A Generalization of Quantum Stein's Lemma

Fernando G.S.L. Brandão* and Martin B. Plenio†
Institute for Mathematical Sciences, Imperial College London, London SW7 2BW, UK and
QOLS, Blackett Laboratory, Imperial College London, London SW7 2BW, UK

Given many independent and identically-distributed (i.i.d.) copies of a quantum system described either by the state ρ or σ (called null and alternative hypotheses, respectively), what is the optimal measurement to learn the identity of the true state? In asymmetric hypothesis testing one is interested in minimizing the probability of mistakenly identifying ρ instead of σ, while requiring that the probability that σ is identified in the place of ρ is bounded by a small fixed number. Quantum Stein's Lemma identifies the asymptotic exponential rate at which the specified error probability tends to zero as the quantum relative entropy of ρ and σ.

We present a generalization of quantum Stein's Lemma to the situation in which the alternative hypothesis is formed by a family of states, which can moreover be non-i.i.d.. We consider sets of states which satisfy a few natural properties, the most important being the closedness under permutations of the copies. We then determine the error rate function in a very similar fashion to quantum Stein's Lemma, in terms of the quantum relative entropy.

Our result has two applications to entanglement theory. First it gives an operational meaning to an entanglement measure known as regularized relative entropy of entanglement. Second, it shows that this measure is faithful, being strictly positive on every entangled state. This implies, in particular, that whenever a multipartite state can be asymptotically converted into another entangled state by local operations and classical communication, the rate of conversion must be non-zero. Therefore, the operational definition of multipartite entanglement is equivalent to its mathematical definition.

I. INTRODUCTION

Hypothesis testing refers to a general set of tools in statistics and probability theory for making decisions based on experimental data from random variables. In a typical scenario, an experimentalist is faced with two possible hypotheses and must decide based on experimental observation which one was actually realized. There are two types of errors in this process, corresponding to mistakenly identifying one of the two options when the other should have been detected. A central task in hypothesis testing is the development of optimal strategies for minimizing such errors and the determination of compact formulae for the minimum error probabilities.

Substantial progress has been achieved both in the classical and quantum settings for i.i.d processes [1–14]. The non-i.i.d. case, however, has proven harder and much less is known. The main result of this paper is a particular instance of quantum hypothesis testing of non-i.i.d. sources for which the optimal separation rate can be fully determined. To the best of the authors knowledge, the complete solution of such a problem was not known even in the classical case.

Suppose we have access to a source that generates independent and identically-distributed random variables according to one of two possible probability distributions. Our aim is to decide which probability distribution is the true one. In the quantum generalization of the problem, we are faced with a source that emits several i.i.d. copies of one of two quantum states ρ and σ, and

*Electronic address: fernando.brandao@imperial.ac.uk
†Electronic address: m.plenio@imperial.ac.uk

we should decide which of them is being produced. Since the quantum setting also encompasses the classical, we will focus on the former.

In order to learn the identity of the state the observer measures a two outcome POVM $\{A_n, \mathbb{I} - A_n\}$ given n realizations of the unknown state. If he obtains the outcome associated to A_n ($\mathbb{I} - A_n$) then he concludes that the state was ρ (σ). The state ρ is seen as the null hypothesis, while σ is the alternative hypothesis. There are two types of errors:

- Type I: The observer finds that the state was σ, when in reality it was ρ. This happens with probability $\alpha_n(A_n) := \mathrm{tr}(\rho^{\otimes n}(\mathbb{I} - A_n))$.
- Type II: The observer finds that the state was ρ, when it actually was σ. This happens with probability $\beta_n(A_n) := \mathrm{tr}(\sigma^{\otimes n} A_n)$.

There are several distinct settings that might be considered, depending on the importance we attribute to the two types of errors [1–14].

In *asymmetric hypothesis testing*, the probability of type II error should be minimized to the extreme, while only requiring that the probability of type I error is bounded by a small parameter ϵ. The relevant error quantity in this case can be written as

$$\beta_n(\epsilon) := \min_{0 \leq A_n \leq \mathbb{I}} \{\beta_n(A_n) : \alpha_n(A_n) \leq \epsilon\}. \tag{1}$$

Quantum Stein's Lemma [5, 6] states that for every $0 < \epsilon < 1$,

$$\lim_{n\to\infty} -\frac{\log(\beta_n(\epsilon))}{n} = S(\rho||\sigma). \tag{2}$$

where $S(\rho||\sigma) = \mathrm{tr}(\rho(\log(\rho) - \log(\sigma)))$ is the quantum relative entropy (or quantum Kullback-Leibler divergence) of ρ and σ. This fundamental result gives a rigorous operational interpretation for the quantum relative entropy and was proven by Hiai and Petz [5] and Ogawa and Nagaoka [6]. Different proofs have since be given in Refs. [7, 8, 13]. The relative entropy is also the asymptotic optimal exponent for the decay of β_n when we require that $\alpha_n \overset{n\to\infty}{\longrightarrow} 0$ [8].

Quantum Stein's Lemma can be generalized in two natural directions. We can consider asymmetric hypothesis testing of *non-i.i.d.* states and, moreover, we can allow the two hypotheses to be composed of sets of states, instead of a single one. In this more general formulation, the problem cannot be solved in simple terms as in quantum Stein's Lemma. It is an interesting line of investigation, therefore, to study under what further assumptions the optimal error exponent can be determined in an illustrative manner.

There are several works that present extensions of quantum Stein's Lemma. Concerning non-i.i.d. sequences, in [15] Bjelaković and Siegmund-Schultze proved that quantum Stein's Lemma is also true if the null hypothesis is an ergodic state, instead of i.i.d.. Further generalizations to particular cases where the null and alternative hypotheses are correlated states were obtained in Refs. [16–18]. Finally, the *information spectrum* approach [12] delivers the achievability and strong converse optimal rate limits in terms of divergence spectrum rates for arbitrary sequence of states. Despite its generality, this method has the drawback that in general no direct connection to the quantum relative entropy is established.

Concerning extensions to sets of states as hypotheses, a generalization of quantum Stein's Lemma, sometimes referred to as quantum Sanov's Theorem, considers the situation in which the null hypotheses are i.i.d extensions of the elements of a family of states $\mathcal{K}$ [7, 19]. It was found that the rate limit of type II error is given by $\inf_{\rho\in\mathcal{K}} S(\rho||\sigma)$, which is a pleasingly direct extension

of the original result. In Ref. [16] generalizations to the case of correlated families of states as the null hypothesis were presented.

The main result of this paper has a similar flavor to the above-mentioned generalizations. We will however be interested in the case where the *alternative hypothesis* is not only composed of a single i.i.d. state, but is actually formed by a family of non-i.i.d. states satisfying certain conditions to be specified in the next section. We will then show that the regularization of the minimum quantum relative entropy over the set of states considered is the optimal rate limit for type II error.

Apart from extending the range of possibilities of the alternative hypothesis, instead of the null hypothesis, the present work differs from previous ones in the assumptions which are imposed on the set of states. Instead of ergodicity and related ideas, we consider as the alternative hypothesis sets of states satisfying five properties outlined in section II, the most important being the closedness under the permutations of the copies of the state. In this way, we will be able to employ recent advances in the characterization of quantum permutation-invariant states, more specifically the exponential de Finetti Theorem due to Renner [20, 21], to reduce the problem from the most general form to particular one closely related to the i.i.d., in which it can be tackled more easily.

The main motivation for considering these particular sets of states comes from entanglement theory [22, 23]. Given a k-partite finite dimensional Hilbert space $\mathcal{H} := \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_k$, we say that a state σ acting on $\mathcal{H}$ is separable if it can be written as

$$\sigma = \sum_j p_j \sigma_{1,j} \otimes ... \otimes \sigma_{k,j}, \tag{3}$$

for local states $\sigma_{i,j} \in \mathcal{D}(\mathcal{H}_i)$ and a probability distribution $\{p_j\}$ [24]. Assuming that the state σ is shared by k parties, each holding a quantum system described by the Hilbert space $\mathcal{H}_j$, it is clear that they can generate it from a completely uncorrelated state by *local quantum operations* on their respective particles and *classical communication* among them (LOCC). If a state cannot be created by LOCC, we say it is *entangled*. To create an entangled state from an uncorrelated state the parties must, in addition to LOCC, exchange quantum particles. As we show, the set of separable states satisfy the conditions we impose on the alternative hypothesis. Therefore, a particular instance of the problem we analyse is the discrimination of tensor powers of an entangled state from an arbitrary sequence of separable states.

Notation: We let $\mathcal{H}$ be a finite dimensional Hilbert space and $\mathcal{D}(\mathcal{H})$ the set of density operators acting on $\mathcal{H}$. Given a pure state $|\theta\rangle \in \mathcal{H}$, $\mathcal{H} \perp |\theta\rangle$ denotes the subspace of $\mathcal{H}$ orthogonal to $|\theta\rangle$. Let $\mathrm{supp}(X)$ be the support of the operator X. For two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ with $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, we define the quantum relative entropy of ρ and σ as

$$S(\rho||\sigma) := \mathrm{tr}(\rho(\log(\rho) - \log(\sigma))).$$

Given a Hermitian operator A, $||A||_1 = \mathrm{tr}(\sqrt{A^\dagger A})$ stands for the trace norm of A, $\mathrm{tr}(A)_+$ for the trace of the positive part of A, i.e. the sum of the positive eigenvalues of A, and $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ for the maximum and the minimum eigenvalue of A, respectively. For two positive semidefinite operators A, B, $F(A,B) := \mathrm{tr}(\sqrt{A^{1/2} B A^{1/2}})$ is their fidelity. The partial trace of $\rho \in \mathcal{D}(\mathcal{H}^{\otimes n})$ with respect to the j-th Hilbert space is denoted by $\mathrm{tr}_j(\rho)$, while $\mathrm{tr}_{\setminus j}(\rho)$ stands for the partial trace of all Hilbert spaces, except the j-th. We denote the binary Shannon entropy by h: $h(x) = -x\log(x) - (1-x)\log(1-x)$.

Given a subset $\mathcal{M} \subseteq \mathbb{R}^n$ we define its associate cone by $\mathrm{cone}(\mathcal{M}) := \{x : x = \lambda y, y \in \mathcal{M}, \lambda \in \mathbb{R}_+\}$ and its dual cone by $\mathcal{M}^* := \{x : y^T x \geq 0 \; \forall \, y \in \mathcal{M}\}$. We denote the ϵ-ball in trace norm around

ρ by $B_\epsilon(\rho) := \{\pi \in \mathcal{D}(\mathcal{H}) : ||\rho - \pi||_1 \leq \epsilon\}$. The Bachmann-Landau notation $g(n) = O(f(n))$ stands for $\exists k > 0, n_0 : \forall n > n_0,\ g(n) \leq kf(n)$, while $g(n) = o(f(n))$ for $\forall k > 0, \exists n_0 : \forall n > n_0,\ g(n) \leq kf(n)$.

A function E is called asymptotically continuous if there is a monotonic increasing function $f : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{x\to 0^+} f(x) = 0$ such that $\forall \rho, \sigma \in \mathcal{D}(\mathcal{H})$, $|E(\rho) - E(\sigma)| \leq \log(\dim(\mathcal{H}))f(||\rho - \sigma||_1)$.

Let $\mathrm{Sym}(\mathcal{H}^{\otimes n})$ denote the symmetric subspace of $\mathcal{H}^{\otimes n}$. For any $|\psi\rangle \in \mathcal{H}^{\otimes n}$ not orthogonal to $\mathrm{Sym}(\mathcal{H}^{\otimes n})$, we define

$$\mathrm{Sym}(|\psi\rangle) := \frac{\sum_{\pi\in S_n} P_\pi|\psi\rangle}{\left\|\sum_{\pi\in S_n} P_\pi|\psi\rangle\right\|} \tag{4}$$

where S_n is the symmetric group of order n and P_π is the representation in $\mathcal{H}^{\otimes n}$ of a permutation $\pi \in S_n$ given by $P_\pi\left(\psi_1 \otimes \psi_2 \otimes ... \otimes \psi_n\right) = \psi_{\pi^{-1}(1)} \otimes \psi_{\pi^{-1}(2)} \otimes ... \otimes \psi_{\pi^{-1}(n)}$. Finally, the symmetrization superoperator $\hat{S}_n : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{H}^{\otimes n})$ is defined as

$$\hat{S}_n(X) := \frac{1}{n!}\sum_{\pi\in S_n} P_\pi X P_\pi^*. \tag{5}$$

II. DEFINITIONS AND MAIN RESULTS

Given a set of states $\mathcal{M} \subseteq \mathcal{D}(\mathcal{H})$ we define

$$E_{\mathcal{M}}(\rho) := \inf_{\sigma\in\mathcal{M}} S(\rho||\sigma), \tag{6}$$

and

$$LR_{\mathcal{M}}(\rho) := \inf_{\sigma\in\mathcal{M}} S_{\max}(\rho||\sigma), \tag{7}$$

where

$$S_{\max}(\rho||\sigma) := \inf\{s : \rho \leq 2^s\sigma\} \tag{8}$$

is the maximum relative entropy [25]. Note that if we take $\mathcal{M}$ to be the set of separable states, $E_{\mathcal{M}}$ and $LR_{\mathcal{M}}$ reduce to two entanglement measures known as the relative entropy of entanglement [26, 27] and the logarithm global robustness of entanglement [28–31]. This connection is the reason for the nomenclature used here.

We will also need the smooth version of $LR_{\mathcal{M}}$, defined as

$$LR^\epsilon_{\mathcal{M}}(\rho) := \min_{\tilde{\rho}\in B_\epsilon(\rho)} LR_{\mathcal{M}}(\tilde{\rho}). \tag{9}$$

We note that smooth versions of other non-asymptotic-continuous measures, such as the min- and max-entropies [20, 32, 33], have been proposed and shown to be useful in non-asymptotic and non-i.i.d. information theory.

