Faithful Squashed Entanglement

Fernando G. S. L. Brandão[1]**, Matthias Christandl**[2]**, Jon Yard**[3]

[1] Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Caixa Postal 702,
   MG 30123-970, Brazil. E-mail: fgslbrandao@gmail.com
[2] Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8057 Zurich, Switzerland.
   E-mail: christandl@phys.ethz.ch
[3] Center for Nonlinear Studies (CNLS), Computer, Computational and Statistical Sciences (CCS-3),
   Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
   E-mail: jtyard@lanl.gov; jtyard@gmail.com

Abstract: Squashed entanglement is a measure for the entanglement of bipartite quantum states. In this paper we present a lower bound for squashed entanglement in terms of a distance to the set of separable states. This implies that squashed entanglement is faithful, that is, it is strictly positive if and only if the state is entangled.

We derive the lower bound on squashed entanglement from a lower bound on the quantum conditional mutual information which is used to define squashed entanglement. The quantum conditional mutual information corresponds to the amount by which strong subadditivity of von Neumann entropy fails to be saturated. Our result therefore sheds light on the structure of states that *almost* satisfy strong subadditivity with equality. The proof is based on two recent results from quantum information theory: the operational interpretation of the quantum mutual information as the optimal rate for state redistribution and the interpretation of the regularised relative entropy of entanglement as an error exponent in hypothesis testing.

The distance to the set of separable states is measured in terms of the LOCC norm, an operationally motivated norm giving the optimal probability of distinguishing two bipartite quantum states, each shared by two parties, using any protocol formed by local quantum operations and classical communication (LOCC) between the parties. A similar result for the Frobenius or Euclidean norm follows as an immediate consequence.

The result has two applications in complexity theory. The first application is a quasi-polynomial-time algorithm solving the weak membership problem for the set of separable states in LOCC or Euclidean norm. The second application concerns quantum Merlin-Arthur games. Here we show that multiple provers are not more powerful than a single prover when the verifier is restricted to LOCC operations thereby providing a new characterisation of the complexity class QMA.

Contents

1. Introduction

The correlations of a bipartite quantum state ρ_{AB} can be measured by the quantum mutual information

$$I(A; B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho, \tag{1}$$

where $H(X)_\rho := -\operatorname{tr}(\rho_X \log \rho_X)$ is the von Neumann entropy. Subadditivity of entropy implies that mutual information is always positive. If $\| * \|_1$ is the trace norm, defined as $\|X\|_1 = \operatorname{tr}\sqrt{X^\dagger X}$, the inequality

$$I(A; B)_\rho \geq \frac{1}{2\ln 2} \| \rho_{AB} - \rho_A \otimes \rho_B \|_1^2, \tag{2}$$

follows from Pinsker's inequality [1] for the relative entropy. Thus a bipartite state has zero mutual information if – and only if – it has no correlations (i.e. it is a product state).

Conditional mutual information, in turn, measures the correlations of two quantum systems relative to a third. Given a tripartite state ρ_{ABE}, it is defined as

$$I(A; B|E)_\rho := H(AE)_\rho + H(BE)_\rho - H(ABE)_\rho - H(E)_\rho. \tag{3}$$

This measure is also always non-negative, as a consequence of the celebrated strong subadditivity inequality for the von Neumann entropy proved by Lieb and Ruskai [2].

As for mutual information, one can ask which states have zero conditional mutual information. Such a characterization was given in Ref. [3]. A state ρ_{ABE} has $I(A; B|E)_\rho = 0$ if and only if it is so-called quantum Markov chain, i.e. there is a decomposition of the E system vector space

$$\mathcal{H}_E = \bigoplus_j \mathcal{H}_{e_j^L} \otimes \mathcal{H}_{e_j^R} \tag{4}$$

into a direct sum of tensor products such that

$$\rho_{ABE} = \sum_j p_j \rho_{Ae_j^L} \otimes \rho_{Be_j^R}, \tag{5}$$

with states $\rho_{Ae_j^L}$ on $\mathcal{H}_A \otimes \mathcal{H}_{e_j^L}$, $\rho_{Be_j^R}$ on $\mathcal{H}_B \otimes \mathcal{H}_{e_j^R}$ and probabilities p_j. One might ask: Is there an inequality analogous to (2) for the conditional mutual information? Here one could expect the minimum distance to quantum Markov chain states to play the role of the distance to the tensor product of the reductions in (2). Up to now, however, only negative results were obtained in this direction [4].

Note, in particular, that the AB reduction of any state of the form (5) is separable [5], i.e.

$$\rho_{AB} = \sum_j p_j \rho_{A,j} \otimes \rho_{B,j}. \tag{6}$$

Rather than finding a lower bound in terms of a distance to Markov chain states, our main result, the theorem in Sect. 2, is a lower bound, similar to the one in (2), in terms of a distance to separable states.

The fact that states ρ_{ABE} with zero conditional mutual information are such that $\rho_{A:B}$ is separable motivated the introduction of the entanglement measure *squashed entanglement* [6],[1] defined as

$$E_{sq}(\rho_{A:B}) = \inf\left\{\frac{1}{2}I(A;B|E)_\rho : \quad \rho_{ABE} \text{ is an extension of } \rho_{AB}\right\}. \tag{7}$$

This measure is known to be an upper bound both for distillable entanglement and distillable key [6,9,10]. It satisfies several useful properties such as monotonicity under local operations and classical communication (LOCC), additivity on tensor products [6], and monogamy [11].

A central open question in entanglement theory, posed already in [6], is whether squashed entanglement is a faithful entanglement measure, i.e. whether it vanishes if and only if the state is separable. On the one hand, an entangled state with zero squashed entanglement would imply the existence of *bound key*, i.e. of an entangled state from which no secret key can be extracted. Analogously, a state with a non-positive partial transpose (NPT) [12,13] and vanishing squashed entanglement would show there are NPT-bound entangled states [14,15]. Both of these are core unresolved problems in quantum information theory. On the other hand, if squashed entanglement turned out to be positive on every entangled state, then we would have the first example of an entanglement measure which is faithful, LOCC monotone and which satisfies the monogamy inequality [16], three properties usually attributed to entanglement but which are not known to be compatible with each other.

The lower bound on conditional mutual information obtained in this paper resolves this question by showing that squashed entanglement is strictly positive on every entangled state. Besides its impact on entanglement theory, this result has a number of unexpected consequences for separability testing, quantum data hiding and quantum complexity theory. In the following section we will present the results of this work in detail, before proceeding to the proofs in subsequent sections.

2. Results

A lower bound on conditional mutual information. Our main result is an *approximate* version of the fact that states ρ_{ABE} with zero conditional mutual information $I(A;B|E)$ are such that $\rho_{A:B}$ is separable: We show that if a tripartite state has small conditional mutual information, its AB reduction is *close* to a separable state.

The appropriate distance measure on quantum states turns out to be crucial for a fruitful formulation of the result and its consequences. By analogy with the definition of the trace norm as the optimal probability of distinguishing two quantum states, our main results involve norms that quantify the distinguishability of quantum states under measurements that are restricted by *locality*. Specifically, given two equiprobable states

---

[1] The functional (without the factor $\frac{1}{2}$) has also been considered in [7,8].

ρ and σ, define the distance[2] $\|\rho - \sigma\|_{\mathsf{LOCC}} = 4\left(P_{\text{succ}} - \frac{1}{2}\right)$, where $P_{\text{succ}} \geq \frac{1}{2}$ is the probability of correctly distinguishing the states, maximised over all decision procedures that can be implemented using only local operations and classical communication (LOCC) [17]. When this distance is small, the states ρ and σ will roughly behave "in the same way" when used in any LOCC protocol. Writing the distance from a state ρ_{AB} to the set $\mathcal{S}_{A:B}$ of separable states on $A : B$ as

$$\|\rho_{AB} - \mathcal{S}_{A:B}\| = \min_{\sigma \in \mathcal{S}_{A:B}} \|\rho_{AB} - \sigma\|, \tag{8}$$

we may state our main result.

Theorem. *For every tripartite finite-dimensional state ρ_{ABE},*

$$I(A; B|E)_\rho \geq \frac{1}{8 \ln 2} \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}}^2. \tag{9}$$

Interestingly, a similar statement is true when replacing the LOCC norm with the Frobenius norm, since there is the dimension-independent lower bound [17]

$$\|X\|_{\mathsf{LOCC}} \geq \frac{1}{\sqrt{153}} \|X\|_2, \tag{10}$$

where $\|X\|_2 = \sqrt{\operatorname{tr} X^\dagger X}$. The Frobenius norm is also known as the Euclidean norm since it measures the Euclidean distance between quantum states when interpreted as elements of the real vector space of Hermitian matrices. Note, however, that the theorem fails dramatically if we replace the LOCC norm by the trace norm, a counterexample being a tripartite extension of the antisymmetric state[3] constructed in [10].[4] On the other hand, the theorem readily implies a lower bound for the trace norm with a dimension-dependent factor. This is because of (10) and the well-known relation between the trace norm and the Frobenius norm

$$\|X\|_2 \geq \frac{1}{\sqrt{|A||B|}} \|X\|_1, \tag{11}$$

where $|A|$ denotes the dimension of system A and $\|X\|_1 = \operatorname{tr} \sqrt{X^\dagger X}$.

A lower bound for squashed entanglement. Combining the theorem with the definition of squashed entanglement we find the following corollary.

Corollary 1. *For every state ρ_{AB},*

$$E_{\text{sq}}(\rho_{A:B}) \geq \frac{1}{16 \ln 2} \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}}^2. \tag{12}$$

Because $\| * \|_{\mathsf{LOCC}}$ is a norm, this implies that squashed entanglement is faithful, i.e. that it is strictly positive on every entangled state. This property had been conjectured in [6] and its resolution here settles the last major open question regarding squashed entanglement. Squashed entanglement is the entanglement measure which – among all known entanglement measures – satisfies most properties that have been proposed as useful for an entanglement measure. A comparison of different entanglement measures is provided in Table 1.

---

[2] This definition extends to a norm for the operators on AB, see (33).

[3] The antisymmetric state is a particular Werner state defined as $\omega_{AB} := \frac{(\mathbb{I} - \mathbb{F})}{d(d-1)}$, with $\mathbb{F}$ the flip, or swap, operator.

[4] Indeed, Ref. [10] presents an extension ω_{ABE} of ω_{AB} such that $I(A; B|E)_\omega \leq \frac{4 \log e}{|A| - 1}$, while a simple calculation gives that ω_{AB} is $\frac{1}{2}$ away from any separable state in trace distance.

Table 1. If no citation is given, the property either follows directly from the definition or was derived by the authors of the main reference. Many recent results listed in this table have significance beyond the study of entanglement measures, such as Hastings's counterexample to the additivity conjecture of the minimum output entropy [33] which implies that entanglement of formation is not strongly superadditive [32]

Measure	E_{sq} [6]	E_D [18,19]	K_D [20,21]	E_C [18,22]	E_F [18]	E_R [23]	E_R^∞ [24]	E_N [25]
normalisation	y	y	y	y	y	y	y	y
faithfulness	y Cor. 1	n [14]	?	y [26]	y	y	y [27]	n
LOCC monotonicity^a	y	y	y	y	y	y	y	y [28]
asymptotic continuity	y [29]	?	?	?	y	y [30]	y [9]	n [9]
convexity	y	?	?	?	y	y	y [31]	n
strong superadditivity	y	y	y	?	n [32,33]	n [34]	?	?
subadditivity	y	?	?	y	y	y	y	y
monogamy	y [11]	?	?	n [10]	n [10]	n [10]	n [10]	?

