

Evaluating convex roof entanglement measures - Supplemental Material

Géza Tóth,^{1,2,3,*} Tobias Moroder,⁴ and Otfried Gühne⁴

¹*Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain*

²*IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain*

³*Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary*

⁴*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Str. 3, 57068 Siegen, Germany*

(Dated: January 24, 2015)

In this supplemental material, we present some of the proofs of the theorems of the main text. We also demonstrate the usefulness of our method by further examples.

PACS numbers: 03.67.Mn, 03.65.Ud, 42.50.St

Proof of Observation 2.

Let us assume that the state has $E_{\text{lin}}^{(\text{ppt})}(\varrho) = 0$. Then, from Eq. (8) it follows that there is a symmetric PPT state ω_{12} such that $\text{Tr}(\mathcal{A}_{AA'}\omega_{12}) = 0$ and $\text{Tr}_1(\omega_{12}) = \varrho$. Hence, for this state $\text{Tr}_{A'B'}(\omega_{12}) = \varrho$ and $\mathcal{F}_{AA'}\omega_{12} = \omega_{12}$. The symmetry of ω_{12} means that $\mathcal{F}_{AA'}\mathcal{F}_{BB'}\omega_{12} = \omega_{12}$. Hence, $\mathcal{F}_{BB'}\omega_{12} = \omega_{12}$ also holds. We can write

$$\omega_{12}^{TAB} = (\mathcal{F}_{AA'}\omega_{12}\mathcal{F}_{AA'})^{TAB} = \mathcal{F}_{AA'}\omega_{12}^{TA'B}\mathcal{F}_{AA'}. \quad (\text{S1})$$

Since $\mathcal{F}_{AA'}$ is unitary, $\omega_{12}^{TAB} \geq 0$ implies $\omega_{12}^{TA'B} \geq 0$. Finally, we obtain

$$\varrho^{TA} = \text{Tr}_{A'B'}(\omega_{12}^{TA'B}). \quad (\text{S2})$$

Hence, $\omega_{12}^{TA'B} \geq 0$ implies $\varrho^{TA} \geq 0$. This proves the first part of our observation.

To prove the second part, note that based on the discussion above ω_{12} is a 2:2 symmetric extension of ϱ . It is not necessarily a PPT symmetric extension since for the $A : BA'B'$ partition it can also be non-PPT.

Finally, the third part can be proved as follows. Let us assume that ϱ has a 2 : 2 PPT symmetric extension denoted by ω_{12} . Hence, $\mathcal{F}_{AA'}\omega_{12} = \omega_{12}$ and $\mathcal{F}_{BB'}\omega_{12} = \omega_{12}$. Moreover, ω_{12} is a PPT state. Hence, $\text{Tr}(\mathcal{A}_{AA'}\omega_{12}) = 0$. ■

Note that Theorem 2 can be generalized to states that have $E_{\text{lin}}^{(n)}(\varrho) > 0$, involving PPT symmetric extensions and symmetric extensions to several parties.

Quantitative entanglement witnesses

In this section, we describe how our method can be used to construct quantitative entanglement witnesses. As an example, we present a condition for entanglement witnesses, such that the expectation value of all witnesses satisfying the condition gives a lower bound on $E_{\text{lin}}^{(\text{ppt})}$ defined in Eq. (8). We also prove that for every state ϱ_{AB} there is a witness of this type that gives not only a lower bound, but gives the value of $E_{\text{lin}}^{(\text{ppt})}$ exactly.

For the linear entropy of entanglement we needed to minimize the expectation value of the operator $M = \mathcal{A}_{AA'} \otimes \mathbb{1}_{BB'}$ over all symmetric separable states ω_{12} with a fixed reduced marginal $\text{Tr}_2(\omega_{12}) = \rho_{AB}$. Consider now an operator $W = W_{AB}$ that acts on the original bipartite Hilbert space. We require that it satisfies

$$M - \Pi_{\text{sym}}(W \otimes \mathbb{1})\Pi_{\text{sym}} = P + \Pi_{\text{sym}}Q^{T_1}\Pi_{\text{sym}} \quad (\text{S3})$$

where $P, Q \geq 0$. Here P is an operator acting only on the symmetric subspace of the two copies $\text{Sym}(\mathcal{H}_{AB}^{\otimes 2})$, while Q acts on the full tensor product $\mathcal{H}_{AB}^{\otimes 2}$ but we only used the projected symmetric part of the partial transpose. For such a decomposition, it can be shown that its expectation value for ω_{12} is

$$\begin{aligned} \text{Tr}\{[M - \Pi_{\text{sym}}(W \otimes \mathbb{1})\Pi_{\text{sym}}]\omega_{12}\} \\ &= \text{Tr}(P\omega_{12}) + \text{Tr}(\Pi_{\text{sym}}\omega_{12}\Pi_{\text{sym}}Q^{T_1}) \\ &= \text{Tr}(P\omega_{12}) + \text{Tr}(\omega_{12}Q^{T_1}) \\ &= \text{Tr}(P\omega_{12}) + \text{Tr}(\omega_{12}^T Q) \geq 0. \end{aligned} \quad (\text{S4})$$