Let us specify the sets of states over which the alternative hypothesis can vary. We will consider any family of sets $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$, with $\mathcal{M}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$, satisfying the following properties

1. Each $\mathcal{M}_n$ is convex and closed.

2. Each $\mathcal{M}_n$ contains $\sigma^{\otimes n}$, for a full rank state $\sigma \in \mathcal{D}(\mathcal{H})$.

3. If $\rho \in \mathcal{M}_{n+1}$, then $\mathrm{tr}_k(\rho) \in \mathcal{M}_n$, for every $k \in \{1, ..., n+1\}$.

4. If $\rho \in \mathcal{M}_n$ and $\nu \in \mathcal{M}_m$, then $\rho \otimes \nu \in \mathcal{M}_{n+m}$.

5. If $\rho \in \mathcal{M}_n$, then $P_\pi \rho P_\pi \in \mathcal{M}_n$ for every $\pi \in S_n$.

We define the regularized version of the quantity given by Eq. (6) as

$$E^{\infty}_{\mathcal{M}}(\rho) := \lim_{n\to\infty} \frac{1}{n} E_{\mathcal{M}_n}(\rho^{\otimes n}). \tag{10}$$

To see that the limit exists in Eq. (10) we use the fact that if a sequence (a_n) satisfies $a_{n+m} \leq a_n + a_m$, then a_n/n is convergent (see e.g. Lemma 4.1.2 in [34]). Using property 4 it is easy to see that our sequence satisfies this condition.

We now turn to the main result of the paper. Suppose we have one of the following two hypothesis:

1. *Null hypothesis*: For every $n \in \mathbb{N}$ we have $\rho^{\otimes n}$ with $\rho \in \mathcal{D}(\mathcal{H})$.

2. *Alternative hypothesis*: For every $n \in \mathbb{N}$ we have an unknown state $\omega_n \in \mathcal{M}_n$, where $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ is a family of sets satisfying properties 1-5.

The next theorem gives the optimal rate limit for the type II error when one requires that type I error vanishes asymptotically.

Theorem I *Let $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ be a family of sets satisfying properties 1-5 and $\rho \in \mathcal{D}(\mathcal{H})$. Then*

(Direct part): For every $\epsilon > 0$ there exists a sequence of POVMs $\{A_n, \mathbb{I} - A_n\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} tr((\mathbb{I} - A_n)\rho^{\otimes n}) = 0 \tag{11}$$

and for every $n \in \mathbb{N}$ and $\omega_n \in \mathcal{M}_n$,

$$-\frac{\log tr(A_n\omega_n)}{n} + \epsilon \geq E^{\infty}_{\mathcal{M}}(\rho). \tag{12}$$

(Strong Converse): If a real number $\epsilon > 0$ and a sequence of POVMs $\{A_n, \mathbb{I} - A_n\}_{n\in\mathbb{N}}$ are such that for every $n \in \mathbb{N}$ and $\omega_n \in \mathcal{M}_n$,

$$-\frac{\log(tr(A_n\omega_n))}{n} - \epsilon \geq E^{\infty}_{\mathcal{M}}(\rho), \tag{13}$$

then

$$\lim_{n\to\infty} tr((\mathbb{I} - A_n)\rho^{\otimes n}) = 1. \tag{14}$$

We note that the converse part of the theorem is a so called *strong converse*, which shows that not only the probability of type I error does not go to zero when we require that type II error rate is larger than $E^{\infty}_{\mathcal{M}}$, but it actually goes to one.

Also note we can recover the original quantum Stein's Lemma by choosing $\mathcal{M}_n := \{\sigma^{\otimes n}\}$, where σ is the alternative hypothesis and ρ is the null hypothesis (Theorem I can only be applied here if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, but this is exactly the non-trivial case of quantum Stein's Lemma).

Theorem I gives an operational interpretation to the *regularized* relative entropy of entanglement [26, 27, 35], defined by

$$E_R^\infty(\rho) := \lim_{n\to\infty} \frac{1}{n} \min_{\sigma\in\mathcal{S}(\mathcal{H}^{\otimes n})} S(\rho^{\otimes n}||\sigma), \tag{15}$$

with $\mathcal{S}(\mathcal{H}^{\otimes n})$ as the set of k-partite separable states over $\mathcal{H}^{\otimes n} := \mathcal{H}_1^{\otimes n} \otimes ... \otimes \mathcal{H}_k^{\otimes n}$, where the j-th local party Hilbert space is given by $\mathcal{H}_j^{\otimes n}$. Taking $\mathcal{M}_n = \mathcal{S}(\mathcal{H}^{\otimes n})$, it is a simple exercise to check that they satisfy conditions 1-5. Therefore, we conclude that $E_R^\infty(\rho)$ gives the asymptotic rate of type II error when we try to decide if we have several realizations of ρ or a sequence of *arbitrary* separable states. This rigorously justifies the use of the regularized relative entropy of entanglement as a measure of distinguishability of quantum correlations from classical correlations, as was originally suggested on heuristic grounds in [27, 36].

On the way to prove Theorem I we establish the following alternative expression for $E_{\mathcal{M}}^\infty$.

Proposition II.1 *For every family of sets $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ satisfying properties 1-5 and every state $\rho \in \mathcal{D}(\mathcal{H})$,*

$$E_{\mathcal{M}}^\infty(\rho) = \lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} LR_{\mathcal{M}_n}^\epsilon(\rho^{\otimes n}) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} LR_{\mathcal{M}_n}^\epsilon(\rho^{\otimes n}). \tag{16}$$

Taking once more $\{\mathcal{M}_n\}$ as the sets of separable states over $\mathcal{H}^{\otimes n}$, Proposition II.1 shows that the regularized relative entropy of entanglement is a smooth asymptotic version of the log global robustness of entanglement [28–31]. Hence we have a connection between the robustness of quantum correlations under mixing and their distinguishability to classical correlations. A different, but related, proof of this fact has been found in Ref. [31].

A corollary of Theorem I is the following.

Corollary II.2 *The regularized relative entropy of entanglement is faithful. For every entangled state $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n)$,*

$$E_R^\infty(\rho) > 0. \tag{17}$$

Recently, Piani found an independent proof of Corollary II.2, using completely different techniques - most notably the insight of defining a new variant of the relative entropy of entanglement, based on the optimal distinguishability of an entangled state to separable states accessible by restricted measurements, e.g. LOCC ones [37].

Corollary II.2 has an interesting consequence to the theory of asymptotic entanglement conversion of multipartite states. Given two states $\rho, \sigma \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n)$, we define the LOCC optimal asymptotic rate of conversion of ρ into σ as

$$R(\rho\to\sigma) := \inf_{\{k_n\}_{n\in\mathbb{N}}} \left\{ \limsup_{n\to\infty} \frac{k_n}{n} : \lim_{n\to\infty} \left(\min_{\Lambda\in LOCC} ||\Lambda(\rho^{\otimes k_n}) - \sigma^{\otimes n}||_1 \right) = 0 \right\}, \tag{18}$$

where the infimum is taken over all sequences of integers $\{k_n\}_{n\in\mathbb{N}}$ and the minimization over all LOCC trace preserving maps Λ. We are therefore interested in the most efficient manner to transform a given entangled state into another, in the regime of many copies, when we only have access to LOCC.

A fundamental question in this context is whether the rate $R(\rho\to\sigma)$ is non-zero whenever σ is entangled. For states composed of two parties, the work of Yang *et al* [38] has provided the answer in the affirmative. The general case of multipartite states, however, remained open. A direct

application of Corollary II.2 shows that indeed the rate function is strictly positive whenever the target state is entangled. We thus find that the mathematical definition of entanglement, as states that cannot be written as in Eq. (3), is equivalent to an operational definition of entangled states, as states which require a non-zero rate of entangled pure states - or any other fixed entangled state in fact - for their formation in the asymptotic limit.

Corollary II.3 *For every two entangled states $\rho, \sigma \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n)$,*

$$R(\rho \to \sigma) > 0. \tag{19}$$

Another application of our main theorem is given in the follow up paper [39] (see also [40, 41]). There, Theorem III.10 is the key technical tool to prove reversibility in the asymptotic manipulation of entangled states under quantum operations which cannot (approximately) generate entanglement.

In the next three sections we provide the proofs of Theorem I, Proposition II.1, Corollary II.2, and Corollary II.3.

III. PROOF OF THEOREM I

We start proving Proposition II.1 and then use it to establish the following auxiliary result.

Proposition III.1 *For every family of sets $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ satisfying properties 1-5 and every state $\rho \in \mathcal{D}(\mathcal{H})$,*

$$\lim_{n\to\infty} \min_{\omega_n \in \mathcal{M}_n} tr(\rho^{\otimes n} - 2^{yn}\omega_n)_+ = \begin{cases} 0, & y > E^\infty_{\mathcal{M}}(\rho), \\ 1, & y < E^\infty_{\mathcal{M}}(\rho). \end{cases} \tag{20}$$

Before proving Propositions II.1 and III.1, let us show how Proposition III.1 implies Theorem I.

Proof (Theorem I) Consider the following family of convex optimization problems

$$\lambda_n(\pi, K) := \max_A \left[\text{tr}(A\pi) : 0 \leq A \leq \mathbb{I}, \;\; \text{tr}(A\sigma) \leq \frac{1}{K} \;\; \forall\, \sigma \in \mathcal{M}_n \right]. \tag{21}$$

The statement of Theorem I is immediately implied by

$$\lim_{n\to\infty} \lambda_n(\rho^{\otimes n}, 2^{ny}) = \begin{cases} 0, & y > E^\infty_{\mathcal{M}}(\rho), \\ 1, & y < E^\infty_{\mathcal{M}}(\rho). \end{cases} \tag{22}$$

In order to see that Eq. (22) holds true, we go to the dual formulation of $\lambda_n(\pi, K)$. We first rewrite it as

$$\lambda_n(\pi, K) := \max_A \left[\text{tr}(A\pi) : 0 \leq A \leq \mathbb{I}, \;\; \text{tr}\left((\mathbb{I}/K - A)\sigma\right) \geq 0 \;\; \forall\, \sigma \in \text{cone}(\mathcal{M}_n) \right], \tag{23}$$

where $\text{cone}(\mathcal{M}_n)$ is the cone of $\mathcal{M}_n$. Then, we note that the second constraint is a generalized inequality (since the set $\text{cone}(\mathcal{M}_n)$ is a convex proper cone) [42] and write the problem as

$$\lambda_n(\pi, K) := \max_A \left[\text{tr}(A\pi) : 0 \leq A \leq \mathbb{I}, \;\; (\mathbb{I}/K - A) \in (\mathcal{M}_n)^* \right], \tag{24}$$

where $(\mathcal{M}_n)^*$ is the dual cone of $\mathcal{M}_n$. The Lagrangian of $\lambda_n(\pi, K)$ is given by

$$L(\pi, K, A, X, Y, \mu) = \text{tr}(A\pi) + \text{tr}(XA) + tr(Y(\mathbb{I} - A)) + \text{tr}((\mathbb{I}/K - A)\mu), \tag{25}$$

where $X, Y \geq 0$ and $\mu \in \text{cone}(\mathcal{M}_n)$ are Lagrange multipliers. It is easy to find a strictly feasible solution for the primal optimization problem given by Eq. (24) (e.g. $A = \mathbb{I}/(2K)$). Therefore, by Slater's condition [42], $\lambda_n(\pi, K)$ is equal to its dual formulation, which reads

$$\lambda_n(\pi, K) = \min_{Y,\mu} \left[\text{tr}(Y) + \text{tr}(\mu)/K : \pi \leq Y + \mu, \;\; Y \geq 0, \;\; \mu \in \text{cone}(\mathcal{M}_n)\right]. \tag{26}$$

Using that $\text{tr}(A)_+ = \min_Y \text{tr}(Y) : Y \geq 0, Y \geq A$, we find

$$\lambda_n(\pi, K) = \min_{\mu} \left[\text{tr}(\pi - \mu)_+ + \text{tr}(\mu)/K : \mu \in \text{cone}(\mathcal{M}_n)\right], \tag{27}$$

which can finally be rewritten as

$$\lambda_n(\pi, K) = \min_{\mu,b} \left[\text{tr}(\pi - b\mu)_+ + b/K : \mu \in \mathcal{M}_n, \;\; b \in \mathbb{R}_+\right]. \tag{28}$$

Let us consider the asymptotic behavior of $\lambda_n(\rho^{\otimes n}, 2^{ny})$. Take $y = E^\infty_{\mathcal{M}}(\rho) + \epsilon$, for any $\epsilon > 0$. Then we can choose $b = 2^{n(E^\infty_{\mathcal{M}}(\rho) + \frac{\epsilon}{2})}$, giving

$$\lambda_n(\rho^{\otimes n}, 2^{ny}) \leq \min_{\mu \in \mathcal{M}_n} \left[\text{tr}(\rho^{\otimes n} - 2^{n(E^\infty_{\mathcal{M}}(\rho) + \frac{\epsilon}{2})}\mu)_+ + 2^{-n\frac{\epsilon}{2}}\right]. \tag{29}$$

From Proposition III.1 we then find that $\lambda_n(\rho^{\otimes n}, 2^{ny}) \to 0$.