^{a} More precisely, we consider "weak" LOCC monotonicity, i.e. monotonicity without averages

Squashed entanglement is the quantum analogue of the intrinsic information, which is defined as

$$I(X; Y \downarrow Z) := \inf_{P_{\bar{Z}|Z}} I(X; Y|\bar{Z}), \tag{13}$$

for a triple of random variables X, Y, Z [35]. The minimisation extends over all conditional probability distributions mapping Z to $\bar{Z}$. It has been shown that the minimisation can be restricted to random variables $\bar{Z}$ with size $|\bar{Z}| = |Z|$[36]. This implies that the minimum is achieved and in particular that the intrinsic information only vanishes if there exists a channel $Z \to \bar{Z}$ such that $I(X; Y|\bar{Z}) = 0$. Whereas our work does not allow us to derive a dimension bound on the system E in the minimisation of squashed entanglement and hence conclude that the minimisation is achieved in general, we can assert such a bound if squashed entanglement vanishes: Corollary 1 implies that squashed entanglement vanishes only for separable ρ_{AB}. By Caratheodory's theorem, the number of terms in the separable decomposition of $\sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$ can be bounded by $|AB|^2$, and thus $\rho_{ABE} = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i} \otimes |i\rangle\langle i|_E$ has vanishing conditional mutual information with $E = |AB|^2$. Equivalently, there exists a channel applied to a purification of ρ_{AB} resulting in ρ_{ABE} such that $I(A; B|E)_\rho$ vanishes.

Positive rate in state redistribution. Quantum conditional mutual information can be given an operational interpretation in the context of the *state redistribution* problem [37,38]. Suppose a sender and a receiver share a quantum state $|\psi\rangle_{ABEE'}$ and the sender (who initially holds subsystems BE) wants to send B to the receiver (who initially holds E'), while preserving the purity of the global state. The purifying subsystem A is unavailable to both of them. Given free entanglement between the sender and receiver, the minimum achievable rate of quantum communication needed to send B, with vanishing error in the limit of a large number of copies of the state, is given by $\frac{1}{2} I(A; B|E)_\psi$ [37,38]. Then from an optimal protocol for state redistribution we find $E_{sq}(\rho_{A:B})$ to be the minimum rate of quantum communication needed to send the B system, optimised over all possible ways of distributing the side-information among E and E' [39].

One can ask: Is a positive rate of quantum communication always required in order to redistribute a system entangled with another (irrespective of the side information available to the sender and receiver)? Corollary 1 allows us to answer this question in the affirmative. The need of a positive rate in state redistribution can then be seen as a new distinctive feature of quantum correlations.

The quantum de Finetti theorem, the LOCC norm and data-hiding states. We say a bipartite state $\rho_{A:B}$ is k-extendible if there is a state $\rho_{A:B_1,...,B_k}$ that is permutation-symmetric in the B systems with $\rho_{A:B} = \mathrm{tr}_{B_2,...,B_k}(\rho_{A:B_1,...,B_k})$. Such a family of states provides a sequence of approximations to the set of separable states. A state is separable if, and only if, it is k-extendible for every k [40–45].

Indeed, quantum versions of the de Finetti theorem [44,45] show that any k-extendible state $\rho_{A:B}$ is such that

$$\| \rho_{AB} - \mathcal{S}_{A:B} \|_1 \leq \frac{4|B|^2}{k}. \tag{14}$$

Moreover, this bound is close to tight, as there are k-extendible states that are $\Omega(|B|k^{-1})$-away from the set of separable states [45].

Unfortunately, for many applications this error estimate – exponentially large in the number of qubits of the state – is too big to be useful. Our next result shows that a significant improvement is possible if we are willing to relax our notion of distance of two quantum states. It shows that in LOCC norm we can obtain an error term that grows as the square root of the number of qubits of the A system:

Corollary 2. *Let $\rho_{A:B}$ be k-extendible. Then*

$$\| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}} \leq \sqrt{\frac{16 \ln 2 \log |A|}{k}}. \tag{15}$$

The key point in the proof of Corollary 2 is the fact that squashed entanglement satisfies the so-called monogamy inequality [11], namely

$$E_{\mathrm{sq}}(\rho_{A:B_1 B_2}) \geq E_{\mathrm{sq}}(\rho_{A:B_1}) + E_{\mathrm{sq}}(\rho_{A:B_2}), \tag{16}$$

for every state $\rho_{AB_1 B_2}$. This, together with the bound $E_{\mathrm{sq}}(\rho_{A:B}) \leq \log |A|$ [6] implies that the squashed entanglement of any k-extendible state must be smaller than $k^{-1} \log |A|$, which combined with Corollary 1 gives Corollary 2.

We do not know if the bound given in Corollary 2 is tight. An indication, however, is given by the following example, which shows that for all k there is a k-extendible state ρ_{AB} with $\log |A| = k$ and $\| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}} \geq 1$.

Example (Lower bound). Consider systems $A = A_1 A_2 \cdots A_k$ and $B = B'B''$ and define

$$\rho_{AB} := \mathrm{tr}_{B_2 \cdots B_k}(\rho_{AB_1 B_2 \cdots B_k}), \tag{17}$$

where

$$\rho_{A_1...A_k:B_1'B_1''\cdots B_k'B_k''} := \mathrm{Id}_{A_1 \cdots A_k} \otimes$$
$$\mathrm{Sym}_{B_1'B_1'',...,B_k'B_k''}\left(\Phi_{A_1 B_1'} \otimes |1\rangle\langle 1|_{B_1''} \otimes \cdots \otimes \Phi_{A_k B_k'} \otimes |k\rangle\langle k|_{B_k''} \right), \tag{18}$$

where $\Phi := |\Phi\rangle\langle\Phi|$ is the projector onto an EPR pair $|\Phi\rangle := (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$, Id is the identity superoperator and $\mathrm{Sym}_{B_1,...,B_k}$ is the symmetrization superoperator defined as

$$\mathrm{Sym}_{B_1,...,B_k}(X) := \frac{1}{k!} \sum_{\pi \in S_k} P_\pi X P_{\pi^{-1}}, \tag{19}$$

with S_k the symmetric group of order k and P_π the representation of the permutation π which acts on a k-partite vector space as $P_\pi |l_1\rangle \otimes \cdots \otimes |l_k\rangle = |l_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |l_{\pi^{-1}(k)}\rangle$. The state ρ_{AB} takes the form

$$\rho_{AB} = \frac{1}{k} \sum_{l=1}^{k} \Phi_{A_l B'} \otimes |l\rangle\langle l|_{B''} \otimes \tau_{A_1 \cdots A_{l-1} A_{l+1} \cdots A_k}, \tag{20}$$

with $\tau_C := \mathbb{I}/|C|$ and is by construction k-extendible. Moreover we can deterministically obtain an EPR pair from ρ_{AB} by the following simple LOCC protocol: Bob measures the B'' register in the computational basis and tells Alice the outcome l obtained. Then she traces out all her systems except the lth. Thus the LOCC distance of ρ_{AB} to separable states is at least the LOCC distance of an EPR pair to separable states, which is known to be one [46], and so

$$\| \rho_{AB} - \mathcal{S}_{A:B} \|_{\text{LOCC}} \geq 1. \tag{21}$$

A direct implication of Corollary 2 concerns data-hiding states [47–50]. Every state ρ_{AB} that can be well-distinguished from separable states by a global measurement, yet is almost completely indistinguishable from a separable state by LOCC measurements is a so-called data-hiding state: it can be used to hide a bit of information (of whether the state prepared is ρ_{AB} or its closest separable state in LOCC norm) that is not accessible by LOCC operations alone. The antisymmetric state of sufficiently high dimension is an example of a data hiding state [49], as are random mixed states with high probability [50] (given an appropriate choice of the dimensions and the rank of the state). Corollary 2 shows that highly extendible states that are far away in trace norm from the set of separable states must necessarily be data-hiding.

A quasipolynomial-time algorithm for separability of quantum states. Given a density matrix describing a bipartite quantum state $\rho_{A:B}$, the separability problem consists of determining if the state is entangled or separable. This is one of the most basic questions in quantum information theory and has been a topic of active research in the past years (see e.g. [15]).

In the weak-membership problem $W_{\text{SEP}}(\epsilon, \| * \|)$ for separability one should decide if a given bipartite state $\rho_{A:B}$ is in the interior of the set of separable states or ϵ-away from any separable state (in norm $\| * \|$), given the promise that one of the two alternatives holds true. The best known algorithms for the problem have worst case complexity $2^{\text{poly}(|A|,|B|) \log(1/\epsilon)}$ (see e.g. [51–54]). In fact, the problem is NP-hard for $\epsilon = 1/\text{poly}(n)$ [55–57].

Corollary 2 implies that for the LOCC and Euclidean norm, one can greatly improve upon the previous known algorithms.

Corollary 3. *There is a quasipolynomial-time algorithm for solving the weak membership problem for separability in the LOCC norm and the Euclidean norm. More precisely, there is an $\exp(O(\epsilon^{-2} \log |A| \log |B|))$-time algorithm for deciding $W_{\text{SEP}}(\epsilon, \| * \|_{\text{LOCC}})$ and $W_{\text{SEP}}(\epsilon, \| * \|_2)$.*

We note that in many applications of the separability algorithm, e.g. accessing the usefulness of a quantum state for violating Bell's inequalities or for performing quantum teleportation, the LOCC norm is actually the relevant quantity to consider. Also, it is intriguing that the complexity of our algorithm matches the best hardness result for the trace-norm version of the problem [58]. It is an open question if a similar hardness result

could be obtained for the LOCC norm case, which would imply that our algorithm is essentially optimal.

The algorithm, which we analyse in more detail in Sect. 4, is very simple and in fact is the basis of a well-known hierarchy of efficient tests for separability [51]. Using semidefinite-programming one looks for a $\Omega(\log|A|\epsilon^{-2})$-extension of $\rho_{A:B}$ and decides that the state is separable if one is found. If the state is separable, then clearly there is such an extension. If it is ϵ-away from any separable state, Corollary 2 shows that no such extension exists.[5]

More detailed information on this algorithm as well as the related problem of optimising a linear function over separable states, relevant for instance in mean-field theory, is given in [59].

Quantum Merlin-Arthur games with multiple Merlins. A final application of the theorem concerns quantum Merlin-Arthur games. The class QMA can be considered the quantum analogue of NP and is formed by all languages that can be decided in quantum polynomial-time by a verifier who is given a quantum system of polynomially many qubits as a proof (see e.g. [60]). It is natural to ask how robust this definition is. A few results are known in this direction: For example, it is possible to amplify the soundness and completeness parameters to exponential accuracy, even without enlarging the proof size [61]. Also, the class does not change if we allow a first round of logarithmic-sized quantum communication from the verifier to the prover [62].

From Corollary 2 we get a new characterisation of QMA, which at first sight might appear to be strictly more powerful: We show QMA to be equal to the class of languages that can be decided in polynomial time by a verifier who is given a constant number k of unentangled proofs and can measure them using any quantum polynomial-time implementable LOCC protocol (among the k proofs). This answers an open question of Aaronson *et al.* [63]. We hope this characterisation of QMA proves useful in devising new QMA verifying systems.

In order to formalise our result we consider the classes $\mathsf{QMA}_\mathsf{M}(k)_{m,s,c}$, defined in analogy to QMA as follows [58,63,64]:

Definition. *Let* M *be a set of POVM elements. A language L is in $\mathsf{QMA}_\mathsf{M}(k)_{m,s,c}$ if for every input $x \in \{0,1\}^n$ there exists a polynomial time implementable two outcome measurement $\{M_x, \mathbb{I} - M_x\}$ with M_x in* M *such that*

- **Completeness:** *If $x \in L$, there exist k witnesses $|\psi_1\rangle, \ldots, |\psi_k\rangle$, each of m qubits, such that*

$$\mathrm{tr}\left(M_x\left(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|\right)\right) \geq c. \tag{22}$$

- **Soundness:** *If $x \notin L$, then for any states $|\psi_1\rangle, \ldots, |\psi_k\rangle$,*

$$\mathrm{tr}\left(M_x\left(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|\right)\right) \leq s. \tag{23}$$

We call $\mathsf{QMA}_\mathsf{M}(k) = \mathsf{QMA}_\mathsf{M}(k)_{\mathrm{poly}(n),2/3,1/3}$.