The projectors onto the symmetric subspace Π_{sym} can be dropped in the third line since ω_{12} is supported only on it. In the last line we used $\text{Tr}(XY^{T_1}) = \text{Tr}(X^{T_1}Y)$, while nonnegativity holds because all occurring operators are positive semidefinite. Hence, Eq. (S4) can be rewritten as

$$\begin{aligned} \text{Tr}(M\omega_{12}) &\geq \text{Tr}(\Pi_{\text{sym}}(W \otimes \mathbb{1})\Pi_{\text{sym}}\omega_{12}) \\ &= \text{Tr}(W\rho_{AB}), \end{aligned} \quad (\text{S5})$$

where we have further simplified the right-hand side using that ω_{12} has a fixed reduced density matrix. Since Eq. (S5) holds for any valid state ω_{12} , it holds in particular for the one yielding the linear entropy of entanglement, thus we arrive at

$$E_{\text{lin}}(\rho_{AB}) \geq \text{Tr}(W\rho_{AB}). \quad (\text{S6})$$

Hence the expectation value of our witness provides a lower bound on the linear entropy of entanglement.

Next, we will show that for a given quantum state ρ_{AB} , if we optimize over all such witness operators, it is always possible to find one that saturates the inequality (S6).

Observation 3.—For the linear entropy of entanglement we obtain

$$E_{\text{lin}}(\rho_{AB}) \geq E_{(\text{lin})}^{\text{(ppt)}}(\rho_{AB}) = \sup_{W \in \mathcal{W}} \text{Tr}(W\rho_{AB}), \quad (\text{S7})$$

with \mathcal{W} being the set of all operators W of Eq. (S3).

Proof. The proof is given by applying the dual form of a semidefinite program [59], which has been employed in a variety of different quantum information problems. In particular we refer to Ref. [18] which explains such a procedure very nicely for the separability criterion based on symmetric extensions. We have structured the proof in two parts: In the first part, we show an equivalent formulation on the two-copy level. Afterwards we further simplify this dual problem to interpret it as an operator acting on a single density operator using techniques that were introduced in Ref. [18].

In the first part, we parse the original problem as given in Observation 1 into the form of a semidefinite program and invoke its dual, which provides the same solution. In order to achieve this one should note that the two conditions, ω_{12} just supported on the symmetric subspace and the linear equations $\text{Tr}_1(\omega) = \text{Tr}_2(\omega) = \rho_{AB}$ can be satisfied automatically with an appropriate ansatz $\omega_{12}(\mathbf{x}) = \omega_{12}^{\text{fix}} + \sum_i x_i F_i$. Here $\omega_{12}^{\text{fix}} = \sum_i s_i B_i$ is the fixed part of two-copy density operator such that the marginals equal to ρ_{AB} (its precise form being discussed later), while the remaining part $\sum_i x_i F_i$ is the yet to be determined part on the symmetric subspace, *i.e.*, the set of operators $\{B_i\}_i \cup \{F_i\}_i$ is a Hermitian operator basis for the symmetric subspace $\text{Sym}(\mathcal{H}^{\otimes 2})$. With this the primal problem reads

$$\begin{aligned} \inf \quad & \text{Tr}(M\omega_{12}^{\text{fix}}) + \sum_i x_i \text{Tr}(MF_i) & (\text{S8}) \\ \text{s.t.} \quad & \omega_{12}(\mathbf{x}) = \omega_{12}^{\text{fix}} + \sum_i x_i F_i \geq 0, \\ & \omega_{12}(\mathbf{x})^{T_1} = (\omega_{12}^{\text{fix}})^{T_1} + \sum_i x_i F_i^{T_1} \geq 0, \end{aligned}$$

where one should note that $\omega_{12}(\mathbf{x})$ acts on the symmetric subspace, while $\omega_{12}(\mathbf{x})^{T_1}$ acts on the full tensor $\mathcal{H}^{\otimes 2}$.

Taking this into account, it is straightforward to invoke the dual and to derive an equivalent optimization problem. That this dual program provides the same solution is certified for instance via the Slater regularity condition [59], which holds since this problem has an inner point, *i.e.*, $\omega_{12} = \Pi_{\text{sym}}(\rho_{AB} \otimes \rho_{AB})\Pi_{\text{sym}} > 0$ if ρ_{AB} has full rank; otherwise one should constrain \mathcal{H} anyway to the range of ρ_{AB} . Since such reformulations have been carried out quite frequently, we refer here only to the literature, and continue with its solution, which is given

by

$$\begin{aligned} \sup \quad & \text{Tr}(Z_{\text{fix}}\omega_{12}^{\text{fix}}) & (\text{S9}) \\ \text{s.t.} \quad & M - Z_{\text{fix}} = P_{\text{sym}} + \Pi_{\text{sym}}Q^{T_1}\Pi_{\text{sym}}, \\ & Z_{\text{fix}} = \sum_i z_i B_i, \\ & P_{\text{sym}} \geq 0, Q \geq 0, \end{aligned}$$

where similarly P_{sym} acts on $\text{Sym}(\mathcal{H}^{\otimes 2})$ and Q on the full tensor product space. This finishes the first part.