We now take $y = E^\infty_{\mathcal{M}}(\rho) - \epsilon$, for any $\epsilon > 0$. The optimal b for each n has to satisfy $b_n \leq 2^{yn}$, otherwise $\lambda_n(\rho^{\otimes n}, 2^{ny})$ would be larger than one, which we know is false. Therefore,

$$\lambda_n(\rho^{\otimes n}, 2^{ny}) \geq \min_{\mu \in \mathcal{M}_n} \text{tr}(\rho^{\otimes n} - 2^{n(E^\infty_{\mathcal{M}}(\rho) - \epsilon)}\mu)_+, \tag{30}$$

which approaches unity again by Proposition III.1. □

A. Proof of Proposition II.1

Proof (Proposition II.1)

We start showing that

$$E^\infty_{\mathcal{M}}(\rho) \leq \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} LR^\epsilon_{\mathcal{M}_n}(\rho^{\otimes n}). \tag{31}$$

Let $\rho^\epsilon_n \in B_\epsilon(\rho^{\otimes n})$ be an optimal state for $\rho^{\otimes n}$ in Eq. (9). For every n there is a state $\sigma_n \in \mathcal{M}_n$ such that $\rho^\epsilon_n \leq s_n \sigma_n$, with $LR^\epsilon_{\mathcal{M}_n}(\rho^{\otimes n}) = LR_{\mathcal{M}_n}(\rho^\epsilon_n) = \log(s_n)$. It follows from the operator monotonicity of the $\log$ function [43] that if $\rho \leq 2^k \sigma$ (where ρ and σ are two states), then $S(\rho||\sigma) \leq k$. Hence,

$$\frac{1}{n} E_{\mathcal{M}_n}(\rho^\epsilon_n) \leq \frac{1}{n} S(\rho^\epsilon_n||\sigma_n) \leq \frac{1}{n} \log s_n = \frac{1}{n} LR_{\mathcal{M}_n}(\rho^\epsilon_n) = \frac{1}{n} LR^\epsilon_{\mathcal{M}_n}(\rho^{\otimes n}). \tag{32}$$

As $\rho^\epsilon_n \in B_\epsilon(\rho^{\otimes n})$, we find from Lemma C.3 (see appendix C) that

$$\frac{1}{n} E_{\mathcal{M}_n}(\rho^{\otimes n}) \leq \frac{1}{n} LR^\epsilon_{\mathcal{M}_n}(\rho^{\otimes n}) + f(\epsilon), \tag{33}$$

where $f : \mathbb{R} \to \mathbb{R}$ is such that $\lim_{\epsilon\to 0} f(\epsilon) = 0$. Taking the limits $n \to \infty$ and $\epsilon \to 0$ in both sides of the equation above,

$$E^{\infty}_{\mathcal{M}}(\rho) = \liminf_{n\to\infty} \frac{1}{n} E_{\mathcal{M}_n}(\rho^{\otimes n}) \leq \lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} LR^{\epsilon}_{\mathcal{M}_n}(\rho^{\otimes n}). \tag{34}$$

To show the converse inequality, namely that

$$E^{\infty}_{\mathcal{M}}(\rho) \geq \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} LR^{\epsilon}_{\mathcal{M}_n}(\rho^{\otimes n}), \tag{35}$$

let $y_k := E_{\mathcal{M}_k}(\rho^{\otimes k}) + \varepsilon = S(\rho^{\otimes k}||\sigma_k) + \varepsilon$ (σ_k is an optimal state for $\rho^{\otimes k}$ in $E_{\mathcal{M}_k}(\rho^{\otimes k})$) with $\varepsilon > 0$. We can write for every $n \in \mathbb{N}$,

$$\rho^{\otimes kn} \leq 2^{y_k n} \sigma_k^{\otimes n} + (\rho^{\otimes kn} - 2^{y_k n} \sigma_k^{\otimes n})_+. \tag{36}$$

From Lemma C.4 (see appendix C) we have

$$\lim_{n\to\infty} tr(\rho^{\otimes kn} - 2^{y_k n} \sigma_k^{\otimes n})_+ = 0. \tag{37}$$

Applying Lemma C.5 (see appendix C) to Eq. (36) we then find that there is a sequence of states $\rho_{n,k}$ such that

$$\lim_{n\to\infty} ||\rho^{\otimes kn} - \rho_{n,k}||_1 = 0 \tag{38}$$

and

$$\rho_{n,k} \leq g(n) 2^{y_k n} \sigma_k^{\otimes n}, \tag{39}$$

where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is such that $\lim_{n\to\infty} g(n) = 1$. It follows that for every $\delta > 0$ there is a sufficiently large n_0 such that for all $n \geq n_0$, $\rho_{n,k} \in B_\delta(\rho^{\otimes kn})$. Moreover, from property 4 of the sets we find $\sigma_k^{\otimes n} \in \mathcal{M}_{kn}$. Hence, for every $\delta > 0$,

$$\limsup_{n\to\infty} \frac{LR^{\delta}_{\mathcal{M}_{nk}}(\rho^{\otimes nk})}{n} \leq \limsup_{n\to\infty} \frac{LR_{\mathcal{M}_{kn}}(\rho_{n,k})}{n} \leq y_k = E_{\mathcal{M}_k}(\rho^{\otimes k}) + \varepsilon. \tag{40}$$

The next step is to note that for every $k \in \mathbb{N}$,

$$\limsup_{n\to\infty} \frac{1}{nk} LR^{\delta}_{\mathcal{M}_{nk}}(\rho^{\otimes nk}) = \limsup_{n\to\infty} \frac{1}{n} LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n}). \tag{41}$$

The $\leq$ inequality follows straightforwardly. For the $\geq$ inequality, let $\{n'\}$ be a subsequence such that

$$M := \lim_{n'\to\infty} \frac{1}{n'} LR^{\delta}_{\mathcal{M}_{n'}}(\rho^{\otimes n'}) \tag{42}$$

is equal to the R.H.S. of Eq. (41). Let n'_k be the first multiple of k larger than n'. Then,

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{nk} LR^{\delta}_{\mathcal{M}_{nk}}(\rho^{\otimes nk}) &\geq \limsup_{n'_k\to\infty} \frac{1}{n'_k} LR^{\delta}_{\mathcal{M}_{n'_k}}(\rho^{\otimes n'_k}) \\ &\geq \limsup_{n'_k\to\infty} \frac{1}{n'_k} LR^{\delta}_{\mathcal{M}_{n'}}(\rho^{\otimes n'}) \\ &= M. \end{aligned} \tag{43}$$

The last inequality follows from $LR^{\delta}_{\mathcal{M}_n}(\pi) \geq LR^{\delta}_{\mathcal{M}_{n-l}}(\mathrm{tr}_{1,..l}(\pi))$, which is a consequence of property 3 of the sets.

From Eq. (40) and the fact that $\varepsilon, \delta > 0$ are arbitrary, it follows that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n}) \leq \frac{1}{k} E_{\mathcal{M}_k}(\rho^{\otimes k}). \tag{44}$$

Finally, since the above equation is true for every $k \in \mathbb{N}$, we find the announced result. □

There is another related quantity that we might consider in this context, in which ϵ and n are not independent. Define

$$LG_{\mathcal{M}}(\rho) := \inf_{\{\epsilon_n\}} \left\{ \limsup_{n \to \infty} \frac{1}{n} LR^{\epsilon_n}_{\mathcal{M}_n}(\rho^{\otimes n}) : \lim_{n \to \infty} \epsilon_n = 0 \right\}. \tag{45}$$

The proof of Proposition II.1 can be straightforwardly adapted to show

Corollary III.2 *For every family of sets $\{\mathcal{M}_n\}_{n \in \mathbb{N}}$ satisfying properties 1-5 and every quantum state $\rho \in \mathcal{D}(\mathcal{H})$,*

$$LG_{\mathcal{M}}(\rho) = E^{\infty}_{\mathcal{M}}(\rho). \tag{46}$$

With Proposition II.1 at hand we are now in position to prove the strong converse part of Proposition III.1, which we restate as a separate corollary for the sake of clarity.

Corollary III.3 *Let $\rho \in \mathcal{D}(\mathcal{H})$. For every $y > E^{\infty}_{R}(\rho)$*

$$\lim_{n \to \infty} \min_{\omega_n \in \mathcal{M}_n} tr(\rho^{\otimes n} - 2^{yn} \omega_n)_+ = 0, \tag{47}$$

while for every $y < E^{\infty}_{\mathcal{M}}(\rho)$,

$$\liminf_{n \to \infty} \min_{\omega_n \in \mathcal{M}_n} tr(\rho^{\otimes n} - 2^{yn} \omega_n)_+ > 0, \tag{48}$$

Proof

We first show that if $y = E^{\infty}_{\mathcal{M}}(\rho) + \epsilon$, then

$$\lim_{n \to \infty} \min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho^{\otimes n} - 2^{yn} \omega_n)_+ = 0. \tag{49}$$

By Proposition II.1 there is a $\delta_0 > 0$ such that

$$\left| E^{\infty}_{\mathcal{M}}(\rho) - \limsup_{n \to \infty} \frac{1}{n} LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n}) \right| \leq \epsilon/2, \tag{50}$$

for every $\delta \leq \delta_0$. Let $\rho_{n,\delta} \in B_{\delta}(\rho^{\otimes n})$ be an optimal state in Eq. (9) for $\rho^{\otimes n}$ realizing the value $LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n})$. Then there must exist a $\sigma_n \in \mathcal{M}_n$ such that

$$\rho_{n,\delta} \leq 2^{LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n})} \sigma_n, \tag{51}$$

from which follows that for every $\lambda \geq LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n})/n$,

$$\min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho^{\otimes n} - 2^{\lambda n} \omega_n)_+ \leq \min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho_{n,\delta} - 2^{\lambda n} \omega_n)_+ + \delta \leq \delta. \tag{52}$$

From Eq. (50) and our choice of y we then find that for every $\delta > 0$ there is a sufficiently large n_0 such that for all $n \geq n_0$,

$$\min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho^{\otimes n} - 2^{yn}\omega_n)_+ \leq \delta, \tag{53}$$

from which Eq. (49) follows.

Now we move to the second part of the proof which aims to show that that if $y = E^{\infty}_{\mathcal{M}}(\rho) - \epsilon$, then

$$\liminf_{n\to\infty} \min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho^{\otimes n} - 2^{yn}\omega_n)_+ > 0. \tag{54}$$

To this end, let us assume by means of a contradiction that this is not the case and that the limit is zero. For each n we have

$$\rho^{\otimes n} \leq 2^{yn}\omega_n + (\rho^{\otimes n} - 2^{yn}\omega_n)_+, \tag{55}$$

where ω_n is the optimal state in $\mathcal{M}_n$ in Eq. (54). Applying Lemma C.5 to Eq. (55) we then find that there is a sequence of states $\tilde{\rho}_n$ (for an increasing subsequence $\mathcal{F} \subseteq \mathbb{N}$, $\{n\}_{n\in\mathcal{F}}$ such that $||\rho^{\otimes n} - \tilde{\rho}_n||_1 \to 0$ and $\tilde{\rho}_n \leq g(n)2^{yn}\omega_n$, for a function g satisfying $\lim_{n\to\infty} g(n) = 1$. It follows that

$$\frac{1}{n} LR_{\mathcal{M}_n}(\tilde{\rho}_n) \leq y + \frac{\log g(n)}{n} \tag{56}$$

and that for every $\delta > 0$ and sufficiently large n, $\tilde{\rho}_n \in B_\delta(\rho^{\otimes n})$. Therefore, for every $\delta > 0$,

$$\liminf_{n\to\infty} \frac{1}{n} LR^{\delta}_{\mathcal{M}_n}(\rho^{\otimes n}) \leq \liminf_{n\to\infty} \frac{1}{n} LR_{\mathcal{M}_n}(\tilde{\rho}_n) \leq y = E^{\infty}_{\mathcal{M}}(\rho) - \epsilon, \tag{57}$$

in contradiction to Eq. (16) of Proposition II.1. □

B. Proof of the direct part of Proposition III.1

We now turn to the proof of the direct part of Proposition III.1, which is the main technical contribution of the paper. Before we start with the proof in earnest, we provide a rough outline of the main steps which will be taken, in order to make the presentation more transparent.

In Corollary III.3 we showed by relatively simple means that $E^{\infty}_{\mathcal{M}}(\rho)$ is the strong converse rate for the hypothesis testing problem which we are analysing. It is more involved to show that $E^{\infty}_{\mathcal{M}}(\rho)$ is also an achievable rate, i.e. that the limit equals unity for every $y < E^{\infty}_{\mathcal{M}}(\rho)$. The difficulty is precisely that the alternative hypothesis is non-i.i.d. and is a set of states, instead of a single one in general. Most of the proof is devoted to circumvent this problem. The main ingredient of the proof is a variant of Renner's exponential version of the quantum de Finetti theorem [20, 21] (see Appendix B), given in Lemma III.5.

Loosely speaking, we will proceed as follows. We will show the reverse implication that if

$$\min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\rho^{\otimes n} - 2^{yn}\omega_n)_+ \overset{n\to\infty}{\longrightarrow} \mu < 1 \tag{58}$$

then $y \geq E^{\infty}_{\mathcal{M}}(\rho) - o(1)$. To this aim we first use Lemma C.5 (see appendix C) to find from the equation above a state ρ_n that possesses non-negligible fidelity with $\rho^{\otimes n}$ and satisfies

$$\rho_n \leq 2^{yn+o(n)}\omega_n, \tag{59}$$

for every n, where $\omega_n \in \mathcal{M}_n$ is the optimal state in the minimization of Eq. (58). Due to property 5 of the sets, we can take ω_n and thus also ρ_n to be permutation-symmetric. Then, tracing a sublinear number of copies $o(n)$ and using Lemmata III.4 and III.5 we will be able to show that the previous equation implies that there is a state $\pi_{\rho,n}$ exponential close to an almost power state along ρ (see Eq. (67) for a definition) such that

$$\pi_{\rho,n} \leq 2^{yn+o(n)} \mathrm{tr}_{1,\ldots,o(n)}(\omega_n). \tag{60}$$

In a second part of the proof, we will argue that the measure $E_{\mathcal{M}_n}(\pi_{\rho,n})$ is not too far away from $E_{\mathcal{M}_n}(\rho^{\otimes n})$, with the difference being upper bounded by a term sublinear in n. This property can be considered as a manifestation of the non-lockability of the measures $E_{\mathcal{M}_n}$, as was proved for the relative entropy of entanglement in Ref. [44].

Finally, using the operator monotonicity of the $\log$ and the asymptotic continuity of both $E_{\mathcal{M}_k}$ and $E^{\infty}_{\mathcal{M}}$ (see Appendix C), we will find from Eq. (60) that, for sufficiently large n,

$$E^{\infty}_{\mathcal{M}}(\rho) = \frac{1}{n} E_{\mathcal{M}_{n-o(n)}}(\pi_{\rho,n}) + o(1) \leq y + o(1). \tag{61}$$

The next lemma is an extension of Uhlmann's theorem on the fidelity [45] to the case of tensor product and symmetric states.

Lemma III.4 *Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$ be such that $\hat{S}_n(\rho_n) = \rho_n$. Then there is a purification $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}$ of ρ and a permutation-symmetric purification $|\Psi_n\rangle \in (\mathcal{H} \otimes \mathcal{H})^{\otimes n}$ of ρ_n such that $|\langle \Psi_n | \theta^{\otimes n} \rangle| = F(\rho_n, \rho^{\otimes n})$.*

Proof Let $|\phi^+\rangle := \sum_{k=1}^{\dim(\mathcal{H})} |k,k\rangle$ and consider the following purifications of ρ and ρ_n, respectively: $|\theta\rangle = \mathbb{I} \otimes \sqrt{\rho}|\phi^+\rangle$ and $|\Psi_n\rangle = \mathbb{I}^{\otimes n} \otimes (\sqrt{\rho_n} U)|\phi^+\rangle^{\otimes n}$, where the unitary U is a particular unitary, to be specified in the next paragraph, such that $\sqrt{\rho_n}\sqrt{\rho^{\otimes n}} = U|\sqrt{\rho_n}\sqrt{\rho^{\otimes n}}|$ [43]. A direct calculation shows that $|\langle \Psi_n | \theta^{\otimes n} \rangle| = F(\rho_n, \rho^{\otimes n})$.