Let $\mathsf{M} = \mathsf{SEP}_\mathsf{YES}$ be the class of (non-normalised) separable POVM elements.[6] Then Harrow and Montanaro showed that $\mathsf{QMA}_{\mathsf{SEP}_\mathsf{YES}}(2) = \mathsf{QMA}(2) = \mathsf{QMA}(k)$ for

---

[5] In [51] one imposed the further constraint that the symmetric-extension has to have a positive partial transpose with respect to all B systems. It is an interesting open question if the worst-case complexity of the algorithm (in the LOCC norm case) can be further improved taking into consideration this extra constraint.

[6] POVM elements A from this class form two-outcome POVMs $\{A, \mathbb{I} - A\}$ known as separable-once-removed, meaning that they are separable measurements once one of the effects is removed. Here we use the notation $\mathsf{SEP}_\mathsf{YES}$ to denote that the POVM element associated to *accept* should be be separable, i.e. proportional to a separable state.

any $k = \text{poly}(n)$ [58], i.e. two proofs are just as powerful as k proofs and one can restrict the verifier's action to $\mathsf{SEP_{YES}}$ without changing the expressive power of the class.

We define $\mathsf{QMA_{LOCC}}(k)$ in an analogous way, but now the verifier can only measure the k proofs with a LOCC measurement. Then we have,

Corollary 4. *For $k = O(1)$,*

$$\mathsf{QMA_{LOCC}}(k) = \mathsf{QMA}. \tag{24}$$

In particular, for any $\epsilon > 0$

$$\mathsf{QMA_{LOCC}}(2)_{m,s,c} \subseteq \mathsf{QMA}_{O(m^2\epsilon^{-2}),s+\epsilon,c}. \tag{25}$$

A preliminary result in the direction of Corollary 4 appeared in [65], where a similar result was shown for $\mathsf{QMA_{LO}}(k)$, a variant of $\mathsf{QMA}(k)$ in which the verifier is restricted to implement only local measurements on the k proofs and jointly post-process the outcomes classically.[7]

It is an open question whether (25) remains true if we consider $\mathsf{QMA}(2)$ instead of $\mathsf{QMA_{LOCC}}(2)$. If this turns out to be the case, then it would imply an optimal conversion of $\mathsf{QMA}(2)$ into QMA in what concerns the proof length (under a plausible complexity-theoretic assumption). For it follows from [58] (based on the $\mathsf{QMA}(\text{poly}\log(n)\sqrt{n})_{\log(n),1,1/3}$ protocol for 3-SAT of [63]) that unless there is a subexponential-time quantum algorithm for 3-SAT, then there is a constant $\epsilon_0 > 0$ such that for every $\delta > 0$,

$$\mathsf{QMA}(2)_{m,s,c} \nsubseteq \mathsf{QMA}_{O(m^{2-\delta}\epsilon_0^{-2}),s+\epsilon_0,c}. \tag{26}$$

Another open question is whether a similar result holds for $\mathsf{QMA_{LOCC}}(2)$. This would be the case if one could construct a LOCC version of the productness test of Harrow and Montanaro [58].

An interesting approach to the $\mathsf{QMA}(2)$ vs. QMA question concerns the existence of *disentangler* superoperators [63], defined as follows.

Definition. *A superoperator $\Lambda : S \to AB$ is a $(\log|S|, \epsilon, \delta)$-disentangler in the $\|*\|$ norm if*

- $\Lambda(\rho)$ *is ϵ-close to a separable state for every ρ, and*
- *for every separable state σ, there is a ρ such that $\Lambda(\rho)$ is δ-close to σ.*

As noted in [63], the existence of an efficiently implementable $(\text{poly}(\log|A|, \log|B|),$ $\epsilon, \delta)$-disentangler in trace norm (for sufficiently small ϵ and δ) would imply $\mathsf{QMA}(2) = \mathsf{QMA}$. Watrous has conjectured that this is not the case and that for every $\epsilon, \delta < 1$, any (ϵ, δ)-disentangler (in trace-norm) requires $|S| = 2^{\Omega(\min(|A|,|B|))}$. For an exponentially large $|S|$, in turn, Matsumoto presented a construction of a quantum disentangler [66] and the quantum de Finetti theorem gives an alternative construction [45]. Corollary 2 can be rephrased as saying that there *is* an efficient disentangler in LOCC norm.

Corollary 5. *Let $k = \lceil 16\ln 2 \log|A|\epsilon^{-2} \rceil$ and $S := AB_1 \cdots B_k$. Define the superoperator $\Lambda : S \to AB$, with $|B| = |B_j|$ for all $j \le k$, as*

$$\Lambda(\rho_{AB_1 \cdots B_k}) := \frac{1}{k} \sum_{i=1}^{k} \rho_{AB_i}. \tag{27}$$

---

[7] $\mathsf{QMA_{LO}}(k)$ is also called $\mathsf{BellQMA}(k)$ [63] since the verifier is basically restricted to perform a *Bell test* on the proofs.

Then Λ is a $(O(\log |A| \log |B| \epsilon^{-2}), \epsilon, 0)$-disentangler in LOCC norm.

For more applications of this work to quantum Merlin-Arthur games, see [59].

Distinguishability under locality-restricted measurements. We will now give a formal definition of the norm $\| * \|_{\mathsf{LOCC}}$ that was used in order to state the results. $\| * \|_{\mathsf{LOCC}}$ can be seen as a restricted version of the trace norm $\| * \|_1$. Introducing the set ALL as consisting of all operators M that satisfy $0 \leq M \leq \mathbb{I}$, the trace norm can be written as

$$\|X\|_1 = \max_{M \in \mathsf{ALL}} \mathrm{tr}((2M - \mathbb{I})X), \tag{28}$$

where $\mathbb{I}$ is the identity matrix. The trace norm is of special importance in quantum information theory as it is directly related to the optimal probability for distinguishing two equiprobable states ρ and σ with a quantum measurement.[8] In analogy with this interpretation of the trace norm, and for operators X on the tensor product space AB, we define the LOCC norm as

$$\|X\|_{\mathsf{LOCC}} := \max_{M \in \mathsf{LOCC}} \mathrm{tr}((2M - \mathbb{I})X), \tag{29}$$

where LOCC is the convex set of matrices $0 \leq M \leq \mathbb{I}$ such that there is a two-outcome measurement $\{M, \mathbb{I} - M\}$ that can be realized by LOCC between Alice and Bob [17]. It will become clear below why this defines a norm. The optimal bias in distinguishing ρ and σ by any LOCC protocol is then $\frac{1}{2}\|\rho - \sigma\|_{\mathsf{LOCC}}$. We note that in many applications of the separability problem, e.g. assessing the usefulness of a quantum state for violating Bell's inequalities or for performing quantum teleportation, the LOCC norm is actually the more relevant quantity to consider.

In the proofs, we will also need versions of the LOCC norm, where the number of rounds of communication is restricted. Similarly to the definition of LOCC above, we therefore introduce $\mathsf{LOCC}^{\rightarrow}(k)$ consisting of operators $0 \leq M \leq \mathbb{I}$ such that there is a two-outcome measurement $\{M, \mathbb{I} - M\}$ that can be realized by k-round LOCC protocol with the first communication from Alice to Bob. A special role is played by the one-round implementable POVM elements $\mathsf{LOCC}^{\rightarrow} \equiv \mathsf{LOCC}^{\rightarrow}(1)$ for which an explicit characterisation exists (Lemma 4). Note that

$$\mathsf{LOCC} = \lim_{k \to \infty} \mathsf{LOCC}^{\rightarrow}(k). \tag{30}$$

Although not strictly needed in this work, one of the lemmas (Lemma 3) will also be valid for the set SEP which is defined as the set of POVM elements M such that both M and $\mathbb{I} - M$ are proportional to a separable state. The inclusions of the sets then imply the following relations between the norms.

$$\|X\|_1 \geq \|X\|_{\mathsf{SEP}} \geq \|X\|_{\mathsf{LOCC}} \geq \cdots \geq \|X\|_{\mathsf{LOCC}^{\rightarrow}(k)}$$
$$\geq \|X\|_{\mathsf{LOCC}^{\rightarrow}(k-1)} \geq \cdots \geq \|X\|_{\mathsf{LOCC}^{\rightarrow}}. \tag{31}$$

Analogous statements hold true for $\mathsf{LOCC}^{\leftarrow}(k)$, where the reverse arrow indicates that the first round of communication is from Bob to Alice.

---

[8] Two-outcome measurements suffice for such tasks, and these are described by a pair of positive semidefinite matrices summing to $\mathbb{I}$, which we write $\{M, \mathbb{I} - M\}$. When the state is ρ, the probabilities of the outcomes are $\Pr(M) = \mathrm{tr}(M\rho)$ and $\Pr(\mathbb{I} - M) = \mathrm{tr}((\mathbb{I} - M)\rho) = 1 - \Pr(M)$. The optimal bias of distinguishing two states ρ and σ is then given by $\max_{0 \leq M \leq \mathbb{I}} \mathrm{tr}(M(\rho - \sigma)) = \frac{1}{2}\|\rho - \sigma\|_1$.

The introduced norms fit into a general framework [17] for restricted norms where one considers M, a closed, convex subset of the operator interval

$$\{M : 0 \leq M \leq \mathbb{I}\} = \{M : M \geq 0, \|M\|_\infty \leq 1\} \tag{32}$$

containing $\mathbb{I}$, having nonempty interior, and such that $M \in \mathsf{M}$ implies $\mathbb{I} - M \in \mathsf{M}$. Then

$$\|X\|_{\mathsf{M}} = \max_{M \in \mathsf{M}} \mathrm{tr}\,((2M - \mathbb{I})X) = \max_{Y \in \mathsf{D}} \mathrm{tr}\,YX, \tag{33}$$

where

$$\mathsf{D} = \mathrm{conv}(\mathsf{M} \cup -\mathsf{M}) = \{2M - \mathbb{I} : M \in \mathsf{M}\} = \{Y : \|Y\|_{\mathsf{M}}^* \leq 1\}$$

is the unit ball for the dual norm $\|\cdot\|_{\mathsf{M}}^*$.

Relating entanglement measures. The core of the proof of the theorem is composed of three steps, Lemmas 1, 2, and 3 below, each of which is a new result about entanglement measures. Together, they establish the weaker inequality

$$I(A; B|E) \geq \frac{1}{8 \ln 2} \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}\leftarrow}^2, \tag{34}$$

and by symmetry, also the same inequality with for $\mathsf{LOCC}^\rightarrow$. In Sect. 3D, we then give an inductive proof that the same bound holds for k-round LOCC, where k is arbitrary:

$$I(A; B|E) \geq \frac{1}{8 \ln 2} \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}^\rightarrow(k)}^2. \tag{35}$$

For $k \to \infty$, this proves the theorem.

A measure that plays a fundamental role in the proof is the regularised relative entropy of entanglement [24,34], defined as

$$E_R^\infty(\rho_{A:B}) := \lim_{n \to \infty} \frac{E_R(\rho_{A:B}^{\otimes n})}{n}, \tag{36}$$

with the relative entropy of entanglement given by

$$E_R(\rho_{A:B}) := \min_{\sigma_{AB} \in \mathcal{S}_{A:B}} S(\rho_{AB} || \sigma_{AB}), \tag{37}$$

where the quantum relative entropy is given by $S(\rho || \sigma) = \mathrm{tr}(\rho(\log \rho - \log \sigma))$ if $\mathrm{supp}\,\rho \subseteq \mathrm{supp}\,\sigma$ and $S(\rho || \sigma) = \infty$ otherwise. A distinctive property of the relative entropy of entanglement among entanglement measures is the fact that it is not lockable, i.e. by discarding a small part of the state, E_R can only drop by an amount proportional to the number of qubits traced out. Indeed, as shown in [67], one has

$$E_R(\rho_{A:BE}) \leq E_R(\rho_{A:E}) + 2H(B)_\rho. \tag{38}$$

While the same is true for E_R^∞, we use an optimal protocol for state redistribution [37,38] to obtain the following improvement of (38), which is valid only for the regularised quantity.