In the remaining part we show how the objective of Eq. (S9) can be interpreted as an operator on the single copy. For that we need some structure of the fixed part ω_{12}^{fix} that is given by the reduced state ρ_{AB} . The idea follows closely the ideas of Ref. [18], though we need to do it here for the symmetric subspace.

To start, note that any given density operator ρ_{AB} can be written as $\rho_{AB} = \mathbb{1}/d + \sum_i \text{Tr}(S_i\rho_{AB})S_i$ with $\{S_i\}_i$ being an operator basis for the traceless Hermitian operators. Next let us define $O_i = \Pi_{\text{sym}}(S_i \otimes \mathbb{1})\Pi_{\text{sym}}$. The expectation values of all these operators O_i are completely determined by the reduced state $\text{Tr}(O_i\omega_{12}(\mathbf{x})) = \text{Tr}(O_i\omega_{12}^{\text{fix}}) = \text{Tr}(S_i\rho_{AB})$, and since all these state coefficients are independent this means that the set $\{O_i\}_i$ is linearly independent. This implies a positive definite Gram matrix $G_{ij} = \text{Tr}(O_iO_j) > 0$, a unique inverse G^{-1} , and the existence of the operators $\tilde{O}_i = \sum_j (G^{-1})_{ij}O_j$. These operators are the corresponding orthogonal operators $\text{Tr}(\tilde{O}_iO_j) = \delta_{ij}$, so that the fixed part becomes

$$\begin{aligned} \omega_{12}^{\text{fix}} &= \Pi_{\text{sym}}/d_{\text{sym}} + \sum_i \text{Tr}(O_i\omega_{12}^{\text{fix}})\tilde{O}_i \\ &= \Pi_{\text{sym}}/d_{\text{sym}} + \sum_i \text{Tr}(S_i\rho_{AB})\tilde{O}_i. \end{aligned} \quad (\text{S10})$$

Note that also the desired dimensionality of d^2 matches, since $\text{Tr}_1(\omega_{12}) = \text{Tr}_2(\omega_{12}) = \rho_{AB}$ are exactly d^2 independent linear equations. To transfer this to the single copy level we write this solution in terms of a map applied to $\Lambda[\rho_{AB}] = \omega_{12}^{\text{fix}}$,

$$\Lambda[\rho_{AB}] = \text{Tr}(\rho_{AB})\Pi_{\text{sym}}/d_{\text{sym}} + \sum_i \text{Tr}(S_i\rho_{AB})\tilde{O}_i. \quad (\text{S11})$$

This map has the adjoint map, *i.e.*, the map satisfying $\text{Tr}(X\Lambda[Y]) = \text{Tr}(\Lambda^\dagger[X]Y)$ for all matrices X, Y ,

$$\Lambda^\dagger[Z] = \text{Tr}(Z)\mathbb{1}/d_{\text{sym}} + \sum_i \text{Tr}(\tilde{O}_iZ)S_i. \quad (\text{S12})$$

Via this we can finally make the connection to the single copy level by

$$\begin{aligned} \text{Tr}(Z_{\text{fix}}\omega_{12}^{\text{fix}}) &= \text{Tr}(Z_{\text{fix}}\Lambda[\rho_{AB}]) \\ &= \text{Tr}(\Lambda^\dagger[Z_{\text{fix}}]\rho_{AB}) \equiv \text{Tr}(W_Z\rho_{AB}), \end{aligned} \quad (\text{S13})$$

where we defined the single copy witness $W_Z = \Lambda^\dagger[Z_{\text{fix}}]$ in the last equation, parametrized in terms of the coefficients of Z_{fix} . However since we want to have the witness

as the open parameter we need to parametrize $Z_{\text{fix}}(W)$ in terms of the witness W . Setting

$$\begin{aligned} Z_{\text{fix}}(W) &= \Pi_{\text{sym}}(W \otimes \mathbb{1})\Pi_{\text{sym}}, \\ &= d\text{Tr}(W)\Pi_{\text{sym}} + \sum_i \text{Tr}(S_i W)O_i \end{aligned} \quad (\text{S14})$$

achieves $\Lambda[Z_{\text{fix}}(W)] = W$. Via this we can finally replace all occurrences of Z_{fix} in Eq. (S9) by W and we obtain the stated result of the observation. ■

Note that one can obtain other quantitative entanglement witnesses if one replaces the decomposable structure, as given in Eq. (S3), by a different entanglement witness condition. It is easy to see that if the operator $M - \Pi_{\text{sym}}(W \otimes \mathbb{1})\Pi_{\text{sym}}$ is non-negative on separable states then $\text{Tr}(W\rho_{AB})$ gives a lower bound. Compared to other possibilities, the advantage of the witness (S3) is that the optimization (S7) to get the lower bound can be carried out with semidefinite programming.