To see that $|\Psi_n\rangle$ is permutation-symmetric, we note that as $\rho^{\otimes n}$ and ρ_n are permutation-invariant, we can take U and thus $\sqrt{\rho_n} U$ to be invariant under permutations too. Indeed, as $\sqrt{\rho_n}\sqrt{\rho^{\otimes n}}$ and $|\sqrt{\rho_n}\sqrt{\rho^{\otimes n}}|$ are permutation invariant, we can write them in the Schur basis [46] as

$$\sqrt{\rho_n}\sqrt{\rho^{\otimes n}} = \bigoplus_{\lambda} A_\lambda \otimes \mathbb{I}_\lambda, \;\; |\sqrt{\rho_n}\sqrt{\rho^{\otimes n}}| = \bigoplus_{\lambda} B_\lambda \otimes \mathbb{I}_\lambda, \tag{62}$$

where λ labels the irreps of S_n, $\mathbb{I}_\lambda$ is the identity on the irrep labelled by λ, and A_λ, B_λ are operators acting on the multiplicity space of the the irrep labelled by λ [46]. We can define the partial isometry V as

$$V := \sqrt{\rho_n}\sqrt{\rho^{\otimes n}}|\sqrt{\rho_n}\sqrt{\rho^{\otimes n}}|^{-1} = \bigoplus_{\lambda} A_\lambda B_\lambda^{-1} \otimes \mathbb{I}_\lambda, \tag{63}$$

where the inverses are taken in the generalized sense. As each $A_\lambda B_\lambda^{-1}$ is a partial isometry, we can extend them to unitaries U_λ. Then we set

$$U := \bigoplus_{\lambda} U_\lambda \otimes \mathbb{I}_\lambda, \tag{64}$$

which is clearly permutation-invariant.

Finally, for every permutation $\pi \in S_n$,

$$P_\pi|\Psi_n\rangle = P_{\pi,S} \otimes P_{\pi,E}(\mathbb{I} \otimes \sqrt{\rho_n}U)|\phi^+\rangle^{\otimes n} = \mathbb{I} \otimes (P_{\pi,E}\sqrt{\rho_n}UP_{\pi,E})(P_{\pi,S} \otimes P_{\pi,E})|\phi^+\rangle^{\otimes n} = |\Psi_n\rangle. \quad (65)$$

□

The next lemma can be seen as a post-selected variant of the exponential de Finetti theorem [20, 21] and is proved by similar techniques. For a $|\theta\rangle \in \mathcal{H}$ and $0 \leq r \leq n$ we define the set of $\binom{n}{r}$-i.i.d states in $|\theta\rangle$ as

$$\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r}) := \{P_\pi(|\theta\rangle^{\otimes n-r} \otimes |\psi_r\rangle) : \pi \in S_n, |\psi_r\rangle \in \mathcal{H}^{\otimes r}\}. \quad (66)$$

Thus for every state in $\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})$ we have the state $|\theta\rangle$ in at least $n-r$ of the copies. The set of almost power states in $|\theta\rangle$ is defined as [47, 48]

$$|\theta\rangle^{[\otimes,n,r]} := \mathrm{Sym}(\mathcal{H}^{\otimes n}) \cap \mathrm{span}(\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})). \quad (67)$$

Finally, we say a mixed state $\rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$ is an almost power state along $\sigma \in \mathcal{D}(\mathcal{H})$, if there is a purification of ρ_n, $|\psi\rangle \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}_E^{\otimes n}$, where $\mathcal{H}_E \cong \mathcal{H}$ is the purifying Hilbert space, such that $|\psi\rangle \in |\theta\rangle^{[\otimes,n,r]}$, for some purification $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}_E$ of σ.

Lemma III.5 *Let $|\Psi_n\rangle \in \mathcal{H}^{\otimes n}$ be a permutation-invariant state and $|\theta\rangle \in \mathcal{H}$. Then for every $m \leq n$ there is a state $|\Psi_{n,m}\rangle \in \mathcal{H}^{\otimes n-m}$ such that*

$$|\Psi_{n,m}\rangle\langle\Psi_{n,m}| \leq |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-2} tr_{1,\ldots,m}(|\Psi_n\rangle\langle\Psi_n|), \quad (68)$$

and for every $r \leq n-m$

$$|||\Psi_{n,m}\rangle\langle\Psi_{n,m}| - |\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|||_1 \leq 2\sqrt{2}|\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1} e^{-\frac{mr}{2n}} \quad (69)$$

for an almost power state $|\Psi_{n,m,r}\rangle \in |\theta\rangle^{[\otimes,n-m,r]}$.

Proof We write $|\Psi_n\rangle = \langle\theta^{\otimes n}|\Psi_n\rangle|\theta\rangle^{\otimes n} + \sqrt{1-|\langle\theta^{\otimes n}|\Psi_n\rangle|^2}|\Phi_n\rangle$, where $|\Phi_n\rangle$ is a permutation-symmetric state orthogonal to $|\theta\rangle^{\otimes n}$. We can expand $|\Phi_n\rangle$ as $|\Phi_n\rangle = \sum_{k=1}^n \beta_k \mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-k})$, where $|\eta_k\rangle$ are permutation-symmetric states which live in $(\mathcal{H} \perp |\theta\rangle)^{\otimes k}$ and $\sum_k |\beta_k|^2 = 1$.

Define $|\Psi_{n,m}\rangle := (\langle\theta|^{\otimes m} \otimes \mathbb{I}^{\otimes n-m})|\Psi_n\rangle / ||(\langle\theta|^{\otimes m} \otimes \mathbb{I}^{\otimes n-m})|\Psi_n\rangle||$. From the inequality

$$||(\langle\theta|^{\otimes m} \otimes \mathbb{I}^{\otimes n-m})|\Psi_n\rangle|| := \langle\Psi_n|(|\theta\rangle\langle\theta|)^{\otimes m} \otimes \mathbb{I}^{\otimes n-m}|\Psi_n\rangle^{1/2} \geq |\langle\Psi_n|\theta^{\otimes n}\rangle| \quad (70)$$

we find

$$\begin{aligned} |\Psi_{n,m}\rangle\langle\Psi_{n,m}| &\leq ||(\langle\theta|^{\otimes m} \otimes \mathbb{I}^{\otimes n-m})|\Psi_n\rangle||^{-2} \mathrm{tr}_{1,\ldots,m}(|\Psi_n\rangle\langle\Psi_n|) \\ &\leq |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-2} \mathrm{tr}_{1,\ldots,m}(|\Psi_n\rangle\langle\Psi_n|). \end{aligned} \quad (71)$$

To estimate how close $|\Psi_{n,m}\rangle$ is to an almost power state, we make use of the following relation, valid for every $m \leq n$,

$$(\langle\theta|^{\otimes m} \otimes \mathbb{I}^{\otimes n-m})\mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-k}) = \binom{n}{k}^{-1/2} \binom{n-m}{k}^{1/2} \mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-k-m}). \quad (72)$$

Define

$$|\Psi'_{n,m,r}\rangle := ||(\langle\theta|^{\otimes m}\otimes\mathbb{I}^{\otimes n-m})|\Psi_n\rangle||^{-1}(\langle\Psi_n|\theta^{\otimes n}\rangle|\theta\rangle^{\otimes n-m} + \sqrt{1-|\langle\Psi_n|\theta^{\otimes n}\rangle|^2}\sum_{k=1}^{r}\beta_k\binom{n}{k}^{-1/2}\binom{n-m}{k}^{1/2}\mathrm{Sym}(|\eta_k\rangle\otimes|\theta\rangle^{\otimes n-k-m})). \quad (73)$$

Note that $|\Psi'_{n,m,n}\rangle = |\Psi_{n,m}\rangle$. Then, from Eq. (70),

$$\begin{aligned}|||\Psi'_{n,m,r}\rangle - |\Psi_{n,m}\rangle|| &\leq |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1}\|\sum_{k=r+1}^{n}\beta_k\binom{n}{k}^{-1/2}\binom{n-m}{k}^{1/2}\mathrm{Sym}(|\eta_k\rangle\otimes|\theta\rangle^{\otimes n-k-m})\| \\ &= |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1}\left(\sum_{k=r+1}^{n}|\beta_k|^2\binom{n}{k}^{-1}\binom{n-m}{k}\right)^{\frac{1}{2}}. \quad (74)\end{aligned}$$

We have

$$\begin{aligned}\binom{n}{k}^{-1}\binom{n-m}{k} &= \frac{(n-m)(n-m-1)...(n-m-k+1)}{n(n-1)...(n-k+1)} \\ &= \left(1-\frac{m}{n}\right)...\left(1-\frac{m}{n-k+1}\right) \\ &\leq \left(1-\frac{m}{n}\right)^k \leq e^{-\frac{mk}{n}}. \quad (75)\end{aligned}$$

where we used that for $\beta\in(0,1]$, $(1-\beta)^{1/\beta}\leq e^{-1}$. Hence

$$|||\Psi'_{n,m,r}\rangle - |\Psi_{n,m}\rangle|| \leq |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1}\left(\sum_{k=r+1}^{n}e^{-\frac{mk}{n}}|\beta_k|^2\right)^{\frac{1}{2}} \leq |\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1}e^{-\frac{mr}{2n}}, \quad (76)$$

where in the last inequality we used that $\sum_{k=r+1}^{n}|\beta|_k^2\leq 1$.

Defining $|\Psi_{n,m,r}\rangle := |\Psi'_{n,m,r}\rangle/|||\Psi'_{n,m,r}\rangle||$, we have $|||\Psi_{n,m,r}\rangle - |\Psi_{n,m}\rangle|| \leq 2|||\Psi'_{n,m,r}\rangle - |\Psi_{n,m}\rangle|| \leq 2|\langle\Psi_n|\theta^{\otimes n}\rangle|^{-1}e^{-\frac{mr}{2n}}$, where we used the estimate

$$\left\|\frac{x}{\|x\|}-y\right\| \leq \|x-y\| + \|\frac{x}{\|x\|}-x\| = \|x-y\|+1-\|x\| = \|x-y\|+\|y\|-\|x\| \leq 2\|x-y\|, \quad (77)$$

with $x := |\Psi'_{n,m,r}\rangle$ and $y := |\Psi_{n,m}\rangle$.

The lemma is now a consequence of the inequality $|||\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|||_1 \leq \sqrt{\langle\psi|\psi\rangle + \langle\phi|\phi\rangle}|||\psi\rangle - |\phi\rangle||$ (see e.g. Lemma A.2.5 of [20]). □

The next lemma is an analogue of a result of Ogawa and Nagaoka [6], stated in Appendix C as Lemma C.4, and originally used to establish the strong converse of quantum Stein's lemma.

Lemma III.6 *Given two states $\rho,\sigma\in\mathcal{D}(\mathcal{H})$ such that supp$(\rho)\subseteq$ supp(σ) and real numbers λ,μ,*

$$tr(\rho^{\otimes n}-2^{\lambda n}\sigma^{\otimes n})_+ \leq 2^{-n(s\mu-\log tr(\rho^{1+s}))} + 2^{-n(s(\lambda-\mu)-s\dim(\mathcal{H})\frac{\log(1+n)}{n}-\log tr(\rho\sigma^{-s}))}. \quad (78)$$

for every $s\in[0,1]$.

Proof Let Q_n be the projector onto the positive part of $(\rho^{\otimes n} - 2^{\lambda n}\sigma^{\otimes n})$. Let $Q_n = \sum_i \lambda_i E_i$ be an eigen-decomposition of Q_n with eigenvalues λ_i (either equal to 0 or 1) and eigen-projectors $\{E_i\}$ whose particular form will be specified later on in the proof.

Define the probability distributions $p_n(i) := \mathrm{tr}(\rho^{\otimes n} E_i)$ and $q_n(i) := \mathrm{tr}(\sigma^{\otimes n} E_i)$. From Lemma C.7 we can write

$$\begin{aligned} \mathrm{tr}(\rho^{\otimes n} - 2^{\lambda n}\sigma^{\otimes n})_+ &= \sum_i \lambda_i \left(p_n(i) - 2^{\lambda n} q_n(i) \right) \\ &\leq \Pr_{\{p_n\}} \left(i : \frac{1}{n} \log \frac{p_n(i)}{q_n(i)} > \lambda \right) \\ &\leq \Pr_{\{p_n\}} \left(i : \frac{1}{n} \log p_n(i) \geq \mu \right) + \Pr_{\{p_n\}} \left(i : -\frac{1}{n} \log q_n(i) \geq \lambda - \mu \right) \end{aligned} \tag{79}$$

for every $\mu \in \mathbb{R}$. Given a discrete probability distribution r, a random variable X, and a real number a, Cramér Theorem gives [49]

$$\Pr_{\{r\}}(X \geq a) \leq 2^{-\Lambda(X,r,a)}, \quad \Lambda(X, r, a) := \sup_{0 \leq s \leq 1} \left(as - \log \sum_i r(i) 2^{sX(i)} \right) \tag{80}$$

Applying it to the two last terms of Eq. (79),

$$\begin{aligned} -\log \left(\Pr_{\{p_n\}} \left(i : \frac{1}{n} \log p_n(i) \geq \mu \right) \right) &\geq \sup_{0 \leq s \leq 1} \left(sn\mu - \log \sum_i p_n(i)^{1+s} \right), \\ -\log \left(\Pr_{\{p_n\}} \left(i : -\frac{1}{n} \log q_n(i) \geq \lambda - \mu \right) \right) &\geq \sup_{0 \leq s \leq 1} \left(sn(\lambda - \mu) - \log \sum_i p_n(i) q_n(i)^{-s} \right). \end{aligned} \tag{81}$$

From the joint convexity of $\mathrm{tr}(A^s B^{1-s})$ for $-1 < s < 0$ [50, 51] we find that the function $g_s(\rho, \sigma) := \mathrm{tr}(\rho^{1+s}\sigma^{-s})$ is monotonic decreasing under trace preserving CP maps for every $0 < s < 1$. Defining the quantum operation $\mathcal{E}(X) = \sum_i E_i X E_i$,

$$\begin{aligned} \sum_i p_n(i)^{1+s} &= \dim(\mathcal{H})^{-ns} g_s \left(\mathcal{E}(\rho^{\otimes n}), \mathcal{E}\left(\frac{\mathbb{I}^{\otimes n}}{\dim(\mathcal{H})^n} \right) \right) \\ &\leq \dim(\mathcal{H})^{-ns} g_s \left(\rho^{\otimes n}, \frac{\mathbb{I}^{\otimes n}}{\dim(\mathcal{H})^n} \right) = \mathrm{tr}((\rho^{\otimes n})^{1+s}). \end{aligned} \tag{82}$$

Applying it to the first inequality in Eq. (81) gives the first term on the right hand side in Eq. (78).