Lemma 1. *Let ρ_{ABE} be a quantum state. Then,*

$$I(A; B|E)_\rho \geq E_R^\infty(\rho_{A:BE}) - E_R^\infty(\rho_{A:E}). \tag{39}$$

The second step in the proof is to obtain a lower bound on $E_R^\infty(\rho_{A:BE}) - E_R^\infty(\rho_{A:E})$, which by Lemma 1 also gives a lower bound on the conditional mutual information. If we could prove that E_R^∞ satisfied the monogamy inequality given by (16), we would be done, since Ref. [68] shows that

$$E_R^\infty(\rho_{A:B}) \geq \frac{1}{2 \ln 2} \| \rho_{A:B} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}}^2. \tag{40}$$

However, the antisymmetric state is a counterexample [10]. Nonetheless, we will show that the regularised relative entropy of entanglement satisfies a weaker form of the monogamy inequality, and this will suffice for us.

In order to state the new inequality, we have to recall an operational interpretation of E_R^∞ in the context of quantum hypothesis testing [27] (see Sect. 3B for more details). Suppose Alice and Bob are given either n copies of an entangled state $\rho_{A:B}$, or an arbitrary state which is separable in the cut $A^n : B^n$. Suppose further they can only measure a POVM from the class M in order to find out which of the two hypotheses is the case. Let $p_e(n)$ be the probability that Alice and Bob mistake a separable state for the n copies of $\rho_{A:B}$, minimised over all possible measurements available to them.[9] Then we define the optimal rate function $D_\mathsf{M}(\rho_{A:B})$ as follows:

$$D_\mathsf{M}(\rho_{A:B}) = \lim_{n \to \infty} -\frac{\log p_e(n)}{n}. \tag{41}$$

The main result of [27] implies

$$D_\mathsf{ALL}(\rho_{A:B}) = E_R^\infty(\rho_{A:B}), \tag{42}$$

i.e. the regularised relative entropy of entanglement is the optimal distinguishability rate when trying to distinguish many copies of an entangled state from (arbitrary) separable states, in the case where there are no restrictions on the measurements available.

The new monogamy-like inequality we prove for E_R^∞, using the connection of E_R^∞ with hypothesis testing given by (42), is the following:

Lemma 2. *For every state $\rho_{A:BE}$,*

$$E_R^\infty(\rho_{A:BE}) - E_R^\infty(\rho_{A:E}) \geq D_{\mathsf{LOCC}^\leftarrow}(\rho_{A:B}). \tag{43}$$

The third and final step is to obtain a lower bound for $D_{\mathsf{LOCC}^\leftarrow}$ in terms of the minimum $\mathsf{LOCC}^\leftarrow$ distance to the set of separable states. In fact, we prove a slightly more general result.

Lemma 3. *For any $\mathsf{M} \in \{\mathsf{LOCC}^\rightarrow, \mathsf{LOCC}^\leftarrow, \mathsf{LOCC}, \mathsf{SEP}\}$ and every state $\rho_{A:B}$,*

$$D_\mathsf{M}(\rho_{A:B}) \geq \frac{1}{8 \ln 2} \| \rho_{A:B} - \mathcal{S}_{A:B} \|_\mathsf{M}^2. \tag{44}$$

In [68] a similar result was proved: consider the following modification of the relative entropy of entanglement:

$$E_{R,\mathsf{M}}(\rho) := \max_{M \in \mathsf{M}} \min_{\sigma \in \mathcal{S}} S(M(\rho)||M(\sigma)), \tag{45}$$

---

[9] We also require that the measurement only mistakes $\rho_{A:B}^{\otimes n}$ for a separable state with a small probability. See Sect. 3B for details.

with $M(X) := (\operatorname{tr} MX)|0\rangle\langle 0| + (1 - \operatorname{tr} MX)|1\rangle\langle 1|$ for $M \in \mathsf{M}$. Then for any class $\mathsf{M} \in \{\mathsf{LOCC}^{\rightarrow}, \mathsf{LOCC}^{\leftarrow}, \mathsf{LOCC}, \mathsf{SEP}\}$, it was shown that [68]

$$E_{R,\mathsf{M}}(\rho_{A:B}) \geq \frac{1}{2\ln 2} \| \rho_{A:B} - \mathcal{S}_{A:B} \|_{\mathsf{M}}^2. \tag{46}$$

It is conceivable that

$$E_{R,\mathsf{M}}^{\infty}(\rho_{A:B}) = D_{\mathsf{M}}(\rho_{A:B}), \tag{47}$$

and if this was the case we could obtain Lemma 3 from (46), since $E_{R,\mathsf{M}}$ is superadditive. Indeed, for $\mathsf{M} = \mathsf{ALL}$, (47) holds true and this is precisely the content of (42). For restricted classes of POVMs, however, it is not known whether (47) remains valid and, alas, the techniques of [27] do not appear to be directly applicable. We note that one direction is easily seen to hold, namely[10]

$$E_{R,\mathsf{M}}^{\infty}(\rho_{A:B}) \geq D_{\mathsf{M}}(\rho_{A:B}). \tag{48}$$

So we can obtain an asymptotic version of (46) from Lemma 3.

A new construction of entanglement witnesses: In the proof of Lemma 3 we derive a relation that might be of independent interest. For any restricted set of POVM elements M,

$$2 \max_{M \in \mathsf{M}} \left(\operatorname{tr}(M\rho) - \max_{\sigma \in \mathcal{S}} \operatorname{tr}(M\sigma) \right) = \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{M}}. \tag{49}$$

The POVM element maximising the left hand side is a so-called *entanglement witness* [15,71]: its expectation value on the entangled state ρ is bigger than its expectation value on any separable state. Therefore it can be used as a witness to attest the entanglement of ρ_{AB}. (49) gives a new construction of entanglement witnesses that is attractive in the following respects:

- The witness is a POVM element and thus can be measured with a *single* measurement setting.
- Every M that (after rescaling) contains an informationally complete POVM can be used to construct witnesses detecting every entangled state.
- The maximum gap in expectation value for an entangled state versus separable states, optimised over all witnesses from a class M, is equal to the minimum distance of the entangled state to the set of separable states in the M norm and thus can also be used to quantify the entanglement of the state (see [72,73] for related results in this direction).

3. Proof of Theorem

A. Proof of Lemma 1. The main idea in the proof of Lemma 1 is to apply the nonlockability of E_R given by (38) to a tensor power state $\rho_{A:BE}^{\otimes n}$, but to use an optimal protocol for performing state redistribution protocol to trace out B in a more efficient way. A general protocol for state redistribution is outlined in Fig. 1. The following proposition follows immediately from the existence of a state redistribution protocol achieving the minimum possible rates of communication and entanglement cost:

---

[10] The proof is very similar to the usual argument, which shows $S(\rho||\sigma)$ to be an upper bound on the optimal distinguishability rate in quantum hypothesis testing (see e.g. [69,70]), and thus we omit it.
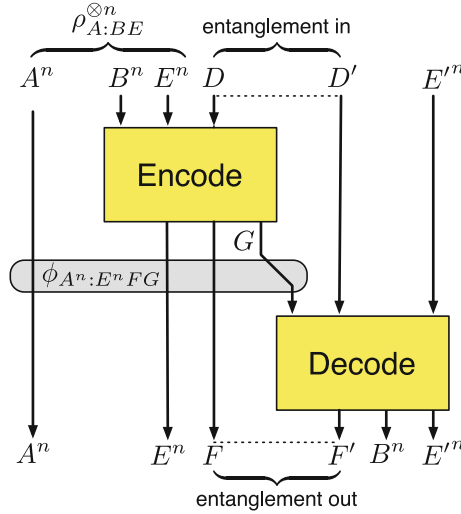
Fig. 1. General protocol for redistributing the B parts of a pure state $|\psi\rangle^{\otimes n}_{ABEE'}$ where the sender holds E and the receiver holds E'. The protocol consumes maximal entanglement on DD', transmits the quantum system G, and produces maximal entanglement on FF'. The dimensions of D, F and G typically depend exponentially on n. The protocol is asymptotically reversible

Proposition 1. ([37,38]) *There is a sequence of asymptotically reversible encoding maps* $\Lambda_n : B^n E^n D \to E^n F G$ *such that if* $\phi_{A^n:E^n F G} = \Lambda_n(\rho^{\otimes n}_{A:BE} \otimes \tau_D)$ *(see* Fig. 1*), then*

$$\frac{\log |G|}{n} \to \tfrac{1}{2} I(A; B|E)_\rho, \tag{50}$$

and

$$\| \phi_{A^n E^n F} - \rho^{\otimes n}_{AE} \otimes \tau_F \|_1 \to 0, \tag{51}$$

where τ_D *and* τ_F *are the maximally mixed states on systems D and F, respectively.*

Proof of Lemma 1.

$$E^\infty_R(\rho_{A:BE}) = \lim_{n\to\infty} \frac{1}{n} E_R(\phi_{A^n:E^n F G}) \tag{52}$$

$$\leq \left(\lim_{n\to\infty} \frac{1}{n} E_R(\phi_{A^n:E^n F}) \right) + I(A; B|E) \tag{53}$$

$$= \left(\lim_{n\to\infty} \frac{1}{n} E_R(\rho^{\otimes n}_{A:E} \otimes \tau_F) \right) + I(A; B|E) \tag{54}$$

$$= E^\infty_R(\rho_{A:E}) + I(A; B|E). \tag{55}$$

The first line holds by asymptotic continuity of E_R [30,74] and the asymptotic reversibility of the encoding operation. The second follows from (38) and (50). The third line holds again by continuity of E_R, applied to the asymptotic equality (51) of the respective arguments of E_R. □

The same argument applied to the mutual information instead of the relative entropy of entanglement can be used to prove the converse of state redistribution (see [75,76] for a similar argument in the context of one-shot quantum state splitting).

B. Proof of Lemma 2. Before proceeding to the proof, we give a precise definition of the rate functions D_M.

Definition. *Let ρ_{AB} be a state and $M \in \{LOCC^{\rightarrow}, LOCC^{\leftarrow}, LOCC, SEP, ALL\}$ a class of POVMs. Then we define $D_M(\rho_{A:B})$ as the unique non-negative number such that*

- *(Direct part): For every $\epsilon > 0$ there exists a sequence of POVM elements $\{M_{A^n B^n}\}_{n \in \mathbb{N}}$, with $M_{A^n B^n} \in M$, such that*

$$\lim_{n \to \infty} \operatorname{tr}(M_{A^n B^n} \rho_{AB}^{\otimes n}) = 1 \tag{56}$$

and for every $n \in \mathbb{N}$ and $\omega_{A^n B^n} \in \mathcal{S}_{A^n:B^n}$,

$$-\frac{\log \operatorname{tr}(M_{A^n B^n} \omega_{A^n B^n})}{n} \geq D_M(\rho_{A:B}) - \epsilon. \tag{57}$$

- *(Converse): If a real number $\epsilon > 0$ and a sequence of POVM elements $\{M_{A^n B^n}\}_{n \in \mathbb{N}}$ are such that for every $n \in \mathbb{N}$ and every $\omega_{A^n B^n} \in \mathcal{S}_{A^n:B^n}$,*

$$-\frac{\log(\operatorname{tr}(M_{A^n B^n} \omega_{A^n B^n}))}{n} \geq D_M(\rho_{A:B}) + \epsilon, \tag{58}$$

then

$$\lim_{n \to \infty} \operatorname{tr}(M_{A^n B^n} \rho_{AB}^{\otimes n}) < 1. \tag{59}$$

Recall (42):

Proposition 2 ([27]). *For every state $\rho_{A:B}$,*

$$D_{ALL}(\rho_{A:B}) = E_R^{\infty}(\rho_{A:B}). \tag{60}$$

The main application of Proposition 2 so far has been to prove that the asymptotic manipulation of entanglement under operations that, approximately, do not generate entanglement is reversible: both the distillation and a cost function in this paradigm are equal to the regularised relative entropy of entanglement [77–79]. In fact, one can show that Proposition 2 is *equivalent* to such reversibility under non-entangling operations, in the sense that the latter can be used to prove the former [80].

