Conv} \mathcal{P}_\ast \subseteq \mathcal{D} = \text{Conv} \mathcal{P} \), and let \( f : \mathcal{P} \to [0, \infty) \), then
\[
\rho \in \mathcal{D}_\ast \quad \iff \quad \rho \in \text{Conv} \mathcal{P}_\ast \\
\iff \quad \rho = \sum_i p_i \pi_i \text{ with } \pi_i \in \mathcal{P}_\ast \\
\iff \quad \rho = \sum_i p_i \pi_i \text{ with } f(\pi_i) = 0 \\
\iff \quad f^\ast(\rho) = 0,
\]
where the first and second implications are from the assumptions above, the third implication is the discrimination property (61a) for the pure state function. The
condition \( f \geq 0 \) is necessary for the last implication: to see the \( \Rightarrow \) direction, note that the minimum in the convex roof extension (64) of a nonnegative function is zero, which is attained if the lefthandside holds, and to see the \( \Leftarrow \) direction, note that if the convex roof extension (64) of a nonnegative function vanishes then there exists a decomposition for pure states for which the function vanishes.

5. Convex roof extension is monotonic

For the proof of (69), note that the \( \Leftarrow \) direction is obvious from (65), while for the \( \Rightarrow \) direction take a \( g \)-optimal decomposition \( \{ (p_i, \pi_i) \} \) of \( g \), that is, \( g = \sum_i p_i \pi_i \) for which the next equality holds, and

\[
g^{ij}(q) = \sum_i p_i g(\pi_i) \geq \sum_i p_i f(\pi_i) \geq \min_{\pi'_i, \pi''_i, \pi_i = \pi} \sum_i p'_i f(\pi'_i) = f^{ij}(q),
\]

where the first inequality is the lefthandside in (69), the second one is because a \( g \)-optimal decomposition is not necessary \( f \)-optimal, and the last equality is the definition (64) of convex roof extension.

Appendix C: On the properties of power-sums and power-means

1. Monotonicity

For the monotonicity of the \( q \)-sums (72a) and \( q \)-means (73a) for \( x > 0 \), we have for the latter one for \( q \neq 0 \) that the Hessian from (C1)

\[
\frac{\partial^2 M_q(x)}{\partial x_j \partial x_i} = \frac{1}{m^{1/q}} \left( \sum_k x_k^{q} \right)^{1/q-1} x_i^{q-1} x_j^{-1} \\
+ \frac{1}{m^{1/q}} (q-1) \left( \sum_k x_k^{q} \right)^{1/q-1} \delta_{ij} x_i^{-2}.
\]  

Then, taking an \( u \in \mathbb{R}^m \), we are interested in the sign of

\[
\sum_j u_j \frac{\partial^2 M_q(x)}{\partial x_j \partial x_i} u_i = \frac{1}{m^{1/q}} (q-1) \left( \sum_k x_k^{q} \right)^{1/q-2} \\
\times \left[ \left( \sum_i x_i^{q} \right) \left( \sum_j x_j^{q-2} \right) - \left( \sum_i x_i^{q-1} \right)^2 \right]
\]

\[
= \frac{1}{m^{1/q}} (q-1) \left( \sum_k x_k^{q} \right)^{1/q-2} \\
\times \left[ \sum_i (x_i^{q/2})^2 \sum_j (x_j^{q/2-1})^2 - \left( \sum_i (x_i^{q/2-1}) \right)^2 \right].
\]  

The square bracket \([ \ ]\) is nonnegative, which is the Cauchy-Bunyakovsky-Schwarz inequality for the vectors of components \( 1 \) and \( x_i^{-1} \). So we have that \( M_q \) convex for \( q \geq 1 \) and \( M_q \) is concave for \( q \leq 1 \). The same holds for \( N_q \).