For the second bound, we first note that the permutation-invariance of $R_n := (\rho^{\otimes n} - 2^{\lambda n}\sigma^{\otimes n})$ allows us to write it in the Schur basis as

$$R_n = \bigoplus_\lambda R_\lambda \otimes \mathbb{I}_\lambda, \tag{83}$$

where, as in the proof of Lemma III.4, λ labels the irreps of S_n, $\mathbb{I}_\lambda$ is the identity on the irrep labelled by λ, and R_λ is a Hermitian operator acting on the multiplicity space of the the irrep labelled by λ [46]. It is then clear that

$$Q_n = \bigoplus_\lambda Q_\lambda \otimes \mathbb{I}_\lambda, \tag{84}$$

where the Q_λ are projectors onto $(R_\lambda)_+$. Likewise,

$$\sigma^{\otimes n} = \bigoplus_\lambda \sigma_\lambda \otimes \mathbb{I}_\lambda, \;\; \rho^{\otimes n} = \bigoplus_\lambda \rho_\lambda \otimes \mathbb{I}_\lambda \tag{85}$$

for positive semidefinite operators $\sigma_\lambda, \rho_\lambda$.

As $\text{supp}(R_n) \subseteq \text{supp}(\sigma^{\otimes n})$, we have that for each λ, $\text{supp}(R_\lambda) \subseteq \text{supp}(\sigma_\lambda)$. We consider an eigen-decomposition of $R_\lambda := \sum_k e_{k,\lambda} E_{k,\lambda}$ with eigenprojectors $E_{k,\lambda}$ divided into three disjoint subsets, with members of the first one being subprojections of $\text{supp}(R_\lambda)$, members of the second one being subprojections of the orthogonal complement of $\text{supp}(R_\lambda)$ in $\text{supp}(\sigma_\lambda)$, and members of the third one being subprojections of $\text{supp}(\sigma_\lambda)^\perp$. Defining the quantum operation $\mathcal{E}_\lambda(X) := \sum_k E_{k,\lambda} X E_{k,\lambda}$, this particular choice of eigen-projectors $E_{k,\lambda}$ ensures that $\text{supp}(\mathcal{E}_\lambda(\sigma_\lambda)) \subseteq \text{supp}(\sigma_\lambda)$, a property which will be used next.

We identify the original eigen-projectors $\{E_k\}$ of Q with $\{\bigoplus_\lambda E_{\lambda,k_\lambda} \otimes \mathbb{I}_\lambda\}$, for all possible combinations of the labels k, λ. Then $\mathcal{E}(X) = \bigoplus_\lambda \mathcal{E}_\lambda \otimes \mathbb{I}_\lambda(X)$ and we can write

$$\begin{aligned} \sum_i p_n(i) q_n(i)^{-s} &= \text{tr}(\mathcal{E}(\rho^{\otimes n})(\mathcal{E}(\sigma^{\otimes n}))^{-s}) \\ &= \text{tr}(\rho^{\otimes n}(\mathcal{E}(\sigma^{\otimes n}))^{-s}) \\ &= \sum_\lambda \text{tr}(\rho_\lambda \mathcal{E}_\lambda(\sigma_\lambda)^{-s}) \dim(\mathbb{I}_\lambda). \end{aligned} \tag{86}$$

From Lemma 9 of Ref. [7] we find for each λ, $\sigma_\lambda \leq \dim \mathcal{H}_\lambda \mathcal{E}_\lambda(\sigma_\lambda)$, where $\mathcal{H}_\lambda$ is the Hilbert space in which σ_λ acts on. As $\text{supp}(\mathcal{E}_\lambda(\sigma_\lambda)) = \text{supp}(\sigma_\lambda)$, we can apply the operator monotonicity of $-u^{-1}$ for $0 < t \leq 1$ to get

$$(\mathcal{E}_\lambda(\sigma_\lambda))^{-s} \leq (\dim \mathcal{H}_\lambda)^s (\sigma_\lambda)^{-s}. \tag{87}$$

Applying the equation above to Eq. (86) and using the bound $\dim(\mathcal{H}_\lambda) \leq (n+1)^{\dim(\mathcal{H})}$ on the dimension of the multiplicity spaces $\mathcal{H}_\lambda$ [46],

$$\begin{aligned} \sum_i p_n(i) q_n(i)^{-s} &\leq (n+1)^{s \dim(\mathcal{H})} \sum_\lambda \text{tr}(\rho_\lambda (\sigma_\lambda)^{-s}) \dim(\mathbb{I}_\lambda) \\ &= (n+1)^{s \dim(\mathcal{H})} \text{tr}(\rho^{\otimes n} (\sigma^{\otimes n})^{-s}), \end{aligned} \tag{88}$$

and we are done. □

We are now in position to prove the direct part of Proposition III.1.

Proof (Direct part Proposition III.1)

We show that

$$\liminf_{n\to\infty} \min_{\omega_n \in \mathcal{M}_n} \text{tr}(\rho^{\otimes n} - 2^{yn} \omega_n)_+ = 1 - \mu, \tag{89}$$

with $\mu > 0$, implies $y \geq E_{\mathcal{M}}^\infty(\rho)$. First, if $\mu = 1$, we find from Corollary III.3 that $y > E_{\mathcal{M}}^\infty(\rho)$. So in the rest of the proof we show that if $0 < \mu < 1$, then $y \geq E_{\mathcal{M}}^\infty(\rho)$.

Let $\{\sigma_n \in \mathcal{M}_n\}_{n\in\mathbb{N}}$ be a sequence of optimal solutions in the minimization of Eq. (54). Note that from Lemma C.2 and property 5 of the sets $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$, we can take the states σ_n to be permutation-symmetric.

For each $n \in \mathbb{N}$ we have $\rho^{\otimes n} \leq 2^{yn}\sigma_n + (\rho^{\otimes n} - 2^{yn}\sigma_n)_+$. Applying Lemma C.5 once more we see that there is an increasing sequence $\mathcal{F}$ of the integers going to infinity and states ρ_n, with $n \in \mathcal{F}$, such that $F(\rho_n, \rho^{\otimes n}) \geq \mu/2 := \lambda$ and

$$\rho_n \leq \frac{2^{yn}}{\lambda}\sigma_n, \tag{90}$$

From Lemma C.2 and the permutation-invariance of σ_n and $\rho^{\otimes n}$, we can also take ρ_n to be permutation-symmetric. Let $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}_E$ be a purification of ρ, where $\mathcal{H}_E \cong \mathcal{H}$ is the purifying Hilbert space. Then, by Lemma III.4 there is a permutation-symmetric purification $|\Psi_n\rangle \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}_E^{\otimes n}$ of ρ_n such that $|\langle\theta^{\otimes n}|\Psi_n\rangle| \geq \lambda$. By Lemma III.5 and Eq. (90), in turn, we find that there is a $|\Psi_{n,m}\rangle$ approximating $|\Psi_{n,m,r}\rangle \in |\theta\rangle^{[\otimes, n-m, r]}$ such that

$$|||\Psi_{n,m}\rangle\langle\Psi_{n,m}| - |\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|||_1 \leq 2\sqrt{2}\lambda^{-1}e^{-\frac{mr}{2n}} \tag{91}$$

and

$$\mathrm{tr}_E(|\Psi_{n,m}\rangle\langle\Psi_{n,m}|) \leq \lambda^{-2}\mathrm{tr}_{1,\ldots,m}(\rho_n) \leq \lambda^{-3}2^{yn}\mathrm{tr}_{1,\ldots,m}(\sigma_n), \tag{92}$$

where the partial trace is taken over the purifying Hilbert space $\mathcal{H}_E^{\otimes n-m}$.

From the operator monotonicity of the $\log$ and property 3 of the sets,

$$\frac{1}{n}E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\Psi_{n,m}\rangle\langle\Psi_{n,m}|)) \leq y - 3\frac{\log(\lambda)}{n} \tag{93}$$

From Lemma C.3

$$\frac{1}{n}E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|)) \leq y - 3\frac{\log(\lambda)}{n} + f(2\sqrt{2}\lambda^{-1}e^{-\frac{mr}{2n}}) \tag{94}$$

for every $r \leq n - m$, where $f : \mathbb{R} \to \mathbb{R}$ is such that $\lim_{x\to 0} f(x) = 0$.

Then, setting $m = r = n^{2/3}$, taking the limit $n \to \infty$ in Eq. (94), and using Lemma III.7, we find that for every ρ with $\lambda_{\max}(\rho) < 1$,

$$E_{\mathcal{M}}^{\infty}(\rho) \leq \liminf_{n\to\infty} \frac{1}{n}E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|)) \leq y. \tag{95}$$

Finally, we show that the result for non-pure states implies its validity to pure states too, completing the proof. Let $|\psi\rangle$ be a pure state and $y < E_{\mathcal{M}}^{\infty}(|\psi\rangle\langle\psi|)$. Asymptotic continuity of $E_{\mathcal{M}}^{\infty}$ (see Lemma C.3) yields the existence of a $\chi > 0$ such that $y < E_{\mathcal{M}}^{\infty}(\zeta)$ for $\zeta := (|\psi\rangle\langle\psi| + \chi\sigma)/(1+\chi)$, where σ is the full rank state from property 2 of the sets $\mathcal{M}_n$. Then, assuming the result for mixed states, we have

$$\lim_{n\to\infty} \min_{\omega_n \in \mathcal{M}_n} \mathrm{tr}(\zeta^{\otimes n} - 2^{yn}\omega_n)_+ = 1. \tag{96}$$

By the asymptotic equipartition theorem [1] we can find a sequence of states $\zeta_n = \sum_i p_{i,n}\zeta_{i,n}$ where $\{p_{i,n}\}$ is a probability distribution and each $\zeta_{i,n}$ is - up to permutations of the copies - of the form $(|\psi\rangle\langle\psi|)^{\otimes n-m_{i,n}} \otimes \sigma^{\otimes m_{i,n}}$, with

$$\lim_{n\to\infty} \max_i \frac{m_{i,n}}{n} = \lim_{n\to\infty} \min_i \frac{m_{i,n}}{n} = \chi/(1+\chi) \tag{97}$$

and $\lim_{n\to\infty} ||\zeta^{\otimes n} - \zeta_n||_1 = 0$. In particular the inequality $\mathrm{tr}(\zeta^{\otimes n} - 2^{yn}\omega_n)_+ \leq \mathrm{tr}(\zeta_n - 2^{yn}\omega_n)_+ + ||\zeta^{\otimes n} - \zeta_n||_1$ yields

$$\lim_{n\to\infty} \min_{\omega\in\mathcal{M}_n} \mathrm{tr}(\zeta_n - 2^{yn})_+ = 1. \tag{98}$$

Note also that $(X, Y) \mapsto \mathrm{tr}(X - Y)_+$ is convex and hence $\rho \mapsto \min_{\omega_n\in\mathcal{M}_n} \mathrm{tr}(\rho - 2^{yn}\omega_n)_+$ is convex too. Therefore

$$\min_{\omega_n\in\mathcal{M}_n} \mathrm{tr}(\zeta_n - 2^{yn}\omega_n)_+ \leq \sum_i p_{i,n} \min_{\omega_n\in\mathcal{M}_n} \mathrm{tr}(\zeta_{i,n} - 2^{yn}\omega_n)_+ \leq \max_i \min_{\omega_n\in\mathcal{M}_n} \mathrm{tr}(\zeta_{i,n} - 2^{yn}\omega_n)_+. \tag{99}$$

Let i^* be a maximizer of the last formula above. Then, $\zeta_{i^*,n}$ can be written as $P_{f_{i^*}}(|\psi\rangle\langle\psi|^{\otimes n-m_n} \otimes \sigma^{\otimes m})P^*_{f_{i^*}}$, for some $m = m(n) \in \mathbb{N}$ and $f_{i^*} \in S_n$. Hence

$$\begin{aligned}\max_i \min_{\omega_n\in\mathcal{M}_n} \mathrm{tr}(\zeta_{i,n} - 2^{yn}\omega_n)_+ &\leq \min_{\omega_n\in\mathcal{M}_{n-m}} \mathrm{tr}(P_{f_{i*}}(|\psi\rangle\langle\psi|^{\otimes n-m} \otimes \sigma^{\otimes m})P^*_{f_{i*}} - P_{f_{i*}}(\omega_n \otimes \sigma^{\otimes m})P^*_{f_{i*}}) \\ &= \min_{\omega_n\in\mathcal{M}_{n-m}} \mathrm{tr}(|\psi\rangle\langle\psi|^{\otimes n-m} - 2^{yn}\omega_n)_+. \end{aligned} \tag{100}$$

By the above,

$$1 \leq \liminf_{n\to\infty} \min_{\omega\in\mathcal{M}_{n-m}} \mathrm{tr}(|\psi\rangle\langle\psi|^{\otimes n-m} - 2^{yn}\omega_n)_+ \leq \liminf_{n\to\infty} \min_{\omega\in\mathcal{M}_n} \mathrm{tr}(|\psi\rangle\langle\psi|^{\otimes n} - 2^{yn}\omega_n)_+, \tag{101}$$

where in the last inequality we used that $\lim_{n\to\infty} n - m = +\infty$, due to the assumption $\lim_{n\to\infty} \frac{1}{n} \max_i m_{i,n} = \frac{\chi}{1+\chi}$. □

The next lemma shows a property of the measures $E_{\mathcal{M}_k}$ analogous to the non-lockability of the relative entropy of entanglement [44], in this case manifested in the almost power states.