The proof of Lemma 2 will utilise the special structure of LOCC measurements with one-way communication:

Lemma 4. *Any operator in $LOCC^{\leftarrow}$ necessarily has the form*

$$\sum_k M_{A,k} \otimes N_{B,k}, \tag{61}$$

where $0 \leq M_{A,k} \leq \mathbb{I}_A$, $0 \leq N_{B,k}$ and $\sum_k N_{B,k} \leq \mathbb{I}_B$.

Proof. Any LOCC measurement consisting of communication only from Bob to Alice has the general form $\{M_{A,a|b} \otimes N_{B,b}\}$, where $0 \leq M_{A,a|b}, 0 \leq N_{B,b}, \sum_a M_{A,a|b} \leq \mathbb{I}_A$ and $\sum_b N_{B,b} \leq \mathbb{I}_B$. The convex hull of such POVM elements consists of all operators of the form given in the lemma. $\square$

Proof of Lemma 2. The main idea in this proof is to start with measurements for $\rho_{A:E}$ and $\rho_{A:B}$ that respectively achieve $D_{\mathsf{ALL}}(\rho_{A:E})$ and $D_{\mathsf{LOCC}^\leftarrow}(\rho_{A:B})$, and to construct a measurement for $\rho_{A:BE}$ distinguishing it from separable states at a rate given by the sum of the individual rates. The lemma then follows from Proposition 2. Here, it is crucial that we only use one-way LOCC measurements to discriminate $\rho_{A:B}$ from separable states, since only in this case are we able to make sure that the composed measurement for $\rho_{A:BE}$ acts as it should.

Let $\delta > 0$ and fix two δ-optimal sequences of POVM elements: an unrestricted sequence $Q_{A^n E^n}$ of POVM elements for distinguishing $\rho_{A:E}$ from separable states, and a sequence $S_{A^n B^n}$ of one-way LOCC (with classical communication from Bob to Alice) elements for distinguishing $\rho_{A:B}$ from separable states. This means that

$$\mathrm{tr}\left(Q_{A^n E^n}\rho_{A:E}^{\otimes n}\right) \to 1 \quad \text{and} \quad \mathrm{tr}\left(S_{A^n B^n}\rho_{A:B}^{\otimes n}\right) \to 1, \tag{62}$$

while

$$-\frac{\log\mathrm{tr}\left(Q_{A^n E^n}\omega_{A:E}^n\right)}{n} \geq D_{\mathsf{ALL}}(\rho_{A:E}) - \delta \tag{63}$$

for every state $\omega_{A^n:E^n}$ separable across $A^n\!:E^n$ and

$$-\frac{\log\mathrm{tr}\left(S_{A^n B^n}\omega_{A^n:B^n}\right)}{n} \geq D_{\mathsf{LOCC}^\leftarrow}(\rho_{A:B}) - \delta \tag{64}$$

for every state $\omega_{A^n:B^n}$ separable across $A^n\!:B^n$. We will use these to construct a sequence of POVM elements $T_{A^n B^n E^n}$ that distinguish $\rho_{A:BE}^{\otimes n}$ from separable states $\omega_{A^n:B^n E^n}$, satisfying

$$\lim_{n\to\infty}\mathrm{tr}(T_{A^n B^n E^n}\rho_{A:BE}^{\otimes n}) = 1 \tag{65}$$

and

$$\lim_{n\to\infty} -\frac{\log\mathrm{tr}(T_{A^n B^n E^n}\omega_{A^n:B^n E^n})}{n} \geq D_{\mathsf{ALL}}(\rho_{A:E}) + D_{\mathsf{LOCC}^\leftarrow}(\rho_{A:B}). \tag{66}$$

Because the left-hand-side is a lower bound to the maximal achievable error rate $D_{\mathsf{ALL}}(\rho_{A:BE})$, Proposition 2 then gives the desired inequality

$$E_R^\infty(\rho_{A:BE}) \geq E_R^\infty(\rho_{A:E}) + D_{\mathsf{LOCC}^\leftarrow}(\rho_{A:B}).$$

To construct $T_{A^n B^n E^n}$, first observe that by Lemma 4, we can write the $\mathsf{LOCC}^\leftarrow$ operator $S_{A^n B^n}$ in the form

$$S_{A^n B^n} = \sum_k M_{A^n,k} \otimes N_{B^n,k}, \tag{67}$$

with $0 \leq M_{A^n,k} \leq \mathbb{I}$, $0 \leq N_{B^n,k}$ and $\sum_k N_{B^n,k} \leq \mathbb{I}$. It will be useful to consider a Naimark extension of the measurement on B, consisting of orthogonal projections $P_{B^n\widehat{B},k}$ such that $\langle 0_{\widehat{B}}|P_{B^n\widehat{B},k}|0_{\widehat{B}}\rangle = N_{B^n\widehat{B},k}$. Define

$$R_{A^n B^n \widehat{B}} := \sum_k M_{A^n,k} \otimes P_{B^n\widehat{B},k}, \tag{68}$$

which by construction is such that $0 \leq R_{A^n B^n \widehat{B}} \leq \mathbb{I}$. Define

$$T_{A^n B^n E^n} := \langle 0_{\widehat{B}} | \sqrt{R_{A^n B^n \widehat{B}}} \ Q_{A^n E^n} \sqrt{R_{A^n B^n \widehat{B}}} | 0_{\widehat{B}} \rangle \tag{69}$$

$$= \sum_k \left(\sqrt{M_{A^n,k}} \ Q_{A^n E^n} \sqrt{M_{A^n,k}} \right) \otimes N_{B^n,k}. \tag{70}$$

Note that we omit identity operators from our notation here and in the remainder of this proof. There should be no room for confusion, as the global Hilbert space will always be clear, and we use superscripts to indicate on which "local systems" each operator acts. For instance, $Q_{A^n E^n} \equiv Q_{A^n E^n} \otimes \mathbb{I}_{B^n \widehat{B}}$ in Eq. (69), while $M_{A^n,k} \equiv M_{A^n,k} \otimes \mathbb{I}_{E^n}$ in Eq. (70).

Observe that $0 \leq T_{A^n B^n E^n} \leq \mathbb{I}$ because the same holds for $Q_{A^n E^n}$ and $R_{A^n B^n \widehat{B}}$. We will now verify the conditions (65) and (66) required to complete the proof. For proving (65), we first write

$$\mathrm{tr}(T_{A^n B^n E^n} \rho_{A:BE}^{\otimes n}) = \mathrm{tr} \left(Q_{A^n E^n} \sqrt{R_{A^n B^n \widehat{B}}} \left(\rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \sqrt{R_{A^n B^n \widehat{B}}} \right). \tag{71}$$

From (62)

$$\lim_{n \to \infty} \mathrm{tr} \left(R_{A^n B^n \widehat{B}} \left(\rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \right) = \lim_{n \to \infty} \mathrm{tr}(S_{A^n B^n} \rho_{A:B}^{\otimes n}) = 1. \tag{72}$$

Hence by the gentle measurement lemma [81,82],

$$\lim_{n \to \infty} \left\| \rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} - \sqrt{R_{A^n B^n \widehat{B}}} \left(\rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \sqrt{R_{A^n B^n \widehat{B}}} \right\|_1 = 0, \tag{73}$$

and so (71) and (62) give

$$\lim_{n \to \infty} \mathrm{tr}(T_{A^n B^n E^n} \rho_{A:BE}^{\otimes n}) \geq \lim_{n \to \infty} \mathrm{tr} \left(Q_{A^n E^n} \left(\rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \right)$$

$$- \lim_{n \to \infty} \left\| \rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} - \sqrt{R_{A^n B^n \widehat{B}}} \left(\rho_{A:BE}^{\otimes n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \sqrt{R_{A^n B^n \widehat{B}}} \right\|_1$$

$$= 1. \tag{74}$$

Let us now prove (66). We write

$$\mathrm{tr}(T_{A^n B^n E^n} \omega_{A^n:B^n E^n}) = \mathrm{tr} \left(Q_{A^n E^n} \tilde{\omega}_{A^n:E^n} \right) \mathrm{tr} \left(R_{A^n B^n \widehat{B}} \left(\omega_{A^n:B^n E^n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \right), \tag{75}$$

with

$$\tilde{\omega}_{A^n:E^n} := \frac{\mathrm{tr}_{B^n \widehat{B}} \left(\sqrt{R_{A^n B^n \widehat{B}}} \left(\omega_{A:BE}^n \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \sqrt{R_{A^n B^n \widehat{B}}} \right)}{\mathrm{tr} \left(R_{A^n B^n \widehat{B}} \left(\omega_{A^n:B^n E^n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \right)}. \tag{76}$$

We claim that $\tilde{\omega}_{A^n:E^n}$ is separable in the cut $A^n : E^n$ whenever $\omega_{A^n:B^n E^n}$ is separable in the cut $A^n : B^n E^n$. Then (75) will imply (66) because

$$\lim_{n \to \infty} -\frac{\log \mathrm{tr}(T_{A^n B^n E^n} \omega_{A^n:B^n E^n})}{n} = \lim_{n \to \infty} -\frac{\log \mathrm{tr}(Q_{A^n E^n} \tilde{\omega}_{A^n:E^n})}{n}$$

$$+ \lim_{n \to \infty} -\frac{\log \mathrm{tr}(S_{A^n B^n} \omega_{A^n:B^n})}{n}$$

$$\geq E_R^\infty(\rho_{A:E}) + D_{\mathsf{LOCC}\leftarrow}(\rho_{A:B}) - 2\delta, \tag{77}$$

where we used Proposition 2 and (63) and (64).

To show that $\tilde{\omega}_{A^n:E^n}$ is separable, we now prove that the completely-positive trace-non-increasing map $\Lambda: A^n B^n E^n \to A^n E^n$ defined as

$$\Lambda(X_{A^n B^n E^n}) := \text{tr}_{B^n \widehat{B}} \left(\sqrt{R_{A^n B^n \widehat{B}}} \left(X_{A^n B^n E^n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \sqrt{R_{A^n B^n \widehat{B}}} \right) \qquad (78)$$

is a separable superoperator. From the form of $R_{A^n B^n \widehat{B}}$ given in (68) we can write

$$\Lambda(X_{A^n B^n E^n}) := \sum_{k,k'} \text{tr}_{B^n \widehat{B}} \left[\left(\sqrt{M_{A^n,k}} \otimes P_{B^n \widehat{B},k} \right) \right.$$
$$\left. \times \left(X_{A^n B^n E^n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \left(\sqrt{M_{A^n,k'}} \otimes P_{B^n \widehat{B},k'} \right) \right].$$

Then, taking the partial trace over $B^n \widehat{B}$ we find that the cross terms in (79) are zero since the projectors $P_{B^n \widehat{B},k}$ are mutually orthogonal and therefore

$$\Lambda(X_{A^n B^n E^n}) := \sum_{k} \text{tr}_{B^n \widehat{B}} \left[\left(\sqrt{M_{A^n,k}} \otimes P_{B^n \widehat{B},k} \right) \right.$$
$$\left. \times \left(X_{A^n B^n E^n} \otimes |0\rangle\langle 0|_{\widehat{B}} \right) \left(\sqrt{M_{A^n,k}} \otimes P_{B^n \widehat{B},k} \right) \right],$$

which is manifestly a separable superoperator. This implies that $\tilde{\omega}_{A^n:E^n} = \Lambda(\omega_{A^n:B^n E^n})$ is separable, since $\omega_{A^n:B^n E^n}$ is separable. $\quad\square$

C. Proof of Lemma 3. The final step is now to prove Lemma 3. For that we construct a particular protocol for distinguishing many copies of an entangled state from arbitrary separable states, with an exponentially small probability of mistaking a separable state for many copies of a particular entangled state. In Corollary II.2 of [27], a similar result was shown using the exponential de Finetti theorem [83]. Here we give a simpler proof with a stronger bound on the distinguishability rate.