We also add that if one only has measured a few observables $\{O_i\}_i$ then to get a lower bound one merely has to add the constraint $W = \sum_i w_i O_i$, which means that the witness is a linear combination of the measured observables with coefficients w_i . Then, we have to optimize $\sum_i w_i v_i$, where v_i are the corresponding expectation values $\langle O_i \rangle_\rho$.

Finally, if one also wants quantitative entanglement witness for the other tasks one can proceed similarly. For instance, if one likes to bound the tangle one demands that $T - \Pi_{\text{sym}}(W \otimes \mathbb{1}^{\otimes 3})\Pi_{\text{sym}}$ is a non-negative on all fully separable states, thus it is an entanglement witness to test against full separability.

Other quantities that can be calculated by our approach

Convex roof of the Meyer-Wallach measure.—The Meyer-Wallach measure is an entanglement measure for pure states defined as [36]

$$Q = \frac{1}{N} \sum_{n=1}^N 2S_{\text{lin}}(\varrho_n), \quad (\text{S15})$$

where ϱ_n is the reduced state of the n^{th} qubits. This measure can be generalized to include the reduced states of multi-qubit groups [37]. Our method can calculate the convex roof of the measure (S15) and the generalized measures as well.

Holevo capacity.—The linear Holevo χ capacity is defined as [65, 66]

$$\chi_2(\Lambda) = \max_{\{p_k, |\Psi_k\rangle\}} \left\{ S_{\text{lin}}(\Lambda(\varrho)) - \sum_k p_k S_{\text{lin}}[\Lambda(|\Psi_k\rangle)] \right\}. \quad (\text{S16})$$

It is a capacity measure for a channel Λ . For qubit channels, explicit formula is given in Ref. [66].

Convex and concave roofs in entanglement conditions with the quantum Fisher information and the variance.—First, let us see simple entanglement conditions with the quantum Fisher information and the variance. We start from the fact that for pure N -qubit states

$$(\Delta J_x)^2 + (\Delta J_y)^2 + (\Delta J_z)^2 = \frac{N}{2} \quad (\text{S17})$$

holds. Next, we need the fundamental properties of the quantum Fisher information $F_Q[\varrho, A]$ in our criteria [67]: (i) For pure states $F_Q[\varrho, A]$ equals four times the variance $(\Delta A)_\varrho^2$. (ii) For mixed states, it is a convex function of the state. Hence, for separable states follows [38]

$$\frac{1}{4} \sum_{l=x,y,z} F_Q[\varrho, J_l] \leq \frac{N}{2}. \quad (\text{S18})$$

Due to the concavity of the variance, we can obtain a similar entanglement condition with variances as [39]

$$(\Delta J_x)^2 + (\Delta J_y)^2 + (\Delta J_z)^2 \geq \frac{N}{2}. \quad (\text{S19})$$

Any state that violates Eq. (S18) or Eq. (S19) is entangled.

The conditions (S18) and (S19) can be improved if we take the concave and convex roofs, respectively, of the left-hand sides of Eq. (S17). Hence, alternative separability conditions arise

$$\min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \sum_{l=x,y,z} (\Delta J_l)_{\Psi_k}^2 \leq \frac{N}{2}, \quad (\text{S20})$$

and

$$\max_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \sum_{l=x,y,z} (\Delta J_l)_{\Psi_k}^2 \geq \frac{N}{2}. \quad (\text{S21})$$

Any state that violates these is entangled. Numerical evidence shows that Eq. (S20) is stronger than Eq. (S18). Moreover, numerical evidence shows also that Eq. (S21) is stronger than Eq. (S19). These ideas can be extended to improve other entanglement conditions based on variances [40].

We note that Ref. [41] shows that 2×2 covariance matrices $C_\varrho(A, B)$ can always be decomposed as the

$$C_\varrho(A, B) = \sum_k p_k C_{\Psi_k}(A, B), \quad (\text{S22})$$

where ϱ has the decomposition as in Eq. (3). Hence, we know that the bound on the sum of two variances cannot be improved this way. However, Ref. [42] demonstrates that such a decomposition is not always possible for 3×3 covariance matrices. This is connected to the fact that the bound for separable states for the sum of three variances can be improved.