Lemma III.7 *Let $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}_E$ and $\rho = tr_E(|\theta\rangle\langle\theta|)$ with $\lambda_{\max}(\rho) < 1$. Let $\{|\Psi_{n,m,r}\rangle \in |\theta\rangle^{[\otimes,n-m,r]}\}_{n,m,r}$ be a sequence of almost power states along $|\theta\rangle$, with $r = o(n)$ and $m = o(n)$. Then*

$$E^\infty_{\mathcal{M}}(\rho) \leq \liminf_{n\to\infty} \frac{1}{n} E_{\mathcal{M}_{n-m}}(tr_E(|\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|)). \tag{102}$$

Proof Write $|\Psi_{n,m,r}\rangle = \sum_{k=0}^r \beta_k \mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-m-k})$, where $|\eta_k\rangle$ are permutation-symmetric states living in $(\mathcal{H}\perp|\theta\rangle)^{\otimes k}$ and $\sum_k |\beta_k|^2 = 1$. Define

$$|\Phi_{n,m,r}\rangle := \sum_{k:|\beta_k|\geq 1/n} \beta_k \mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-m-k}) \tag{103}$$

and $|\tilde{\Phi}_{n,m,r}\rangle := |\Phi_{n,m,r}\rangle/|||\Phi_{n,m,r}\rangle||$. Note that $\lim_{n\to\infty} |||\tilde{\Phi}_{n,m,r}\rangle - |\Psi_{n,m,r}\rangle|| = 0$. Thus, from the asymptotic continuity of the measures $E_{\mathcal{M}_k}$ (Lemma C.3) it follows

$$\liminf_{n\to\infty} \frac{1}{n} E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\Psi_{n,m,r}\rangle\langle\Psi_{n,m,r}|)) = \liminf_{n\to\infty} \frac{1}{n} E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\tilde{\Phi}_{n,m,r}\rangle\langle\tilde{\Phi}_{n,m,r}|)), \tag{104}$$

and thus it suffices to show that the R.H.S. of the equation above is larger or equal to $E^\infty_{\mathcal{M}}(\rho)$.

From Lemma III.8 we find

$$\begin{aligned}(|\theta\rangle\langle\theta|)^{\otimes n-m-r} &\leq 2^{nh(\frac{r}{n-m})} n^2 \mathrm{tr}_{1,\ldots,r}(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|) \\ &\leq 2^{nh(\frac{r}{n-m})} n^2 \mathrm{tr}_{1,\ldots,r}(|\tilde{\Phi}_{n,m,r}\rangle\langle\tilde{\Phi}_{n,m,r}|), \end{aligned} \tag{105}$$

where the last inequality follows from $|||\Phi_{n,m,r}\rangle|| \leq 1$.

For simplicity of notation we define $\pi_n := \mathrm{tr}_{1,\ldots,r}\mathrm{tr}_E(|\tilde{\Phi}_{n,m,r}\rangle\langle\tilde{\Phi}_{n,m,r}|)$. Tracing out the environment Hilbert space in Eq. (105),

$$\rho^{\otimes n-m-r} \leq 2^{nh(\frac{r}{n-m})} n^2 \pi_n. \tag{106}$$

Let $\tilde{\omega}_n \in \mathcal{M}_{n-m-r}$ be such that

$$E_{\mathcal{M}_{n-m-r}}(\pi_n) = S(\pi_n||\tilde{\omega}_n). \tag{107}$$

and set

$$\omega_n := \frac{1}{1+\tau}\tilde{\omega}_n + \frac{\tau}{1+\tau}\sigma^{\otimes n-m-r}, \tag{108}$$

where and $\tau > 0$. We introduce ω_n in order to have a non-negligible lower bound on the minimum eigenvalue of a close-to-optimal state for π_n, which will show useful later on.

From the previous equation and the operator monotonicity of the $\log$ function,

$$E_{\mathcal{M}_{n-m-r}}(\pi_n) = S(\pi_n||\tilde{\omega}_n) \geq S(\pi_n||\omega_n) - \log(1+\tau). \tag{109}$$

Let $\lambda_{n,\nu} = E_{\mathcal{M}_{n-m-r}}(\pi_n) + n\nu + \log(1+\tau) \geq S(\pi_n||\omega_n) + n\nu$, for $\nu > 0$. For every integer l

$$\begin{aligned}
\rho^{\otimes(n-m-r)l} &\leq n^{2l}2^{nh(\frac{r}{n-m})l}\pi_n^{\otimes l} \\
&\leq n^{2l}2^{nh(\frac{r}{n-m})l}2^{\lambda_{n,\nu}l}\omega_n^{\otimes l} + n^{2l}2^{nh(\frac{r}{n-m})l}(\pi_n^{\otimes l} - 2^{\lambda_{n,\nu}l}\omega_n^{\otimes l})_+.
\end{aligned} \tag{110}$$

From Lemma III.9 we find that for every $\nu > 0$ there is a constant $\gamma > 0$ with the property that for every $n \in \mathbb{N}$, there is an integer l_n such that

$$\mathrm{tr}(\pi_n^{\otimes l} - 2^{\lambda_{n,\nu}l}\omega_n^{\otimes l})_+ \leq 2^{-\gamma nl}. \tag{111}$$

for every $l \geq l_n$.

Then applying Lemma C.5 to Eq. (110), we find that for every n sufficiently large, there is a sequence of states $\rho_{l,n}$ such that $\lim_{l\to\infty} ||\rho_{l,n} - \rho^{\otimes(n-m-r)l}||_1 = 0$ and

$$\rho_{l,n} \leq g(l)(n^2 2^{nh(\frac{r}{n-m})})^l 2^{\lambda_{n,\nu}l}\omega_n^{\otimes l}, \tag{112}$$

for a function $g(l)$ such that $\lim_{l\to\infty} g(l) = 1$. Then we have

$$\begin{aligned}
(n-m-r)E_{\mathcal{M}}^{\infty}(\rho) &= E_{\mathcal{M}}^{\infty}(\rho^{\otimes n-m-r}) = \lim_{l\to\infty}\frac{1}{l}E_{\mathcal{M}_{(n-m-r)l}}(\rho^{\otimes(n-m-r)l}) \\
&= \lim_{l\to\infty}\frac{1}{l}E_{\mathcal{M}_{(n-m-r)l}}(\rho_{l,n}) \leq \lim_{l\to\infty}\frac{1}{l}S_{\max}(\rho_{l,n}||\omega_n^{\otimes l}) \\
&\leq \lim_{l\to\infty}\frac{1}{l}\log g(l) + 2\log(n) + nh\left(\frac{r}{n-m}\right) + \lambda_{n,\nu} \\
&= 2\log(n) + nh\left(\frac{r}{n-m}\right) + E_{\mathcal{M}_{n-m-r}}(\pi_n) + \nu n + \log(1+\tau)
\end{aligned} \tag{113}$$

and, since, $E_{\mathcal{M}_{n-m-r}}(\pi_n) \leq E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\tilde{\Phi}_{n,m,r}\rangle\langle\tilde{\Phi}_{n,m,r}|))$,

$$\begin{aligned}
E_{\mathcal{M}}^{\infty}(\rho) &= \liminf_{n\to\infty}\frac{1}{n-m-r}\left(2\log(n) + nh\left(\frac{r}{n-m}\right) + E_{\mathcal{M}_{n-m-r}}(\pi_n) + \nu n + \log(1+\tau)\right) \\
&\leq \liminf_{n\to\infty} E_{\mathcal{M}_{n-m}}(\mathrm{tr}_E(|\tilde{\Phi}_{n,m,r}\rangle\langle\tilde{\Phi}_{n,m,r}|)) + 2\nu.
\end{aligned}$$

Taking ν to zero and using Eq. (104) we find Eq. (102). □

As in the proof above, let $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}_E$ and $\rho := \mathrm{tr}_E(|\theta\rangle\langle\theta|)$ be such that $\lambda_{\max}(\rho) < 1$. The next three lemmata concern the following states:

$$|\Phi_{n,m,r}\rangle := \sum_{k:|\beta_k| \geq 1/n} \beta_k \mathrm{Sym}(|\eta_k\rangle \otimes |\theta\rangle^{\otimes n-m-k}), \tag{114}$$

for complex-valued coefficients β_k and states $|\eta_k\rangle$ living in $(\mathcal{H} \perp |\theta\rangle)^{\otimes k}$, and

$$\pi_n := \mathrm{tr}_{1,\ldots,r}\mathrm{tr}_E(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|)/\langle\Phi_{n,m,r}|\Phi_{n,m,r}\rangle. \tag{115}$$

Lemma III.8 *Let $k_{\max} \leq (n-m)/2$ be the maximum k appearing in Eq. (114). Then, for $r \geq k_{\max}$,*

$$(|\theta\rangle\langle\theta|)^{\otimes n-m-r} \leq 2^{nh(\frac{r}{n-m})} n^2 tr_{1,\ldots,r}(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|), \tag{116}$$

Proof Let $|\phi\rangle := |\eta_{k_{\max}}\rangle \otimes |\theta\rangle^{\otimes n-m-k_{\max}}$. Then

$$|\Phi_{n,m,r}\rangle = c|\phi\rangle + c' e^{i\vartheta}|\phi^\perp\rangle, \tag{117}$$

where

$$c := \binom{n-m}{k_{\max}}^{-1/2} \beta_{k_{\max}}, \tag{118}$$

$\vartheta \in \mathbb{R}$, $c' \geq 0$, and $|\phi^\perp\rangle$ is a state orthogonal to $|\phi\rangle$. From Eq. (114), we can write $|\phi^\perp\rangle$ as a superposition of states of the form $|f_1\rangle \otimes \ldots \otimes |f_{n-m}\rangle$, where at least in one of the first $k_{\max}$ registers, $|f_i\rangle = |\theta\rangle$. Therefore, as $|\eta_{k_{\max}}\rangle$ lives in $(\mathcal{H} \perp |\theta\rangle)^{\otimes k_{\max}}$, we get $\mathrm{tr}_{1,\ldots,k_{\max}}(|\phi\rangle\langle\phi^\perp|) = 0$ and thus

$$\begin{aligned} \mathrm{tr}_{1,\ldots k_{\max}}(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|) &= |c|^2 \mathrm{tr}_{1,\ldots k_{\max}}(|\phi\rangle\langle\phi|) + (c')^2 \mathrm{tr}_{1,\ldots k_{\max}}(|\phi^\perp\rangle\langle\phi^\perp|) \\ &\geq |c|^2 \mathrm{tr}_{1,\ldots k_{\max}}(|\phi\rangle\langle\phi|) \\ &= |c|^2 (|\theta\rangle\langle\theta|)^{\otimes n-m-k_{\max}}. \end{aligned} \tag{119}$$

From Eq. (118),

$$(|\theta\rangle\langle\theta|)^{\otimes n-m-k_{\max}} \leq \binom{n-m}{k_{\max}} |\beta_{k_{\max}}|^{-2} \mathrm{tr}_{1,\ldots k_{\max}}(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|). \tag{120}$$

Note that $|\beta_{k_{\max}}|^{-2} \leq n^2$ and the entropic bound $\binom{n}{k} \leq 2^{nh(k/n)}$ (see e.g. Lemma 17.5.1 of [1]). Moreover, from the monotonicity of the binary entropy in the interval $[0, 1/2]$, $h(k_{\max}/(n-m)) \leq h(r/(n-m))$. Therefore,

$$(|\theta\rangle\langle\theta|)^{\otimes n-m-k_{\max}} \leq 2^{nh(\frac{r}{n-m})} n^2 \mathrm{tr}_{1,\ldots k_{\max}}(|\Phi_{n,m,r}\rangle\langle\Phi_{n,m,r}|). \tag{121}$$

The lemma follows by tracing out the first $r - k_{\max}$ registers in the equation above. □

As in the proof of the direct part of Proposition III.1, let $\tilde{\omega}_n$ be such that $E_{\mathcal{M}_{n-m-r}}(\pi_n) = S(\pi_n||\tilde{\omega}_n)$ and define

$$\omega_n := \frac{1}{1+\tau}\tilde{\omega}_n + \frac{\tau}{1+\tau}\sigma^{\otimes n-m-r}, \tag{122}$$

with $\tau > 0$.

Lemma III.9 *Let ω_n be given by Eq. 122, π_n by Eq. (115), and λ be such that*

$$\lambda = \lambda_{n,\nu} \geq S(\pi_n||\omega_n) + \nu n, \tag{123}$$

for $\nu > 0$. Then, there is a $\gamma > 0$ and a sequence $\{l_n\}_{n\in\mathbb{N}}$ such that for sufficiently large n and $l \geq l_n$,

$$tr(\pi_n^{\otimes l} - 2^{\lambda_{n,\nu} l}\omega_n^{\otimes l})_+ \leq 2^{-\gamma n l}, \tag{124}$$

Proof From Lemma III.6,

$$\text{tr}(\pi_n^{\otimes l} - 2^{\lambda l}\omega_n^{\otimes l})_+ \leq 2^{-lp(s)} + 2^{-lq(s)}, \tag{125}$$

with $p_n(s) := (s\mu - \log \text{tr}(\pi_n^{1+s}))$ and $q_n(s) := (s(\lambda-\mu) - sD^{n-m-r}\frac{\log(1+l)}{l} - \log \text{tr}(\pi_n\omega_n^{-s}))$. We set $\mu = (\nu/2 - S(\rho))n$ and show that each of the two bounds in the equation above is smaller than $2^{-\gamma n l}$, for a given constant γ and sufficiently large n and $l \geq l_n$.

From Eq. (103) we can write $\pi_n = \text{tr}_{1,\ldots,r}\text{tr}_E(|\Psi_{\pi_n}\rangle\langle\Psi_{\pi_n}|)$ (identifying $|\Psi_{\pi_n}\rangle$ and $|\Phi_{n,m,r}\rangle/|||\Phi_{n,m,r}\rangle||$), with

$$|\Psi_{\pi_n}\rangle := \sum_{k=0}^{r} \alpha_k \text{Sym}(|\chi_k\rangle \otimes |\theta\rangle^{\otimes n-m-k}), \tag{126}$$

where $\sum_{k=0}^{r} |\alpha_k|^2 = 1$ and

$$|\chi_k\rangle \in (\mathcal{H} \perp |\theta\rangle)^{\otimes k}. \tag{127}$$

Each $\text{Sym}(|\chi_k\rangle \otimes |\theta\rangle^{\otimes n-m-k})$ is a superposition of $\binom{n-m}{k}$ terms which, up to permutation of the copies and normalization, have the form $|\chi_k\rangle \otimes |\theta\rangle^{\otimes n-m-k}$; let us denote these by $|\psi_{k,j}\rangle$. from Eq. (127), we get $|\langle\psi_{k,j}|\psi_{k',j'}\rangle| = \delta_{kk'}\delta_{jj'}$. Therefore we can write

$$|\Psi_{\pi_n}\rangle = \sum_{k=0}^{r} \sum_{j=1}^{\binom{n-m}{k}} \varsigma_{k,j}|\psi_{k,j}\rangle, \tag{128}$$

with $\sum_{k,j} |\varsigma_{k,j}|^2 = 1$. By Lemma C.6,

$$|\Psi_{\pi_n}\rangle\langle\Psi_{\pi_n}| \leq (r+1)\binom{n-m}{r} \sum_{k,j} |\varsigma_{k,j}|^2 |\psi_{k,j}\rangle\langle\psi_{k,j}|, \tag{129}$$

where we used that since $k, m, r = o(n)$, $\binom{n-m}{k} \leq \binom{n-m}{r}$ for every $k \leq r$. Tracing out E and the first r copies in both sides of the equation above, we find

$$\pi_n \leq (r+1)\binom{n-m}{r} \sum_j p_j\rho_j \leq (r+1)2^{(n-m)h(\frac{r}{n-m})} \sum_j p_j\rho_j, \tag{130}$$

where $\{p_j\}$ is a probability distribution and each ρ_j is of the form $\rho^{\otimes n-m-r} \otimes \sigma_r$, up to permutations of the copies, with an arbitrary state σ_r acting on $\mathcal{H}^{\otimes r}$.