A useful result we will employ is the following simple extension of [84], which in turn is a consequence of the minimax theorem, which we quote below the proof.

Lemma 5. *Let* M *be a closed convex set of POVM elements and* $\mathcal{C}$ *a closed convex set of states. Then given* $\rho \notin \mathcal{C}$,

$$2 \max_{M \in \mathsf{M}} \left(\text{tr}(M\rho) - \max_{\sigma \in \mathcal{C}} \text{tr}(M\sigma) \right) = \|\rho - \mathcal{C}\|_{\mathsf{M}}. \qquad (79)$$

Proof. Following Jain [84], let the function $u : (\mathsf{M}, \mathcal{C}) \to \mathbb{R}$ be defined as $u(M, \sigma) = \text{tr}(M\rho) - \text{tr}(M\sigma)$. The sets M and $\mathcal{C}$ are compact convex subsets of the (real) vector space of Hermitian matrices. Since the function u is furthermore affine in both arguments, the conditions of the minimax theorem are satisfied (see Proposition 3 just below) and we find

$$2 \max_{M \in \mathsf{M}} \left(\text{tr}(M\rho) - \max_{\sigma \in \mathcal{C}} \text{tr}(M\sigma) \right) = 2 \max_{M \in \mathsf{M}} \min_{\sigma \in \mathcal{C}} (\text{tr}(M\rho) - \text{tr}(M\sigma)) \qquad (80)$$
$$= 2 \min_{\sigma \in \mathcal{C}} \max_{M \in \mathsf{M}} (\text{tr}(M\rho) - \text{tr}(M\sigma)) \qquad (81)$$
$$= \min_{\sigma \in \mathcal{C}} \|\rho - \sigma\|_{\mathsf{M}}. \qquad (82)$$

Proposition 3 *(Minimax theorem). Let C_1, C_2 be non-empty, convex and compact subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $u : C_1 \times C_2 \to \mathbb{R}$ be convex in the first argument and concave in the second and continuous in both. Then,*

$$\min_{x_1 \in C_1} \max_{x_2 \in C_2} u(x_1, x_2) = \max_{x_2 \in C_2} \min_{x_1 \in C_1} u(x_1, x_2). \tag{83}$$

In the proof of Lemma 3 we will also make use of Azuma's inequality, a large deviation bound for correlated random variables. We say a sequence of random variables $X_0, X_1, X_2, \ldots$ is a supermartingale if for all $k \in \mathbb{N}$,

$$\mathbb{E}(X_{k+1}|X_0, \ldots, X_k) \leq X_k. \tag{84}$$

Then we have

Proposition 4 *(Azuma's inequality). Let $\{X_k\}_k$ be a supermartingale with $|X_{k+1} - X_k| \leq c_k$. Then for all positive integers n and all positive reals t,*

$$\Pr(X_n - X_0 \geq t) \leq \exp\left(-\frac{t^2}{2\left(\sum_{k=1}^{n} c_k^2\right)}\right). \tag{85}$$

Proof of Lemma 3. By Lemma 5, there is $M_{AB} \in \mathsf{M}$ such that

$$b = a + \frac{1}{2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_\mathsf{M}, \quad \text{where} \quad a := \max_{\sigma_{AB} \in \mathcal{S}_{A:B}} \operatorname{tr}(M_{AB}\sigma_{AB}), \quad b := \operatorname{tr}(M_{AB}\rho_{AB}), \tag{86}$$

and without loss of generality $\|\rho_{AB} - \mathcal{S}_{A:B}\|_\mathsf{M} > 0$, or $b > a$.

The element $M_{A^n B^n} \in \mathsf{M}$ is constructed as follows. Fix $\delta > 0$. One measures the POVM $\{M_{AB}, \mathbb{I} - M_{AB}\}$ in each of the n copies, obtaining n binary random variables $\{Z_i\}_{i=1}^n$, where $Z_i = 1$ is associated with effect M_{AB} and $Z_i = 0$ with effect $\mathbb{I} - M_{AB}$. Then if

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \geq a + (1 - \delta)(b - a), \tag{87}$$

we accept (this corresponds to $M_{A^n B^n}$). Otherwise we reject.

First, by the law of large numbers it is clear that

$$\lim_{n \to \infty} \operatorname{tr}(M_{A^n B^n}\rho_{AB}^{\otimes n}) = 1. \tag{88}$$

We now show that

$$\lim_{n \to \infty} -\frac{\log \operatorname{tr}(M_{A^n B^n}\omega_{A^n B^n})}{n} \geq \frac{(1 - \delta)^2}{8 \ln 2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_\mathsf{M}^2. \tag{89}$$

The key observation is that, in case we have a separable state $\omega_{A^n B^n}$, for all $i \in \{1, \ldots, n\}$,

$$\Pr(Z_i = 1|Z_1, \ldots, Z_{i-1}) \leq a. \tag{90}$$

This is so because by assumption $M_{AB} \in \mathsf{SEP}$ implies that both POVM elements in the measurement $\{M_{AB}, \mathbb{I} - M_{AB}\}$ are separable operators. Therefore, the state to

which the kth measurement is applied is always separable, even conditioning on previous measurement outcomes. Then by Proposition 4,

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} Z_i \geq a + (1-\delta)(b-a)\right) \leq 2^{-\frac{(1-\delta)^2}{2\ln 2}n(b-a)^2}. \tag{91}$$

In more detail, define the random variables

$$X_k := \sum_{i=1}^{k} (Z_i - a). \tag{92}$$

Note that

$$\mathbb{E}\left(X_k | X_1, \ldots, X_{k-1}\right) = \mathbb{E}\left(Z_k | Z_1, \ldots, Z_{k-1}\right) - a + X_{k-1} \tag{93}$$

$$= \Pr\left(Z_k = 1 | Z_1, \ldots, Z_{k-1}\right) - a + X_{k-1} \tag{94}$$

$$\leq X_{k-1}, \tag{95}$$

i.e. $\{X_k\}_k$ form a supermartingale. Moreover, for all k, $|X_k - X_{k-1}| = |Z_k - a| \leq 1$. Hence by Proposition 4, setting $X_0 = 0$, we have

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} Z_i \geq a + (1-\delta)(b-a)\right) = \Pr\left(X_n \geq n(1-\delta)(b-a)\right) \tag{96}$$

$$\leq \exp\left(-\frac{(1-\delta)^2}{2}n(b-a)^2\right) \tag{97}$$

$$= 2^{-\frac{(1-\delta)^2}{2\ln 2}n(b-a)^2}. \tag{98}$$

From (86),

$$-\frac{1}{n}\log\left(\Pr\left(\frac{1}{n}\sum_{i=1}^{n} Z_i \geq a + (1-\delta)(b-a)\right)\right) \geq \frac{(1-\delta)^2}{8\ln 2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_{\mathsf{M}}^2, \tag{99}$$

and thus

$$D_{\mathsf{M}}(\rho_{A:B}) \geq \frac{(1-\delta)^2}{8\ln 2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_{\mathsf{M}}^2. \tag{100}$$

The result then follows from the fact that $\delta > 0$ is arbitrary. $\quad\square$

D Proof of Theorem

Proof. We show by induction that for every ρ_{ABE} and every $k \in \mathbb{N}$,

$$I(A; B|E) \geq \frac{1}{8\ln 2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_{\mathsf{LOCC}^{\rightarrow}(k)}^2, \tag{101}$$

which is equivalent to the statement of the theorem. As we have already established the base case $k = 1$ for all ρ_{ABE} as a consequence of Lemmas 1, 2, and 3,

$$I(A; B|E) \geq \frac{1}{8\ln 2}\|\rho_{AB} - \mathcal{S}_{A:B}\|_{\mathsf{LOCC}^{\rightarrow}}^2, \tag{102}$$

we only need to assume that (101) holds for $k - 1$ and prove it for k.

By Lemma 5, we have for all $\rho_{AB} \notin \mathcal{S}_{A:B}$,

$$\|\rho_{AB} - \mathcal{S}_{A:B}\|_{\text{LOCC}^{\rightarrow}(k)} = \text{tr}(M_{AB}\rho_{AB}) - \max_{\sigma_{AB} \in \mathcal{S}_{A:B}} \text{tr}(M_{AB}\sigma_{AB}) \qquad (103)$$

for some $M_{AB} \in \text{LOCC}^{\rightarrow}(k)$. Without loss of generality, we may assume that M_{AB} is a sum of POVM elements from a measurement of the following form: First Alice performs a projective measurement with mutually orthogonal projectors $P_{A,j}$ on her share of the state (we assume that the ancilla register needed to realise a general POVM is contained in the state $\rho_{A:B}$) and communicates the outcome to Bob, obtaining the state

$$\Lambda(\rho_{AB}) := \sum_j (P_{A,j} \otimes \mathbb{I}_B)\rho_{AB}(P_{A,j} \otimes \mathbb{I}_B) \otimes |j\rangle\langle j|_{\widehat{B}} \qquad (104)$$

on $AB\widehat{B}$. Then they measure a $(k - 1)$-round LOCC POVM on $AB\widehat{B}$ with the first round of communication from Bob to Alice, so that $M_{AB} = \Lambda^{\dagger}(\widetilde{M}_{AB\widehat{B}})$ for some $\widetilde{M}_{AB\widehat{B}} \in \text{LOCC}^{\leftarrow}(k - 1)$, where $\Lambda^{\dagger}$ is the dual map to Λ.

Let $\omega_{AB\widehat{B}E\widehat{E}}$ be the extension of $(\mathbb{I}_E \otimes \Lambda)(\rho_{ABE})$ obtained by letting $\widehat{E}$ be a copy of the classical register $\widehat{B}$. It satisfies

$$I(A; B|E)_{\rho} \geq I(A\widehat{B}; B|E)_{\omega} \geq I(A; B\widehat{B}|E\widehat{E})_{\omega}. \qquad (105)$$

Indeed, the first inequality is by the data processing inequality, or equivalently, by strong subadditivity [2], while the second follows from the chain

$$\begin{aligned} I(A\widehat{B}; B|E)_{\omega} &= I(\widehat{B}; B|E)_{\omega} + I(A; B|\widehat{B}E)_{\omega} \\ &\geq I(A; B|\widehat{B}E)_{\omega} = I(A; B\widehat{B}|E\widehat{E})_{\omega}. \end{aligned} \qquad (106)$$

Here, we have used the chain rule for conditional mutual information, strong subadditivity, and the fact that $\widehat{E}$ is a copy of $\widehat{B}$. Assuming that (101) holds with $k - 1$ and applying it to this extension (with swapped roles of A and $B\widehat{B}$ in order to account for reverse arrow) gives

$$I(A; B\widehat{B}|E\widehat{B})_{\omega} = I(B\widehat{B}; A|E\widehat{E})_{\omega} \geq \frac{1}{8\ln 2}\|\Lambda(\rho_{AB}) - \mathcal{S}_{A:B\widehat{B}}\|^2_{\text{LOCC}^{\leftarrow}(k-1)}. \qquad (107)$$

Because $M_{AB} = \Lambda^{\dagger}(\widetilde{M}_{AB\widehat{B}})$, we have the following, which, together with the inequalities (105) and (107), shows the desired inequality (101):

$$\|\Lambda(\rho_{AB}) - \mathcal{S}_{A:B\hat{B}}\|_{\text{LOCC}^{\leftarrow}(k-1)} = \max_{M_{AB\widehat{B}} \in \text{LOCC}^{\leftarrow}(k-1)} \qquad (108)$$

$$\times \left(\text{tr}(M_{AB\widehat{B}}\Lambda(\rho_{AB})) - \max_{\sigma_{AB\widehat{B}} \in \mathcal{S}_{A:B\widehat{B}}} \text{tr}\left(M_{AB\widehat{B}}\sigma_{AB\widehat{B}}\right)\right)$$

$$\geq \text{tr}(\widetilde{M}_{AB\widehat{B}}\Lambda(\rho_{AB})) - \max_{\sigma_{AB\widehat{B}} \in \mathcal{S}_{A:B\widehat{B}}} \text{tr}(\widetilde{M}_{AB\widehat{B}}\sigma_{AB\widehat{B}}) \qquad (109)$$

$$\geq \text{tr}(\widetilde{M}_{AB\widehat{B}}\Lambda(\rho_{AB})) - \max_{\sigma_{AB} \in \mathcal{S}_{A:B}} \text{tr}(\widetilde{M}_{AB\widehat{B}}\Lambda(\sigma_{AB})) \qquad (110)$$

$$= \text{tr}(M_{AB}\rho_{AB}) - \max_{\sigma_{AB} \in \mathcal{S}_{A:B}} \text{tr}(M_{AB}\sigma_{AB}) \qquad (111)$$

$$= \|\rho_{AB} - \mathcal{S}_{A:B}\|_{\text{LOCC}^{\rightarrow}(k)}. \qquad (112)$$

The only nonobvious step here is (110), but before we show it, we remark that since the reverse inequality is obvious, we will actually have established the equality

$$\| \Lambda(\rho_{AB}) - \mathcal{S}_{A:B\hat{B}} \|_{\mathsf{LOCC}^{\leftarrow}(k-1)} = \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}^{\rightarrow}(k)}, \qquad (113)$$

which, together with (105) and (107) proves the induction step.