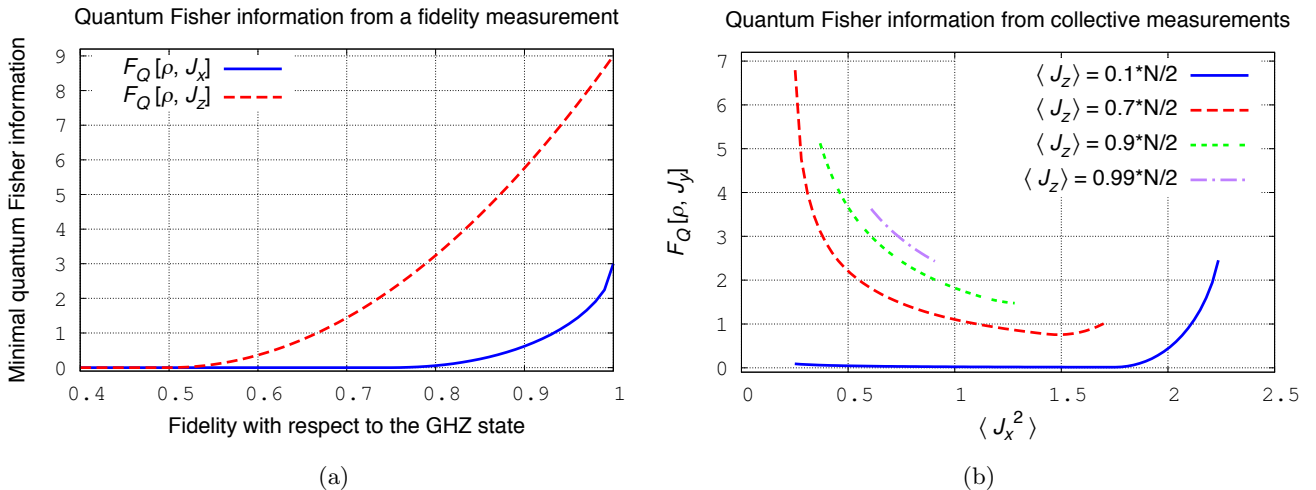


FIG. S1: (a) Lower bound on $F_Q[\varrho, J_z]$ based on the fidelity with respect to the three-qubit GHZ state. (b) Lower bound on $F_Q[\varrho, J_y]$ based on $\langle J_x^2 \rangle$ for various values of $\langle J_z \rangle$ and for $\langle J_x \rangle = 0$, for $N = 3$ qubits.

Quantum Fisher information based on incomplete data.— The quantum Fisher information can be bounded from below from partially known data. That is, we know the expectation value of some operators, and want to find a lower bound for the quantum Fisher information. The problem can be mapped to a semidefinite optimization in the two-copy space. A very good lower bound can be obtained if we optimize over PPT states.

For that we can use that the quantum Fisher information is, apart from a constant factor, the convex roof of the variance [43]

$$F_Q[\varrho, A] = 4 \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k (\Delta A)_{\Psi_k}^2. \quad (\text{S23})$$

The variance of a pure state $|\Psi\rangle$ can be expressed on two copies as

$$(\Delta A)_{\Psi}^2 = \text{Tr}[(A^2 \otimes \mathbb{1} - A \otimes A)|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|]. \quad (\text{S24})$$

Hence, a lower bound on the quantum Fisher information can be obtained as

$$F_Q^{(\text{ppt})}[\varrho, A] = \min_{\omega_{12}} \text{Tr}[(A^2 \otimes \mathbb{1} - A \otimes A)\omega_{12}], \quad (\text{S25})$$

s.t. ω_{12} symmetric, PPT,
 $\text{Tr}(O_i \omega_1) = v_i,$

where the constraints are given with the expectation values $v_i = \langle O_i \rangle_{\varrho}$. The optimization (S25) can straightforwardly be carried out with semidefinite programming.

In Fig. S1(a), we present a simple example where a lower bound on the quantum Fisher information $F_Q[\varrho, J_z]$ is shown based on measurements of the fidelity with respect to the GHZ state. Below a fidelity of 1/2, the bound for $F_Q[\varrho, J_z]$ is zero. This is due to the fact that the

product state $|11..11\rangle$ reaches this fidelity value, while $F_Q[\varrho, J_z]$ is zero for this state. If the fidelity is 1, we obtain $F_Q[\varrho, J_z] = N^2$ and $F_Q[\varrho, J_x] = N$, as expected [38].

In Fig. S1(b), we present a bound on the quantum Fisher information based on collective measurements, relevant to spin squeezing. Note that for well polarized ensembles, increasing $\langle J_x^2 \rangle$ leads to decreasing $F_Q[\varrho, J_y]$. On the other hand, for small $\langle J_z \rangle$, increasing $\langle J_x^2 \rangle$ leads to increasing $F_Q[\varrho, J_y]$. Some of the curves have points only in certain ranges of $\langle J_x^2 \rangle$, as there are no physical states corresponding to measurement results outside of these ranges, assuming a given value for $\langle J_z \rangle$ and $\langle J_x \rangle = 0$.