For \( q = 0 \), the Hessian from (C2)

\[
\frac{\partial^2 M_0(x)}{\partial x_j \partial x_i} = \frac{1}{m^2} \left( \prod_k x_k \right)^{1/m} x_i^{-1} x_j^{-1} \\
- \frac{1}{m} \left( \prod_k x_k \right)^{1/m} \delta_{ij} x_i^{-2}
\]  

Then, taking an \( u \in \mathbb{R}^m \), we are interested in the sign of

\[
\sum_j u_j \frac{\partial^2 M_0(x)}{\partial x_j \partial x_i} u_i = \frac{1}{m^2} \left( \prod_k x_k \right)^{1/m} \\
\times \left[ \left( \sum_i u_i x_i^{-1} \right) \left( \sum_j u_j x_j^{-1} \right) - m \sum_i u_i^2 x_i^{-2} \right]
\]

\[
= \frac{1}{m^2} \left( \prod_k x_k \right)^{1/m} \\
\times \left[ \left( \sum_i (u_i x_i^{-1}) \right)^2 - \left( \sum_i (1)^2 \right) \left( \sum_i (u_i x_i^{-1})^2 \right) \right].
\]  

The square bracket \([ \ ]\) is nonpositive, which is the Cauchy-Bunyakovsky-Schwarz inequality for the vectors of components \( 1 \) and \( u_i x_i^{-1} \), so we have that \( M_0 \) is concave.
Appendix D: On constructions of entanglement monotonies

1. By concavity

For the proof of Lemma 3, let \( A = \sum_i A_i \) a LOCC with the decomposition into pure suboperations \( A_i \), and \( \pi'_i = \frac{1}{p_i} A_i(\pi) \) with probability \( p_i = \text{tr} A_i(\pi) \). Then we can write

\[
\sum_i p_i G(f_1, \ldots, f_m)(\pi'_i) = \sum_i p_i G(f_1(\pi'_i), \ldots, f_m(\pi'_i)) \\
\leq G(\sum_i p_i f_1(\pi'_i), \ldots, \sum_i p_i f_m(\pi'_i)) \\
\leq G(f_1, \ldots, f_m)(\pi) = G(f_1, \ldots, f_m)(\pi)
\]

where the first inequality is the concavity of \( G \), the second inequality is the assumption (60) together with the monotonicity \( G \).

Then the Hessian is

\[
\frac{\partial^2 M_{\text{in} \circ \rho}}{\partial x_j \partial x_i} = \left[ \frac{\partial}{\partial x_j} \frac{1}{1 - \Pi^{1/m}} \right] \frac{1}{m} \Pi^{1/m} \frac{e^{-x_i}}{1 - e^{-x_i}} + \frac{1}{1 - \Pi^{1/m}} \frac{1}{m} \left[ \frac{\partial}{\partial x_j} \Pi^{1/m} \right] \frac{e^{-x_i}}{1 - e^{-x_i}} + \frac{1}{1 - \Pi^{1/m}} \frac{1}{m} \Pi^{1/m} \left[ \frac{\partial}{\partial x_j} \frac{e^{-x_i}}{1 - e^{-x_i}} \right]
\]

\[
= \frac{1}{m^2} \frac{\Pi^{1/m}}{(1 - \Pi^{1/m})^2} \left[ \frac{\Pi^{1/m}}{1 - e^{-x_i}} \frac{e^{-x_i}}{1 - e^{-x_i}} + (1 - \Pi^{1/m}) \frac{e^{-x_i}}{1 - e^{-x_i}} \right] m(1 - \Pi^{1/m}) \frac{\partial}{\partial x_j} \frac{e^{-x_i}}{1 - e^{-x_i}}
\]

\[
= \frac{1}{m^2} \frac{\Pi^{1/m}}{(1 - \Pi^{1/m})^2} \left[ \frac{e^{-x_i}}{1 - e^{-x_i}} - m(1 - \Pi^{1/m}) \frac{\partial}{\partial x_j} \frac{e^{-x_i}}{1 - e^{-x_i}} \right].
\]