Then, by the Schur-convexity of the function $h(x) = x^{1+s}$ ($s \geq 0$),

$$\begin{aligned} \text{tr}(\pi_n^{1+s}) &\leq (r+1)^{1+s}2^{(n-m)h(\frac{r}{n-m})(1+s)}\text{tr}((\sum_j p_j\rho_j)^{1+s}) \\ &\leq (r+1)^{1+s}2^{(n-m)h(\frac{r}{n-m})(1+s)} \sum_j p_j \text{tr}(\rho_j^{1+s}), \end{aligned} \tag{131}$$

from which follows that, with $h_{n,m,r,s} := -(1+s)(\log(r+1) + (n-m)h\left(\frac{r}{n-m}\right))$,

$$\begin{aligned} -\log \mathrm{tr}(\pi_n^{1+s}) &\geq h_{n,m,r,s} - \max_j \log \mathrm{tr}(\rho_j^{1+s}) \\ &= h_{n,m,r,s} - \max_j \log \mathrm{tr}((\sigma_j)^{1+s}) - (n-m-r)\log \mathrm{tr}(\rho^{1+s}) \\ &\geq h_{n,m,r,s} + (m+r)\log \mathrm{tr}(\rho^{1+s}) - n \log \mathrm{tr}(\rho^{1+s}), \end{aligned} \tag{132}$$

where the last inequality follows from $\mathrm{tr}((\sigma_j)^{1+s}) \leq 1$. Note that the first two terms in the equation above are $o(n)$. Therefore

$$-\log \mathrm{tr}(\pi_n^{1+s}) \geq -n \log \mathrm{tr}(\rho^{1+s}) - o(n). \tag{133}$$

Letting $g(s) := -\log \mathrm{tr}(\rho^{1+s})$, we see that $g(0) = 0$ and $g'(0) = S(\rho)$. Then,

$$\begin{aligned} p_n(s) &= s(\nu/2 - S(\rho))n - \log \mathrm{tr}(\pi_n^{1+s}) \\ &\geq s(\nu/2 - S(\rho))n - n \log \mathrm{tr}(\rho^{1+s}) - o(n) \\ &\geq ns\nu/2 - n \max_{0\leq t\leq s} |g''(t)| s^2 - o(n). \end{aligned} \tag{134}$$

Thus there is a s small enough, independent of n, such that for sufficiently large n, $p_n(s) \geq ns\nu/4$.

Considering the second bound in Eq. (125), let $f_n(s) := -\frac{1}{n}\log \mathrm{tr}\left(\pi_n \omega_n^{-s}\right)$. As ω_n is full rank, we find from Taylor's Theorem,

$$-\frac{1}{n}\log \mathrm{tr}\left(\pi_n \omega_n^{-s}\right) = f_n(0) + f_n'(0)s + f_n''(t_{s,n})s^2/2, \tag{135}$$

for some real number $t_{s,n} \leq s$. A simple calculation shows that $f_n(0) = 0$,

$$f_n'(0) = \frac{1}{n}\mathrm{tr}(\pi_n \log \omega_n), \tag{136}$$

and

$$f_n''(s) = -\frac{1}{n}\left(\frac{\mathrm{tr}(\pi_n \omega_n^{-s}(\log \omega_n)^2)}{\mathrm{tr}(\pi_n \omega_n^{-s})} - \left(\frac{\mathrm{tr}(\pi_n \omega_n^{-s}\log \omega_n)}{\mathrm{tr}(\pi_n \omega_n^{-s})}\right)^2\right). \tag{137}$$

We next show that there is a s sufficiently small, but independent of n, such that

$$\max_{0\leq t\leq s} |f_n''(t)| \leq 1 \tag{138}$$

for n sufficiently large. Hence

$$\begin{aligned} q_n(s) &\geq s(n\nu/2 + S(\pi_n||\omega_n) + nS(\rho) + \mathrm{tr}(\pi_n \log \omega_n)) - sD^{n-m-r}\frac{\log(1+l)}{l} - n\max_{0\leq t\leq s}|f_n''(t)|s^2 \\ &\geq \frac{s\nu n}{2} + s(nS(\rho) - S(\pi_n)) - sD^{n-m-r}\frac{\log(1+l)}{l} - ns^2. \end{aligned} \tag{139}$$

Using Lemma III.10, choosing s sufficiently small and l_n such that $D^{n-m-r}\frac{\log(1+l_n)}{l_n} = o(n)$, we find $q_n(s) \geq ns\nu/4$, for sufficiently large n and $l \geq l_n$.

In order to prove Eq. (138), we consider the basis where π_n is diagonal

$$\pi_n = \mathrm{Diag}(\lambda_{1,n}, \lambda_{2,n}, ...). \tag{140}$$

and write ω_n in this basis

$$\omega_n = U\text{Diag}(\mu_{1,n}, \mu_{2,n}, ...)U^\dagger, \tag{141}$$

where U is a unitary. Note that Eq. (122) gives

$$\omega_n = \frac{1}{1+\tau}\tilde{\omega}_n + \frac{\tau}{1+\tau}\sigma^{\otimes n-m-r} \geq \frac{\tau}{1+\tau}\sigma^{\otimes n-m-r} \geq \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}. \tag{142}$$

where $\lambda_{\min}(\sigma) > 0$ is the minimum eigenvalue of σ.

From Eq. (137) it follows that we can write

$$|f_n''(s)| = \frac{1}{n}\left(\sum_j t_{j,n}(\log \mu_{j,n})^2 - \left(\sum_j t_{j,n}\log \mu_{j,n}\right)^2\right), \tag{143}$$

where $\{t_{j,n}\}$ is the probability distribution given by

$$t_{j,n} := \frac{\mu_{j,n}^{-s}\sum_i \lambda_{i,n}|U_{i,j}|^2}{\sum_{i,j}\lambda_{i,n}\mu_{j,n}^{-s}|U_{i,j}|^2}. \tag{144}$$

Clearly we can upper bound the function $|f_n''(s)|$ by maximizing over the $\mu_{j,n}$ while keeping the probabilities $t_{j,n}$ fixed. We extend the set of allowed $\mu_{j,n}$ even more and consider all probability distributions for which $\mu_{j,n} \geq \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$. We are hence interested in maximizing the function

$$g(\mu_{1,n}, \mu_{2,n}, ...) = \frac{1}{n}\left(\sum_j t_{j,n}(\log \mu_{j,n})^2 - \left(\sum_j t_{j,n}\log \mu_{j,n}\right)^2\right) \tag{145}$$

over the set of probability distributions $\{\mu_{j,n}\}$ such that

$$\mu_{j,n} \geq \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}, \tag{146}$$

for all j.

The function g will reach its maximum either on its extreme points or on the boundary of the set in which the maximization is performed. A simple calculation gives

$$\frac{\partial g}{\partial \mu_{k,n}} = \frac{1}{n}\left(2t_{k,n}\frac{\log \mu_{k,n}}{\mu_{k,n}} - 2\left(\sum_j t_{j,n}\log \mu_{j,n}\right)\frac{t_{k,n}}{\mu_{k,n}}\right) = 0 \Rightarrow \log \mu_{k,n} = \sum_i t_{i,n}\log \mu_{i,n}. \tag{147}$$

Hence, in the extreme points of g all the $\mu_{k,n}$ are equal and it is then easy to see that $g(\mu, \mu, ...) = 0$. As g is positive, it then follows that the maximum of g is attained on the boundary of the set in which the maximization is performed. Such boundary is composed of subsets of the original set given by Eq. (146) in which at least one of the $\mu_{j,n}$ is equal to $\frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$. Setting $\mu_{k,n} = \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$, the new function to be maximized is

$$\tilde{g}(\mu_{1,n}, ..., \mu_{k-1,n}, \mu_{k+1,n}, ...) = \frac{1}{n}\left(\sum_j t_{j,n}(\log \mu_{j,n})^2 - \left(\sum_j t_{j,n}\log \mu_{j,n}\right)^2\right), \tag{148}$$

where now $\mu_{k,n} = \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$ is a constant. Proceeding exactly as before, we find again that all the extreme points of $\tilde{g}$ are again minima of the function and, hence, the maximum of $\tilde{g}$ is attained once more on the boundary of the the set of probabilities allowed. This, in turn, is given by the union of subsets of the set given by Eq. (146) in which at least two of the $\mu_{k,n}$ are equal to $\frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$. We can continue with this process to show that all $\mu_{k,n}$ except one are equal to $\frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$. We hence find that the optimal choice of parameters is given by

$$\begin{cases} \tilde{\mu}_{j,n} = \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r} & \text{if } j \neq k, \\ \tilde{\mu}_{k,n} = 1 + \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r} - \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}, & \text{otherwise} \end{cases} \tag{149}$$

for some integer k. Let

$M := \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$ and $N := 1 + \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r} - \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$. It then follows that

$$\begin{aligned} g(\tilde{\mu}_{1,n}, \tilde{\mu}_{2,n}, ...) &= \frac{1}{n}\Big((1-t_{k,n})t_{k,n}\left(\log M\right)^2 + t_{k,n}\left(\log N\right)^2 \\ &- t_{k,n}^2\left(\log N\right)^2 - 2t_{k,n}(1-t_{k,n})\left(\log M \log N\right)\Big) \end{aligned} \tag{150}$$

We have

$$\left|\log M\right|, \left|\log N\right| \leq 2\log(\lambda_{\min}^{-1}(\sigma))n, \tag{151}$$

for sufficiently large n, and

$$\begin{aligned} t_{k,n} &= \frac{\mu_{k,n}^{-s}\sum_i \lambda_{i,n}|U_{i,k}|^2}{\sum_{i,j}\lambda_{i,n}\mu_{j,n}^{-s}|U_{i,j}|^2} \\ &\leq \frac{\lambda_{\max}(\pi_n)\sum_i |U_{i,k}|^2}{(\tau/((1+\tau)D^n))^s \sum_{i,j}\lambda_{i,n}|U_{i,j}|^2} \\ &= \lambda_{\max}(\pi_n)\left(\frac{(1+\tau)\lambda_{\min}(\sigma)^{-n+m+r}}{\tau}\right)^s, \end{aligned} \tag{152}$$

where the second inequality follows from $1 \geq \mu_{j,n} \geq \frac{\tau}{1+\tau}\lambda_{\min}(\sigma)^{n-m-r}$, which is a direct consequence of Eq. (142).

From Eq. (130), we have the bound

$$\lambda_{\max}(\pi_n) \leq 2^{o(n)}\lambda_{\max}(\sum_i p_i\rho_j) \leq 2^{o(n)}\lambda_{\max}(\rho)^{n-o(n)}. \tag{153}$$

Thus

$$t_{k,n} \leq 2^{o(n)}\left(\frac{(1+\tau)}{\tau}\right)^s (\lambda_{\min}(\sigma)^{-s}\lambda_{\max}(\rho))^n \lambda_{\max}(\rho)^{-o(n)}. \tag{154}$$

As by assumption $\lambda_{\max}(\rho) < 1$, choosing $s < \log(\lambda_{\max}(\rho))/\log(\lambda_{\min}(\sigma))$, we get that for n sufficiently large, $t_{k,n} \leq (10\log\lambda_{\min}^{-1}(\sigma)n)^{-1}$. Then, from Eqs. (150) and (151),

$$\begin{aligned} g(\tilde{\mu}_{1,n}, \tilde{\mu}_{2,n}, ...) &\leq 2\log\lambda_{\min}^{-1}(\sigma)n\left((1-t_{k,n})t_{k,n} + t_{k,n} + t_{k,n}^2 + 2(1-t_{k,n})t_{k,n}\right) \\ &\leq 10\log\lambda_{\min}^{-1}(\sigma)t_{k,n} \leq 1, \end{aligned} \tag{155}$$

and we are done. □

The final lemma of this section relates the entropy of an almost power state along ρ with its own entropy.