Similar methods can be used for bounding the variance of an observable from above based on the expectation value of other observables. We can use that the variance is the concave roof of itself [43]

$$(\Delta A)^2 = \max_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k (\Delta A)_{\Psi_k}^2. \quad (\text{S26})$$

The difference between the two cases is that for the quantum Fisher information we have to look for the minimum, while for the variance we have to look for the maximum.

Genuine multipartite entanglement.—It is possible to define quantities that detect true multipartite entanglement and can be evaluated with our method. Let us define

$$G = \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \prod_n S_{\text{linn}}(|\Psi\rangle_k) \quad (\text{S27})$$

where $S_{\text{linn}}(|\Psi\rangle)$ is the linear entropy for the n^{th} bipartition of the qudits. To be more precise, $S_{\text{linn}}(|\Psi\rangle)$ is the linear entropy of the reduced state of the qudits in one of

the two partitions for the n^{th} bipartition. If $G = 0$ then the state is biseparable, otherwise it is genuine multipartite entangled.

Similar idea can work such that only a sum of entropies must be computed by defining

$$H = \min_{\{\varrho_n\}} \sum_n p_n E_{\text{lin}}(\varrho_n), \quad (\text{S28})$$

$$\text{s.t.} \quad \sum_n p_n \varrho_n = \varrho,$$

where E_{lin} is linear entropy for the n^{th} bipartition. If $H = 0$ then the state is biseparable, otherwise it is genuine multipartite entangled. If, instead of $E_{\text{lin}}(\varrho)$, we calculate $E_{\text{lin}}^{(\text{ppt})}$ given in Eq. (8) then Eq. (S28) can be obtained via a semidefinite program. The advantage of Eq. (S28) is that only two copies of the original state are needed to calculate the value with our approach, while for the formula (S27) we need much more copies. The formalism of Eq. (S28) is in the spirit of the PPT mixer detecting genuine multipartite entanglement [44].

Note that a three-qubit state mixed from states that are PPT with respect to some partitions have been found that is genuine multipartite entangled [45]. Thus, detecting genuine multipartite entanglement is a non-trivial task.

Device independent programs

In this section, we explain the methods to obtain lower bounds on the linear entropy of entanglement for the device independent scenarios; either in the steering case where only the apparatus of one side is uncharacterized, or in Bell-type scenarios where both sides are unknown.

We will use the tool presented in Ref. [15], resting on ideas from Ref. [46, 47], which transforms the problem of estimating entanglement in a device independent scenario into the more common problem to lower bound the entanglement of a given fixed finite-dimensional system having only partial information. The method uses instead of the quantum state ϱ of unknown dimension, a finite dimensional object χ which captures most of the properties of the state.

To set the stage, let us assume that on a given side, say system A , one only knows the number of settings $x = 1, \dots, n$ and respective outcomes $a = 1, \dots, m$. This measurement scheme is described by a collection of POVM elements $M_{a|x}$, which act on a Hilbert space \mathcal{H}_A of unknown dimension. To this measurement scenario one now associates a specific completely positive local map: $\Lambda_A(\varrho_A) = \sum_k F_k \varrho_A F_k^\dagger$ with Kraus operators $F_k = \sum_{\vec{i}} |\vec{i}\rangle_{\bar{A}A} \langle k| M_{\vec{i}}$. Here $|k\rangle_A$ and $|\vec{i}\rangle_{\bar{A}}$ are respective basis states of the input and output Hilbert spaces, while $M_{\vec{i}}$ are operators out of a chosen set \mathcal{M} on which we comment shortly. However, via this structure, first observe

that this map transforms a given input state ϱ_A to

$$\chi_{\bar{A}} = \Lambda_A[\varrho_A] = \sum_{\vec{i}, \vec{j}} |\vec{i}\rangle \langle \vec{j}| \text{Tr}(M_{\vec{j}}^\dagger M_{\vec{i}} \varrho_A), \quad (\text{S29})$$

hence an output with matrix elements given by certain expectation values. At this stage the specific operator set \mathcal{M} becomes important, since so far we know nothing about $\chi_{\bar{A}}$ because we neither know ϱ_A nor $M_{\vec{i}}$. The only knowledge that we have are certain generic properties of the POVM elements $M_{a|x}$, more precisely we have (i) positivity $M_{a|x} \geq 0$, (ii) normalization $\sum_a M_{a|x} = \mathbb{1}$ and (iii) that each operator $M_{a|x}$ is a projector. Here note that by Naimark's extension any measurement can be written as a projector onto a larger dimensional space. Since for most device independent tasks this extension does not change the underlying tasks this property can be assumed without loss of generality. $M_{a|x} M_{a'|x} = \delta_{aa'} M_{a|x}$. In addition note that the expectation values of each measurement operator is observable, (iv) $\text{Tr}(M_{a|x} \varrho_A) = P(a|x)$.