Then, taking \( u \in \mathbb{R}^m \), we are interested in the sign of

\[
\sum_{ij} u_j \frac{\partial^2 M_{\text{in} \circ \rho}}{\partial x_j \partial x_i} u_i = \frac{1}{m^2} \frac{\Pi^{1/m}}{(1 - \Pi^{1/m})^2} \left[ \left( \sum_i u_i e^{-x_i} \right)^2 - m(1 - \Pi^{1/m}) \sum_i u_i^2 e^{-x_i} \right].
\]

We claim that the square bracket \([\cdot]\) is nonpositive, what follows from

\[
\left( \sum_i \frac{u_i e^{-x_i}}{1 - e^{-x_i}} \right)^2 \leq \sum_i \left( \frac{u_i \sqrt{e^{-x_i}}}{1 - e^{-x_i}} \right)^2 \sum_j \left( \frac{\sqrt{e^{-x_j}}}{1 - e^{-x_i}} \right)^2 \leq \sum_i \left( \frac{u_i \sqrt{e^{-x_i}}}{1 - e^{-x_i}} \right)^2 m(1 - \Pi^{1/m})
\]

where the first inequality is the Cauchy-Bunyakovsky-Schwarz inequality for the vectors of components \( \frac{u_i \sqrt{e^{-x_i}}}{1 - e^{-x_i}} \) and
\[ \sqrt{e^{-x_j}}, \text{ and the second inequality is } \sum_j e^{-x_j} \leq m(1 - \Pi^{1/m}), \text{ which is rearranged as } \Pi^{1/m} \leq 1 - \frac{1}{m} \sum_j e^{-x_j}, \]

\[ \left( \prod_j (1 - e^{-x_j}) \right)^{1/m} = \Pi^{1/m} \leq 1 - \frac{1}{m} \sum_j e^{-x_j} = \frac{1}{m} \sum_j (1 - e^{-x_j}), \] (D8)