To show the inequality (110), first observe that

$$M_{AB} = \sum_j \left(P_{A,j} \otimes \mathbb{I}_B \otimes |j\rangle\langle j|_{\hat{B}} \right) \widetilde{M}_{AB\hat{B}} \left(P_{A,j} \otimes \mathbb{I}_B \otimes |j\rangle\langle j|_{\hat{B}} \right). \qquad (114)$$

We may thus assume without loss of generality that

$$\widetilde{M}_{AB\hat{B}} = \sum_j \widetilde{M}_{AB,j} \otimes |j\rangle\langle j|_{\hat{B}} \qquad (115)$$

for operators $\widetilde{M}_{AB,j} \in \mathsf{LOCC}^{\leftarrow}(k-1)$ satisfying $(P_{A,j} \otimes \mathbb{I}_B)\widetilde{M}_{AB,j}(P_{A,j} \otimes \mathbb{I}_B) = \widetilde{M}_{AB,j}$. Now let $\sigma_{AB\hat{B}} \in \mathcal{S}_{A:B\hat{B}}$ achieve the maximum in (109). Because of the assumed form of the $\widetilde{M}_{AB,j}$, we may assume that

$$\sigma_{AB\hat{B}} = \sum_j p_j \sigma_{AB,j} \otimes |j\rangle\langle j|_{\hat{B}} \qquad (116)$$

for $\sigma_{AB,j} \in \mathcal{S}_{A:B}$ and probabilities p_j. Then the average state $\overline{\sigma}_{AB} = \sum_j p_j \sigma_{AB,j}$ satisfies

$$\mathrm{tr}\left(\widetilde{M}_{AB\hat{B}} \sigma_{AB\hat{B}} \right) = \mathrm{tr}\left(\widetilde{M}_{AB\hat{B}} \Lambda(\overline{\sigma}_{AB}) \right) \leq \max_{\sigma'_{AB} \in \mathcal{S}_{A:B}} \mathrm{tr}\left(\widetilde{M}_{AB\hat{B}} \Lambda(\sigma'_{AB}) \right). \qquad (117)$$

□

4 Proofs of Corollaries

We start with the proof of Corollary 2, which is straightforward.

Proof of Corollary 2. Squashed entanglement satisfies the monogamy inequality [11]

$$E_{\mathrm{sq}}(\rho_{A:BB'}) \geq E_{\mathrm{sq}}(\rho_{A:B}) + E_{\mathrm{sq}}(\rho_{A:B'}). \qquad (118)$$

Repeatedly applying it to $\rho_{A:B_1\cdots B_k}$ and noting that $E_{\mathrm{sq}}(\rho_{A:B_1\cdots B_k}) \leq \log|A|$ [6], we get

$$\log|A| \geq E_{\mathrm{sq}}(\rho_{A:B_1\ldots B_k}) \geq k E_{\mathrm{sq}}(\rho_{A:B}), \qquad (119)$$

where we used that $\rho_{AB_j} = \rho_{AB}$ for all j. The result then follows from the lower bound for squashed entanglement given in Corollary 1.   □

Here, we give an outline of the proof of Corollary 3, concerning the quasipolynomial-time algorithm for deciding separability. A more careful treatment can be found in [59].

Proof of Corollary 3. Given a density matrix ρ_{AB}, we use semidefinite programming [85] in order to search for a $\left\lceil \frac{16 \ln 2 \log |A|}{\epsilon^2} \right\rceil$-extension of $\rho_{A:B}$. If $\rho_{AB} \in \mathcal{S}_{A:B}$, then there is such an extension (since separable states have a k-extension for every k). If $\| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}} \geq \epsilon$, Corollary 2 implies that there is no such extension.

The computational cost of solving a feasibility semidefinite problem with n variables and a matrix of dimension m is $O(n^2 m^2)$. A k-extension of ρ_{AB} has dimension $|A||B|^k$ and less than $|A|^2 |B|^{2k}$ variables. Hence the time complexity of searching for a $\left\lceil \frac{16 \ln 2 \log |A|}{\epsilon^2} \right\rceil$-extension is of order $\exp\left(O(\epsilon^{-2} \log |A| \log |B|) \right)$. $\quad\square$

Proof of Corollary 4. We start by proving (25). Consider a protocol in $\mathsf{QMA}_{\mathsf{LOCC}}(2)_{m,s,c}$ given by the LOCC measurement $\{M_{AB}, \mathbb{I} - M_{AB}\}$. We construct a protocol in $\mathsf{QMA}_{O(m^2 \epsilon^{-2}), s+\epsilon, c}$ that can simulate it: The verifier asks for a proof of the form $|\psi\rangle_{AB_1 \cdots B_k}$, where $|A| = |B_i| = 2^m$ for all i (each register consists of m qubits) and $k = \lceil 16 \ln 2 \epsilon^{-2} m \rceil$. He then symmetrises the B systems obtaining the state $\rho_{AB_1 \cdots B_k}$ and measures $\{M_{AB}, \mathbb{I} - M_{AB}\}$ in the subsystems $AB \equiv AB_1$.

Let us analyse the completeness and soundness of the protocol. For completeness, the prover can send $|\psi\rangle_A \otimes |\phi\rangle_B^{\otimes k}$, for states $|\psi\rangle, |\phi\rangle$ such that

$$\operatorname{tr} |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_B M_{AB} \geq c. \tag{120}$$

Thus the completeness parameter of the QMA protocol is at least c.

For soundness, we note that by Corollary 2,

$$\| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}} \leq \epsilon. \tag{121}$$

Thus, as $\{M_{AB}, \mathbb{I} - M_{AB}\} \in \mathsf{LOCC}$ the soundness parameter for the QMA protocol can only be ϵ away from s. Indeed, for every $\rho_{AB_1 \cdots B_k}$ symmetric in B,

$$\operatorname{tr}(M_{AB}\rho_{AB}) \leq \max_{\sigma_{AB} \in \mathcal{S}_{A:B}} \operatorname{tr}(M_{AB}\sigma_{AB}) + \| \rho_{AB} - \mathcal{S}_{A:B} \|_{\mathsf{LOCC}} \leq s + \epsilon. \tag{122}$$

Equation (24) now follows easily from the protocol above. Given a protocol in $\mathsf{QMA}_{\mathsf{LOCC}}(l)$ with each proof of size m qubits, we can simulate it in $\mathsf{QMA}_{\mathsf{LOCC}}(l-1)$ as follows: The verifier asks for $l-1$ proofs, the first proof consisting of registers $AB_1 \cdots B_k$, each of size m qubits and $k = \lceil 16 \ln 2m\epsilon^{-2} \rceil$, and all the $l-2$ other proofs in systems C_j ($3 \leq j \leq l$) of size m qubits. Then he symmetrises the B systems and traces out all of them except the first. Finally he applies the original measurement from the $\mathsf{QMA}_{\mathsf{LOCC}}(l)$ protocol to the resulting state.

The completeness of the protocol is unaffected by the simulation. For the soundness let $\rho_{AB_1 \cdots B_k} \otimes \rho_{C_3} \otimes \cdots \otimes \rho_{C_l}$ be an arbitrary state sent by the prover (after the symmetrisation step has been applied to $B_1, \ldots, B_m$). Let $\{M_{C^l}, \mathbb{I} - M_{C^l}\} \in \mathsf{LOCC}$ be the l-partite LOCC verification measurement from the $\mathsf{QMA}_{\mathsf{LOCC}}(l)$ protocol (where for notational simplicity we define $C_1 = A$, $C_2 = B_1$ and $C^l = C_1 \cdots C_l$). Then

$$\operatorname{tr}\left(M_{C^l}(\rho_{C_1 C_2} \otimes \rho_{C_3} \otimes \cdots \otimes \rho_{C_l})\right) \leq \max_{\sigma_{C^l} \in \mathcal{S}_{C_1:\cdots:C_l}} \operatorname{tr}(M_{C^l}\sigma_{C^l}) \tag{123}$$

$$+ \min_{\sigma_{C^l} \in \mathcal{S}_{C_1:\cdots:C_l}} \| \rho_{C_1 C_2} \otimes \rho_{C_3} \otimes \cdots \otimes \rho_{C_l} - \sigma_{C^l} \|_{\mathsf{LOCC}}$$

$$= \max_{\sigma_{C^l} \in \mathcal{S}_{C_1:\cdots:C_l}} \operatorname{tr}(M_{C^l}\sigma_{C^l}) \tag{124}$$

$$+ \min_{\sigma_{C_1 C_2} \in \mathcal{S}_{C_1:C_2}} \| \rho_{C_1 C_2} - \sigma_{C_1 C_2} \|_{\mathsf{LOCC}}$$

$$\leq s + \epsilon. \tag{125}$$

The equality in the second line follows since we can assume that the states $\rho_{C_3}, \ldots, \rho_{C_l}$ belong to Alice and adding local states does not change the minimum LOCC-distance to separable states.