Via these four properties one can thus choose specific operator sets \mathcal{M} such that one has at least some partial information on $\chi_{\bar{A}}$. For instance, if one chooses \mathcal{M} to consist of the measurement operators $\mathcal{M} = \{M_{a|x}\}_{a,x}$ one knows for instance

$$\begin{aligned} \text{Tr}(M_{a|x} M_{a'|x} \varrho_A) &= \delta_{aa'} \text{Tr}(M_{a|x} \varrho_A) \\ &= \delta_{aa'} P(a|x), \end{aligned} \quad (\text{S30})$$

while other entries like $\text{Tr}(M_{a|x} M_{a'|x'} \varrho_A)$ with $x \neq x'$ are still unknown. Nevertheless via this one gets some partial knowledge and some structure of $\chi_{\bar{A}}$, which can be captured by an explicit parametrization as

$$\begin{aligned} \chi_{\bar{A}}[P, u] &= \chi^{\text{fix}}[P] + \chi^{\text{open}}[u] \\ &= \sum_{a|x} P(a|x) Z_{a|x} + \sum_v u_v Z_v, \end{aligned} \quad (\text{S31})$$

using appropriate operators $Z_{a|x}$ and Z_v . Here the first part represents the known part of $\chi_{\bar{A}}$, while the second one is the restricted open unknown part.

Such a structure can be inferred for any choice of \mathcal{M} . For instance, one could remove some linear dependencies of the just given example set if one adds the identity and erases the last outcome for each measurement setting $\mathcal{M}_1 = \{\mathbb{1}\} \cup \{M_{a|x}\}_{a < n, x}$. In addition note that one could also enlarge this set by including also products up to N POVM elements \mathcal{M}_N , so for instance $\mathcal{M}_2 = \mathcal{M}_1 \cup \{M_{a|x} M_{a'|x'}\}_{a, a' < n, x \neq x'}$, already removing trivial parts. In this way one gets further relations like

$$\sum_{a'} \text{Tr}(M_{a|x} M_{a'|x'} M_{a'|x'} M_{a|x}) = P(a|x) \quad (\text{S32})$$

if $x \neq x'$. The advantage of including products is that one gets a tighter, more constrained, description. This set of

operators \mathcal{M}_N is precisely the one which has been mostly used [15], since it is very straightforward to “decode” all the known structure. Still there are other possibilities, like $\mathcal{M}_t = \{M_{a_1|1}M_{a_2|2}\dots M_{a_n|n}\}_{a_1,\dots,a_n}$. Here it might be harder to deduce all the structure but it has for instance the advantage that the associated map Λ_A is then even trace-preserving, thus $\chi_{\bar{A}}$ can be completely interpreted as an output quantum state; something which is not directly possible if one uses \mathcal{M}_N .

Now let us come to the concrete cases. At first let us discuss the fully device independent case where both sides are completely uncharacterized. If we locally apply the just described trace-preserving physical map (using for instance the choice \mathcal{M}_t) we transform any state ϱ_{AB} into another bipartite state $\chi_{\bar{A}\bar{B}} = \Lambda_A \otimes \Lambda_B(\varrho_{AB})$. Since an entanglement monotone does not increase under local operators and classical communication, we get $E(\varrho_{AB}) \geq E(\chi_{\bar{A}\bar{B}})$ and thus we obtain a valid lower bound by estimating the entanglement of the output state. Hence if we want to bound the linear entropy of entanglement by seeing a certain value of a Bell inequality $I \cdot P = V$ we use

$$\begin{aligned} E_{\min}(I \cdot P = V) & \\ \geq \min_{\omega_{12}, u, P} & \text{Tr}(\mathcal{A}_{12}\omega_{12}) \\ \text{s.t. } & \omega_{12} \text{ is } d_{\bar{A}} \times d_{\bar{B}}, \text{ symmetric, separable,} \\ & \omega_1 = \chi_{\bar{A}\bar{B}} = \chi_{\bar{A}\bar{B}}^{\text{fix}}[P] + \chi_{\bar{A}\bar{B}}^{\text{open}}[u] \geq 0, \\ & I_{\text{chsh}} \cdot P = V, \end{aligned} \quad (\text{S33})$$

Now let us turn to the steering case, where we assume that Alice’s side is uncharacterized while Bob obtains complete tomography. Then the data are given by the collection of unnormalized density operators $\mathcal{E} = \{\varrho_{a|x}\}_{a,x}$ for Bob with $P(a|x) = \text{Tr}(\varrho_{a|x})$. In principle we can use the same method as for the fully device independent case by employ the trace-preserving local map only on one side $\chi_{\bar{A}\bar{B}} = \Lambda_A \otimes \text{id}[\varrho_{AB}]$ and then bounding the linear entropy of entanglement of the output state.