which is the inequality between the geometric and arithmetic means (see equation (80) for \( q = 0 \) and \( q' = 1 \)). So we can conclude that the function \( M_{\ln og} \) is concave. Now, \( M_{\ln og} \) is an entanglement monotone, since Lemma 3 holds for that, what completes the proof.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. 81, 865 (2009).
[2] S. Szalay, Quantum entanglement in finite-dimensional Hilbert spaces, PhD dissertation (2013), http://arxiv.org/pdf/1302.4654.
[3] S. Szalay and Z. Kökényesi, Phys. Rev. A 86, 032341 (2012).
[4] E. Schmidt, Mathematische Annalen 63, 433 (1907).
[5] A. Uhlmann, Entropy 12, 1799 (2010).
[6] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A 54, 3824 (1996).
[7] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach, Phys. Rev. Lett. 85, 1560 (2000).
[8] G. Vidal, Phys. Rev. Lett. 91, 147902 (2003).
[9] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza, arXiv:physics.chem-ph , arXiv:1412.5829 (2015), http://arxiv.org/pdf/1412.5829.
[10] B. A. Davey and H. A. Priestley, Introduction to Lattices and Order, 2nd ed. (Cambridge University Press, 2002).
[11] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press, New York, NY, USA, 2006).
[12] S. Boyd and L. Vandenberghe, Convex Optimization (Cambridge University Press, 2004).
[13] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, 1st ed. (Cambridge University Press, 2000).
[14] J. Preskill, “Lecture notes on quantum computation,” Available on the homepage of the author: http://www.theory.caltech.edu/people/preskill/ph229.
[15] D. Petz, Quantum Information Theory and Quantum Statistics (Springer, 2008).
[16] M. M. Wilde, Quantum Information Theory (Cambridge University Press, New York, NY, USA, 2013).
[17] M. M. Wolf, “Quantum channels & operations guided tour,” (2012), unpublished manuscript, http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QuantumChannels.pdf.
[18] J. von Neumann, Mathematical Foundations of Quantum Mechanics (Princeton University Press, 1996).
[19] R. F. Werner, Phys. Rev. A 40, 4277 (1989).
[20] E. Chitambar, D. Leung, L. Mancinska, M. Ozols, and A. Winter, arXiv [quant-ph] , 1210.4583 (2012).
[21] S. Szalay, Phys. Rev. A 83, 062337 (2011).
[22] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935).
[23] E. Schrödinger, Naturwissenschaften 23, 807 (1935).
[24] E. Schrödinger, Math. Proc. Camb. Phil. Soc. 31, 555 (1935).
[25] “The On-Line Encyclopedia of Integer Sequences, A000110,” Bell or exponential numbers: ways of placing n labeled balls into n indistinguishable boxes.
[26] M. Seevinck and J. Uffink, Phys. Rev. A 78, 032101 (2008).
[27] W. Dür, J. I. Cirac, and R. Tarrach, Phys. Rev. Lett. 83, 3562 (1999).
[28] W. Dür and J. I. Cirac, Phys. Rev. A 61, 042314 (2000).
[29] M. Horodecki, Quant. Inf. Comp. 1, 3 (2001).
[30] J. v. Neumann, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse 1927, 273 (1927).
[31] M. Ohya and D. Petz, Quantum Entropy and Its Use, 1st ed. (Springer Verlag, 1993).
[32] A. Rényi, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics (University of California Press, Berkeley, Calif., 1961) pp. 547–561.
[33] E. A. Carlen, in Entropy and the Quantum, Contemporary Mathematics, Vol. 529, edited by R. Sims and D. Ueltschi (American Mathematical Society, 2010) pp. 73–140.
[34] G. A. Raggio, J. Math. Phys. 36, 4785 (1995).
[35] K. M. R. Audenaert, J. Math. Phys. 48, 083507 (2007).
[36] W. van Dam and P. Hayden, arXiv:quant-ph , arXiv:quant-ph (2002), http://arxiv.org/pdf/quant-ph/0204093.
[37] H. Umegaki, Kodai Math. Sem. Rep. 14, 59 (1962).
[38] T. Sagawa, arXiv [cond-mat] , 1202.0983 (2012), http://arxiv.org/pdf/1202.0983.
[39] M. B. Plenio and S. Virmani, Quant. Inf. Comp. 7, 1 (2007).
[40] G. Vidal, J. Mod. Opt 47, 355 (2000).
[41] A. Uhlmann, Phys. Rev. A 62, 032307 (2000).
[42] B. Röthlisberger, J. Lehmann, and D. Loss, Phys. Rev. A 80, 042301 (2009).
[43] B. Röthlisberger, J. Lehmann, and D. Loss, arXiv [quant-ph] , 1107.4497 (2011), http://arxiv.org/pdf/1107.4497.
[44] E. Schrödinger, Math. Proc. Camb. Phil. Soc. 32, 446 (1936).
[45] N. Gisin, Helv. Phys. Acta 62, 363 (1989).
[46] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A 183, 14 (1993).
[47] A. Uhlmann, Open Sys. Inf. Dyn. 5, 209 (1998).
[48] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998).
[49] S. Hill and W. K. Wootters, Phys. Rev. Lett. 78, 5022 (1997).

[50] A. N. Kolmogorov, in Selected Works of A. N. Kolmogorov, Volume I: Mathematics and Mechanics, Mathematics and its Applications, Vol. 25, edited by V. M. Tikhomirov (Springer, 1991) pp. 144–146.

[51] J.-L. Marichal, aequationes mathematicae 59, 74 (2000).

[52] F. Herbut, Journal of Physics A: Mathematical and General 37, 3535 (2004).

[53] K. Modi, T. Paterek, W. Son, V. Vedral, and M. Williamson, Phys. Rev. Lett. 104, 080501 (2010).

[54] C. Eltschka, T. Bastin, A. Osterloh, and J. Siewert, Phys. Rev. A 85, 022301 (2012).

[55] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd Edition (John Wiley & Sons, 2006).

[56] R. P. Stanley, Enumerative Combinatorics, Volume 1, 2nd ed. (Cambridge University Press, 2012).