Since for going from $\mathsf{QMA}_{\mathsf{LOCC}}(l)$ to $\mathsf{QMA}_{\mathsf{LOCC}}(l-1)$ we had to blow up one of the proof's size only by a quadratic factor, we can repeat the same protocol a constant number of times and still get each proof of polynomial size. In the end, the completeness parameter of the QMA procedure is the same as the original one for $\mathsf{QMA}_{\mathsf{LOCC}}(l)$, while the soundness is smaller than $s + l\epsilon$, which can be taken to be a constant away from c by choosing ϵ sufficiently small. To reduce the soundness back to the original value s we then use the standard amplification procedure for QMA (see e.g. [60]), which works in this case since the verification measurement is LOCC [63]. $\quad\square$

References

1. Ohya, M., Petz, D.: *Quantum Entropy and Its Use*. Berlin-Heidelberg-New York: Springer-Verlag, 2004
2. Lieb, E.H., Ruskai, M.B.: Proof of the strong subadditivity of quantum-mechanical entropy. J. Math. Phys. **14**, 1938 (1973)
3. Hayden, P., Jozsa, R., Petz, D., Winter, A.: Structure of states which satisfy strong subadditivity of quantum entropy with equality. Commun. Math. Phys. **246**, 359 (2004)
4. Ibinson, B., Linden, N., Winter, A.: Robustness of quantum Markov chains. Commun. Math. Phys. **277**, 289 (2008)
5. Werner, R.F.: Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. Phys. Rev. A **40**, 4277 (1989)
6. Christandl, M., Winter, A.: "Squashed entanglement": an additive entanglement measure. J. Math. Phys. **45**, 829 (2004)
7. Tucci, R.R.: Quantum entanglement and conditional information transmission. http://arxiv.org/abs/quant-ph/9909041v2, 1999
8. Tucci, R.R.: Entanglement of distillation and conditional mutual information. http://arxiv.org/abs/quant-ph/0202144v2 2002
9. Christandl, M.: The Structure of Bipartite Quantum States - Insights from Group Theory and Cryptography. PhD thesis, Cambridge University, 2006
10. Christandl, M., Schuch, N., Winter, A.: Highly entangled states with almost no secrecy. Phys. Rev. Lett. **104**, 240405 (2010)
11. Koashi, M., Winter, A.: Monogamy of entanglement and other correlations. Phys. Rev. A **69**, 022309 (2004)
12. Peres, A.: Separability criterion for density matrices. Phys. Rev. Lett. **77**, 1413 (1996)
13. Horodecki, M., Horodecki, P., Horodecki, R.: Separability of mixed states: Necessary and sufficient conditions. Phys. Lett. A **223**, 1 (1996)
14. Horodecki, M., Horodecki, P., Horodecki, R.: Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? Phys. Rev. Lett. **80**, 5239 (1998)
15. Horodecki, R., Horodecki, P., Horodecki, M., Horodecki, K.: Quantum entanglement. Rev. Mod. Phys. **81**, 865 (2009)
16. Coffman, V., Kundu, J., Wootters, W.K.: Distributed entanglement. Phys. Rev. A **61**, 052306 (2000)
17. Matthews, W., Wehner, S., Winter, A.: Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding. Commun. Math. Phys. **291**, 813 (2009)
18. Bennett, C.H., DiVincenzo, D.P., Smolin, J.A., Wootters, W.K.: Mixed state entanglement and quantum error correction. Phys. Rev. A **54**, 3824 (1996)
19. Rains, E.M.: Rigorous treatment of distillable entanglement. Phys. Rev. A **60**, 173 (1999)

20. Devetak, I., Winter, A.: Distillation of secret key and entanglement from quantum states. Proc. Roy. Soc. Lond. Ser. A **461**, 207 (2004)
21. Horodecki, K., Horodecki, M., Horodecki, P., Oppenheim, J.: Secure key from bound entanglement. Phys. Rev. Lett. **94**, 160502 (2005)
22. Hayden, P., Horodecki, M., Terhal, B.M.: The asymptotic entanglement cost of preparing a quantum state. J. Phys. A **34**, 6891 (2001)
23. Vedral, V., Plenio, M.B., Rippin, M.A., Knight, P.L.: Quantifying entanglement. Phys. Rev. Lett. **78**, 2275 (1997)
24. Vedral, V., Plenio, M.B.: Entanglement measures and purification procedures. Phys. Rev. A **57**, 1619 (1998)
25. Vidal, G., Werner, R.F.: A computable measure of entanglement. Phys. Rev. A **65**, 032314 (2001)
26. Yang, D., Horodecki, M., Horodecki, R., Synak-Radtke, B.: Irreversibility for all bound entangled states. Phys. Rev. Lett. **95**, 190501 (2005)
27. Brandão, F.G.S.L., Plenio, M.B.: A generalization of quantum Stein's lemma. Commun. Math. Phys. **295**, 791 (2010)
28. Plenio, M.B.: Logarithmic negativity: A full entanglement monotone that is not convex. Phys. Rev. Lett. **95**, 090503 (2005)
29. Alicki, R., Fannes, M.: Continuity of quantum conditional information. J. Phys. A: Math. Gen. **37**, L55 (2004)
30. Donald, M.J., Horodecki, M.: Continuity of relative entropy of entanglement. Phys. Lett. A **264**, 257 (1999)
31. Donald, M.J., Horodecki, M., Rudolph, O.: The uniqueness theorem for entanglement measures. J. Math. Phys. **43**, 4252 (2002)
32. Shor, P.W.: Equivalence of additivity questions in quantum information theory. Commun. Math. Phys. **246**, 453 (2003)
33. Hastings, M.B.: Superadditivity of communication capacity using entangled inputs. Nature Physics **5**, 255 (2009)
34. Vollbrecht, K.G.H., Werner, R.F.: Entanglement measures under symmetry. Phys. Rev. A **64**, 062307 (2001)
35. Maurer, U.M., Wolf, S.: Unconditionally secure key agreement and the intrinsic conditional information. IEEE Trans. Inform. Theory **2**, 499 (1999)
36. Christandl, M., Renner, R., Wolf, S.: A property of the intrinsic mutual information. In: *Proc. 2003 IEEE Int. Symp. Inform. Theory*, 2003, p. 258
37. Devetak, I., Yard, J.: Exact cost of redistributing quantum states. Phys. Rev. Lett **100**, 230501 (2008)
38. Yard, J., Devetak, I.: Optimal quantum source coding with quantum information at the encoder and decoder. IEEE Trans. Inform. Theory **55**, 5339 (2009)
39. Oppenheim, J.: A paradigm for entanglement theory based on quantum communication. http://arxiv.org/abs/0801.0458v1 [quantph], 2008
40. Hudson, R.L., Moody, G.R.: Locally normal symmetric states and an analogue of de Finetti's theorem. Z. Wahrschein. Verw. Geb. **33**, 343 (1976)
41. Størmer, E.: Symmetric states of infinite tensor products of *C**-algebras. J. Funct. Anal. **3**, 48 (1969)
42. Raggio, G.A., Werner, R.F.: Quantum statistical mechanics of general mean field systems. Helv. Phys. Acta. **62**, 980 (1989)
43. Werner, R.F.: An application of Bell's inequalities to a quantum state extension problem. Lett. Math. Phys. **17**, 359 (1989)
44. König, R., Renner, R.: A de Finetti representation for finite symmetric quantum states. J. Math. Phys. **46**, 122108 (2005)
45. Christandl, M., König, R., Mitchison, G., Renner, R.: One-and-a-half quantum de Finetti theorems. Commun. Math. Phys. **273**, 473 (2007)
46. Virmani, S., Plenio, M.B.: Construction of extremal local positive operator-valued measures under symmetry. Phys. Rev. A **67**, 062308 (2003)
47. DiVincenzo, D.P., Leung, D.W., Terhal, B.M.: Quantum data hiding. IEEE Trans. Inform. Theory **48**, 580 (2002)
48. DiVincenzo, D.P., Hayden, P., Terhal, B.M.: Hiding quantum data. Found. Phys. **33**, 1629 (2003)
49. Eggeling, T., Werner, R.F.: Hiding classical data in multi-partite quantum states. Phys. Rev. Lett. **89**, 097905 (2002)
50. Hayden, P., Leung, D., Winter, A.: Aspects of generic entanglement. Commun. Math. Phys. **265**, 95 (2006)
51. Doherty, A.C., Parrilo, P.A., Spedalieri, F.M.: A complete family of separability criteria. Phys. Rev. A **69**, 022308 (2004)
52. Brandão, F.G.S.L., Vianna, R.O.: Separable multipartite mixed states - operational asymptotically necessary and sufficient conditions. Phys. Rev. Lett. **93**, 220503 (2004)

53. Ioannou, L.M.: Computational complexity of the quantum separability problem. Quant. Inform. Comp. **7**, 335 (2007)
54. Navascues, M., Owari, M., Plenio, M.B.: A complete criterion for separability detection. Phys. Rev. Lett. **103**, 160404 (2009)
55. Gurvits, L.: Classical complexity and quantum entanglement. J. Comp. Sys. Sci **69**, 448 (2004)
56. Gharibian, S.: Strong NP-hardness of the quantum separability problem. Quant. Inform. Comp. **10**, 343 (2010)
57. Beigi, S.: NP vs QMA$_{\log}$(2). Quant. Inform. Comp. **10**, 141 (2010)
58. Harrow, A., Montanaro, A.: An efficient test for product states, with applications to quantum Merlin-Arthur games. In: *Proc. Found. Comp. Sci. (FOCS)*, 2010, p. 633
59. Brandão, F.G.S.L., Christandl, M., Yard, J.: A quasipolynomial-time algorithm for the quantum separability problem. In: *Proc. ACM Symp. on Theoretical Computer Science (STOC)*, 2011, p. 343
60. Watrous, J.: Quantum computational complexity. In: *Encyclopedia of Complexity and System Science*. Berlin-Heidelberg-New York: Springer, 2009
61. Marriott, C., Watrous, J.: Quantum Arthur-Merlin games. Computational Complexity **14**, 122 (2005)
62. Beigi, S., Shor, P.W., Watrous, J.: Quantum interactive proofs with short messages. Theory of Computing **7**, 201 (2011)
63. Aaronson, S., Beigi, S., Drucker, A., Fefferman, B., Shor, P.: The power of unentanglement. Theory of Computing **5**, 1 (2009)
64. Kobayashi, H., Matsumoto, K., Yamakami, T.: Quantum Merlin-Arthur proof systems: Are multiple Merlins more helpful to Arthur? In: *Lecture Notes in Computer Science*, Volume **2906**, Berlin-Heidelberg-Newyork: Springer, 2003, p. 189
65. Brandão, F.G.S.L.: *Entanglement Theory and the Quantum Simulation of Many-Body Physics*. PhD thesis, Imperial College, 2008
66. Matsumoto, K.: Can entanglement efficiently be weakened by symmetrization? http://arxiv/org/abs/quant-ph/0511240v3, 2005
67. Horodecki, K., Horodecki, M., Horodecki, P., Oppenheim, J.: Locking entanglement measures with a single qubit. Phys. Rev. Lett. **94**, 200501 (2005)
68. Piani, M.: Relative entropy of entanglement and restricted measurements. Phys. Rev. Lett. **103**, 160504 (2009)
69. Hiai, F., Petz, D.: The proper formula for the relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**, 99 (1991)
70. Ogawa, T., Nagaoka, H.: Strong converse and Stein's lemma in the quantum hypothesis testing. IEEE Trans. Inform. Theory **46**, 2428 (2000)
71. Gühne, O., Toth, G.: Entanglement detection. Phys. Rep. **474**, 1 (2009)
72. Brandão, F.G.S.L.: Quantifying entanglement with witness operators. Phys. Rev. A **72**, 022310 (2005)
73. Brandão, F.G.S.L.: Entanglement activation and the robustness of quantum correlations. Phys. Rev. A **76**, 030301(R) (2007)
74. Synak-Radtke, B., Horodecki, M.: On asymptotic continuity of functions of quantum states. J. Phys. A: Math. Gen. **39**, 423 (2006)
75. Berta, M., Christandl, M., Renner, R.: A conceptually simple proof of the quantum reverse Shannon theorem. In: *Lecture Notes in Computer Science*, Volume **6519**, Berlin-Heidelberg-New York: Springer, 2011, p. 131
76. Berta, M., Christandl, M., Renner, R.: The quantum reverse Shannon theorem based on one-shot information theory. *Commun. Math. Phys.*, 2011. to appear, http://arxiv.org/abs/0912.3805v2
77. Brandão, F.G.S.L., Plenio, M.B.: Entanglement theory and the second law of thermodynamics. Nature Physics **4**, 873 (2008)
78. Brandão, F.G.S.L., Plenio, M.B.: A reversible theory of entanglement and its relation to the second law. Commun. Math. Phys. **295**, 829 (2010)
79. Brandão, F.G.S.L., Datta, N.: One-shot rates for entanglement manipulation under non-entangling maps. IEEE Trans. Inform. Theory **57**, 1754 (2011)
80. Brandão, F.G.S.L.: A reversible framework for resource theories. In preparation, 2011
81. Winter, A.: Coding theorem and strong converse for quantum channels. IEEE Trans. Inform. Theory **45**, 2481 (1999)
82. Ogawa, T., Nagaoka, H.: A new proof of the channel coding theorem via hypothesis testing in quantum information theory. In: *Proc. 2002 IEEE Int. Symp. Inform. Theory* 2002, p. 73
83. Renner, R.: Symmetry implies independence. Nature Physics **3**, 645 (2007)
84. Jain, R.: Distinguishing sets of quantum states. http://arxiv.org/abs/quant-ph/0506205v1, 2005
85. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Review **38**, 49 (1996)