However in this case we can do slightly better, since it is possible to bound the linear entropy of entanglement more directly on the original state ϱ_{AB} . This is in similar spirit as the negativity of Ref. [15] and Ref. [16]. Suppose we apply the same local, not necessarily trace-preserving, local map to the two copies $\chi_{12} \equiv \chi_{\bar{A}\bar{A}'\bar{B}\bar{B}'} = \Lambda_A \otimes \Lambda_A \otimes \text{id}[\omega_{12}]$. Then we can relax the constraint that ω_{12} is the symmetric PPT state of two qudits by $\chi_{12} \geq 0$, $\chi_{12}^{T_1} \geq 0$ and the permutation invariance

$$\begin{aligned} \mathcal{F}\chi_{12}\mathcal{F} &= \Lambda_A \otimes \Lambda_A \otimes \text{id}[\mathcal{F}\omega_{12}\mathcal{F}] \\ &= \Lambda_A \otimes \Lambda_A \otimes \text{id}[\omega_{12}] = \chi_{12}. \end{aligned} \quad (\text{S34})$$

Note that since the identity is within the set \mathcal{M} we have that the data of $\chi_{AB} = \Lambda_A \otimes \text{id}[\varrho_{AB}]$ are included in χ_{12} , thus we can directly parametrize $\chi_{12} = \chi_{12}^{\text{fix}}[\mathcal{E}] + \chi_{12}^{\text{open}}[u]$. The key difference compared to the previous

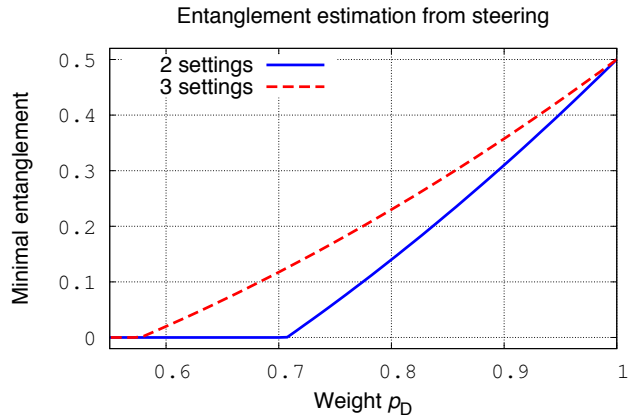


FIG. S2: Entanglement quantification via the simplest steering cases according to Eq. (S36).

case is that the objective value is still accessible. Due to the symmetry of the linear entropy of pure states, it is not surprising that $\mathcal{A}_{AA'}\mathbb{1}_{BB'} = \mathbb{1}_{AA'}\mathcal{A}_{BB'}$ holds on the symmetric subspace of the two copies. However, because the identity is included in \mathcal{M} and because Bob’s side is characterized, the expectation value of $\mathbb{1}_{AA'}\mathcal{A}_{BB'}$ is given by a linear function of the values χ_{12} ; this linear function is denoted by $\text{Tr}(\hat{E}\chi_{12})$ as a shorthand. Then we get as a lower bound

$$\begin{aligned} E_{\min}(\mathcal{E}) &\geq \min_u \text{Tr}(\hat{E}\chi_{12}) \\ \text{s.t. } & \chi_{12} = \chi_{12}^{\text{fix}}[\mathcal{E}] + \chi_{12}^{\text{open}}[u] \geq 0, \chi_{12}^{T_1} \geq 0. \end{aligned} \quad (\text{S35})$$

Before we conclude, let us point out that for both programs one can obtain sharper bounds if one includes more products into the generating set \mathcal{M} . This is very straightforward for the steering case, but even such programs quickly reach the border of being feasible. Here it remains to investigate which particular sets \mathcal{M} are more suitable than others. We leave this open for further investigation. If one combines these ideas with, for instance, the Schmidt number program then one could access even the Schmidt number (often taken as a synonym for quantum dimension) also in a device independent way.

As an example we were investigating the simplest steering scenarios, where Alice, the uncharacterized side, has either two or three dichotomic measurements and Bob performs full tomography on his qubit. In this case the available data are given by the corresponding probabilities for Alice $P(r|s)$, with $r \in \{\pm 1\}$ denoting the outcome and $s \in \{1, 2, 3\}$ the setting, and the corresponding conditional states for Bob $\rho_{r|s}$.

For the examples we assume the following data

$$P(r|s) = \frac{1}{2}, \quad \rho_{r|s} = \frac{1}{2}(\mathbb{1} - r p_D \sigma_s), \quad (\text{S36})$$

with σ_s denoting the standard Pauli matrices. Such data are for instance generated by measuring the noisy singlet

state of two qubits in the standard spin directions. Figure S2 shows the corresponding lower bound for the linear entropy of entanglement. We remark that the points where the lower bound becomes trivial coincide with the

analytic cut-off value $p_D \leq 1/\sqrt{2}$ and $p_D \leq 1/\sqrt{3}$ respectively, while the lower bound at the maximum, $E_{\min} = 1$ is exact, since such observations require at least a maximally entangled two-qubit state.