Lemma III.10 *Let π_n be given by Eq. (115) with $k, r = o(n)$. Then*

$$S(\pi_n) \leq nS(\rho) + o(n). \tag{156}$$

Proof Let $\rho = \sum_{i=1}^{d} p_i |i\rangle\langle i|$, with $d = \text{rank}(\rho)$, and

$$\rho^{\otimes n} := \sum_{i^n} p_{i^n} |i^n\rangle\langle i^n| \tag{157}$$

with $i^n := i_1 ... i_n$, $p_{i^n} := p_{i_1} ... p_{i_n}$, and $|i^n\rangle := |i_1\rangle ... |i_n\rangle$. For $\delta > 0$ define the set of typical sequences by $\mathcal{T}_\delta^n := \{i^n : |-\log p_{i^n} - nS(\rho)| \leq n\delta\}$, and the typical projector by

$$\Pi_\delta^n := \sum_{i^n \in \mathcal{T}_\delta^n} |i^n\rangle\langle i^n|. \tag{158}$$

Then from e.g. [52] (appendix C) we have

$$\text{tr}(\rho^{\otimes n}\Pi_\delta^n) \geq 1 - e^{-b\delta^2 n}, \tag{159}$$

and

$$\Pi_\delta^n \rho^{\otimes n} \Pi_\delta^n \geq 2^{-n(S(\rho)+\delta)} \Pi_\delta^n. \tag{160}$$

Let $\Pi_n' := (\mathbb{I}^{\otimes r} \otimes \Pi_{n^{-1/4}}^{n-m-r}) \otimes \mathbb{I}_E$, where the first identity is applied to the first r register of $\mathcal{H}^{\otimes n-m}$, while the second is applied to the purifying Hilbert space$\mathcal{H}_E^{\otimes n-m}$. Writing $|\Phi_{n,m,r}\rangle$ as in Eq. (117), we can define

$$|\Phi'_{n,m,r}\rangle = c\Pi_n'|\phi\rangle + \sqrt{1-c^2}e^{i\vartheta}|\phi^\perp\rangle \tag{161}$$

and follow the argument in the proof of Lemma III.8 (which applies unchanged to $|\Phi'_{n,m,r}\rangle$) to get

$$\Pi_{n^{-1/4}}^{n-m-r} \text{tr}_E \left((|\theta\rangle\langle\theta|)^{\otimes n-m-r} \right) \Pi_{n^{-1/4}}^{n-m-r} \leq 2^{nh\left(\frac{r}{n-m}\right)} n^2 \text{tr}_{1,...r} \text{tr}_E (|\Phi'_{n,m,r}\rangle\langle\Phi'_{n,m,r}|). \tag{162}$$

Hence from Eq. (160),

$$\lambda_{\min}\left(\text{tr}_{1,...r}\text{tr}_E\left(|\Phi'_{n,m,r}\rangle\langle\Phi'_{n,m,r}|\right)\right) \geq 2^{o(n)} \lambda_{\min}(\Pi_{n^{-1/4}}^n \rho^{\otimes n} \Pi_{n^{-1/4}}^n) \geq 2^{-n(S(\rho)+o(n))}. \tag{163}$$

Moreover, Eqs. (159) and (161) give

$$\begin{aligned} |\langle\Phi'_{n,m,r}|\Phi_{n,m,r}\rangle| &= c^2\langle\phi|\Pi_n'|\phi\rangle + (1-c^2) \\ &= c^2 \text{tr}(\rho^{\otimes n-m-r}\Pi_{n^{-1/4}}^{n-m-r}) + (1-c^2) \\ &\geq 1 - e^{-n^{1/8}}, \end{aligned} \tag{164}$$

for sufficiently large n. Defining,

$$\pi_n' := \text{tr}_{1,...,r}\text{tr}_E(|\Phi'_{n,m,r}\rangle\langle\Phi'_{n,m,r}|)/\langle\Phi'_{n,m,r}|\Phi'_{n,m,r}\rangle, \tag{165}$$

we get from Eq. (164) that

$$||\pi_n - \pi_n'||_1 = o(1). \tag{166}$$

Furthermore, from Eq. (163), $\lambda_{\min}(\pi_n') \geq 2^{-n(S(\rho)+o(n))}$, and thus

$$S(\pi_n') \leq -\log \lambda_{\min}(\pi_n') \leq nS(\rho) + o(n). \tag{167}$$

The lemma follows from Eqs. (166), (167) and Fannes inequality [53]. □

IV. PROOF OF COROLLARY II.2

In this section we prove that the regularized relative entropy of entanglement is faithful. The idea is to combine Theorem I with the exponential de Finetti theorem [20, 21].

Proof (Corollary II.2)

In the following paragraphs we prove that for every entangled state $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_m)$, there is a $\mu(\rho) > 0$ and a sequence of POVM elements $0 \leq A_n \leq \mathbb{I}$, where A_n acts on $(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_m)^{\otimes n}$, such that

$$\lim_{n\to\infty} \mathrm{tr}(A_n \rho^{\otimes n}) = 1, \tag{168}$$

and for all sequences of separable states $\{\omega_n\}_{n\in\mathbb{N}}$,

$$-\frac{\log \mathrm{tr}(A_n \omega_n)}{n} \geq \mu(\rho), \tag{169}$$

From Theorem I it will then follows that $E_R^\infty(\rho) \geq \mu(\rho) > 0$ (actually we only need Corollary III.3 here).

The A_n's are defined as follows. We apply the symmetrization operation $\hat{S}_n$ to the n individual Hilbert spaces, trace out the first αn systems ($0 < \alpha < 1$), and then measure a LOCC informationally complete POVM $\{M_k\}_{k=1}^L$ in each of the remaining $(1-\alpha)n$ systems, obtaining an empirical frequency distribution $p_{k,n}$ of the possible outcomes $\{k\}_{k=1}^L$ (see Appendix A). Using this probability distribution, we form the operator

$$L_n := \sum_{k=1}^{L} p_{k,n} M_k^*, \tag{170}$$

where $\{M_k^*\}$ is the dual set of the family $\{M_k\}$. If

$$||L_n - \rho||_1 \leq \epsilon/2, \tag{171}$$

where

$$\epsilon := \min_{\sigma\in\mathcal{S}} ||\rho - \sigma||_1 > 0, \tag{172}$$

we accept, otherwise we reject. Then we set $A_n := \hat{S}_n(\mathbb{I}^{\otimes \alpha n} \otimes \tilde{A}_n)$ as the POVM element associated to the event that we accept, where $\tilde{A}_n$ is the POVM element associated to measuring $\{M_k\}_{k=1}^L$ on each of the $(1-\alpha)n$ copies and accepting.

First, by the law of large numbers [54] and the definition of informationally complete POVMs, it is clear that $\lim_{n\to\infty} \mathrm{tr}(A_n \rho^{\otimes n}) = 1$. It thus remains to show that $\mathrm{tr}(A_n \omega_n) = \mathrm{tr}(\mathbb{I}^{\otimes \alpha n} \otimes \tilde{A}_n)\hat{S}_n(\omega_n)) \leq 2^{-\mu n}$, for a positive number μ and every sequence of separable states $\{\omega_n\}_{n\in\mathbb{N}}$.

Applying Theorem II with $k = \alpha n$ and $r = \beta n$ to $\mathrm{tr}_{1,...,\alpha n}(\hat{S}_n(\omega_n))$, we find that there is a probability measure ν such that

$$\mathrm{tr}_{1,...,\alpha n}(\hat{S}_n(\omega_n)) = \int_{\sigma\in D(\mathcal{H})} \int_{|\theta\rangle\supset\sigma} \nu(d|\theta\rangle) \pi_n^{|\theta\rangle} + X_n, \tag{173}$$

where $||X_n||_1 \leq 2^{\frac{\alpha\beta n}{3}}$ for sufficiently large n,

$$\pi_n^{|\theta\rangle} := \mathrm{tr}_E \left(|\psi_{(1-\alpha)n}^{|\theta\rangle}\rangle\langle\psi_{(1-\alpha)n}^{|\theta\rangle}| \right), \tag{174}$$

and $|\psi^{|\theta\rangle}_{(1-\alpha)n}\rangle \in |\theta\rangle^{[\otimes,(1-\alpha)n,\beta n]}$.

In the next paragraphs we show that only an exponentially small portion of the volume of ν is in a neighborhood of purifications of ρ.

Since we are measuring local POVMs, the operation $\pi \mapsto \text{tr}_{\setminus 1}(\hat{S}_n(\pi)\mathbb{I}^{\otimes \alpha n} \otimes \tilde{A}_n)$ is a stochastic LOCC map (see e.g. [23]). It hence follows from Eq. (173) that

$$\begin{aligned}\text{tr}_{\setminus 1}(\hat{S}_n(\omega_n)\mathbb{I} \otimes \tilde{A}_n) \;&=\; \int_{\sigma \in B_{2\epsilon}(\rho)} \int_{|\theta\rangle \supset \sigma} \nu(d|\theta\rangle)\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n) \\ &+\; \int_{\sigma \in \notin B_{2\epsilon}(\rho)} \int_{|\theta\rangle \supset \sigma} \nu(d|\theta\rangle)\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n) \\ &+\; \text{tr}_{\setminus 1}(X_n\mathbb{I} \otimes \tilde{A}_n) \in \text{cone}(\mathcal{S}). \end{aligned} \tag{175}$$

As $||X_n|| \leq 2^{-\alpha\beta n/3}$, we find $||\text{tr}_{\setminus 1}(X_n\mathbb{I} \otimes \tilde{A}_n)||_1 \leq 2^{-\alpha\beta n/3}$.

Furthermore, from Lemma B.1 we have that if $\text{tr}_E(|\theta\rangle\langle\theta|) \notin B_{2\epsilon}(\rho)$,

$$||\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n)||_1 = \text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n) \leq n^{d^2} 2^{-(\epsilon/K - h(\beta))(1-\alpha)n}, \tag{176}$$

where K is given by Eq. (A2) and can be taken to be such that $K \leq \dim(\mathcal{H})^4$.

Putting it all together,

$$\text{tr}_{\setminus 1}(\hat{S}_n(\omega_n)\mathbb{I} \otimes \tilde{A}_n) = \int_{\sigma \in B_{2\epsilon}(\rho)} \int_{|\theta\rangle \supset \sigma} \nu(d|\theta\rangle)\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n) + \tilde{X}_n \in \text{cone}(\mathcal{S}). \tag{177}$$

with $\tilde{X_n}$ given by the sum of the two last terms in Eq. (175), which satisfies $||\tilde{X}_n||_1 \leq 2^{-\alpha\beta n/3} + n^{d^2} 2^{-(\epsilon/K - h(\beta))(1-\alpha)n}$.

For each $\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n)$, with $\text{tr}_E(|\theta\rangle\langle\theta|) \in B_{2\epsilon}(\rho)$, we can write

$$\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes \tilde{A}_n) = \text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n) + \text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes (\tilde{A}_n - B_n)), \tag{178}$$

where B_n is the sum of the POVM elements for which the post-selected state is δ-close from the empirical state.

From Lemma B.2 we find that $\text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes (\tilde{A}_n - B_n)) \leq 2^{-M(1-\alpha)\delta^2 n}$. Therefore,

$$\begin{aligned}\text{tr}_{\setminus 1}(\hat{S}_n(\omega_n)\mathbb{I} \otimes \tilde{A}_n) \;&=\; \int_{\sigma \in D(\mathcal{H})} \int_{|\theta\rangle \supset \sigma \in B_{2\epsilon}(\rho)} \nu(d|\theta\rangle)\text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n)\rho^{|\theta\rangle} \\ &+\; \hat{X}_n \in \text{cone}(\mathcal{S}). \end{aligned} \tag{179}$$

where $\hat{X}_n$ is such that $||\hat{X}_n||_1 \leq 2^{-\alpha\beta n/3} + n^{d^2} 2^{-(\epsilon/K - h(\beta))(1-\alpha)n} + 2^{-M(1-\alpha)\delta^2 n}$ and

$$\rho^{|\theta\rangle} := \frac{\text{tr}_{\setminus 1}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n)}{\text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n)}. \tag{180}$$

Note that we have $||\rho^{|\theta\rangle} - \rho|| \leq \delta + \epsilon/2$ for every $\rho^{|\theta\rangle}$ appearing in the integral of Eq. (179). Define

$$\Lambda := \int_{\sigma \in D(\mathcal{H})} \int_{|\theta\rangle \supset \sigma \in B_{2\epsilon}(\rho)} \nu(d|\theta\rangle)\text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n). \tag{181}$$

Then,

$$\left\| \Lambda^{-1} \int_{\sigma \in D(\mathcal{H})} \int_{|\theta\rangle \supset \sigma \in B_{2\epsilon}(\rho)} \nu(d|\theta\rangle)\text{tr}(\pi_n^{|\theta\rangle}\mathbb{I} \otimes B_n)\rho^{|\theta\rangle} - \rho \right\| \leq \delta + \epsilon/2, \tag{182}$$

From Eqs. (172) and (182) it follows that $\Lambda^{-1}\int_{\sigma\in D(\mathcal{H})}\int_{|\theta\rangle\supset\sigma\in B_{2\epsilon}(\rho)}\nu(d|\theta\rangle)\mathrm{tr}(\pi_n^{|\theta\rangle}\mathbb{I}\otimes B_n)\rho^{|\theta\rangle}$ is at least $\epsilon/2-\delta$ far away from the separable states set. Using Eq. (179) we thus find that

$$\Lambda \leq (\epsilon/2-\delta)^{-1}(2^{-\alpha\beta n/3}+n^{d^2}2^{-(\epsilon/K-h(\beta))n}+n2^{-((1-\alpha)n-1)\delta^2M^{-2}}). \tag{183}$$

With this bound we finally see that

$$\begin{aligned}\mathrm{tr}(\omega_n A_n) &= \mathrm{tr}(\hat{S}_n(\omega_n)\mathbb{I}\otimes\tilde{A}_n)\\ &= \Lambda+\mathrm{tr}(\hat{X})\\ &\leq (1+(\epsilon/2-\delta)^{-1})(2^{-\alpha\beta n/3}+n^{d^2}2^{-(\epsilon/K-h(\beta))n}+n2^{-((1-\alpha)n-1)\delta^2M^{-2}})\\ &\leq 2^{-\mu n},\end{aligned} \tag{184}$$

for appropriately chosen $\alpha,\beta\in[0,1]$ and $\mu>0$. □

In the proof above the only property of the set of separable states that we used, apart from the five properties required for Theorem I to hold, was its closedness under SLOCC. It is an interesting question if such a property is really needed, or if actually the positiveness of the rate function is a generic property of any $\rho\notin\mathcal{M}$ for every family of sets satisfying Theorem I. The following example shows that this is not the case; for some choices of sets $\{\mathcal{M}_k\}$ the rate function can be zero for a state $\rho\notin\mathcal{M}$. In fact, in our example the rate function is zero for every state.

A bipartite state σ_{AB} is called n-extendible if there is a state $\tilde{\sigma}_{AB_1\ldots B_n}$ symmetric under the permutation of the B systems and such that $\mathrm{tr}_{B_2,\ldots,B_n}(\tilde{\sigma})=\sigma$. Let us denote the set of n-extendible states acting on $\mathcal{H}=\mathcal{H}_A\otimes\mathcal{H}_B$ by $\mathcal{E}_k(\mathcal{H})$. It is clear that the sets $\{\mathcal{E}_k(\mathcal{H}^{\otimes n})\}_{n\in\mathbb{N}}$ satisfy conditions 1-5 and therefore we can apply Theorem I to them. Corollary II.2 however does not hold in this case, as the sets are not closed under two-way LOCC, even though they are closed under one-way LOCC. In fact, the statement of the corollary fails dramatically in this case as it turns out that the measures $E^\infty_{\mathcal{E}_k}$ are zero for every state. This can be seen as follows: Given a state ρ, let us form the k-extendible state

$$\tilde{\rho}_{AB_1,\ldots,B_k}:=\mathbb{I}_A\otimes\hat{S}_{B_1,\ldots,B_k}\left(\rho_{AB}\otimes\left(\frac{\mathbb{I}}{d^2}\right)^{\otimes k-1}\right) \tag{185}$$

We have $\tilde{\rho}_{AB_1,\ldots,B_k}\geq\rho_{AB}\otimes\frac{\mathbb{I}}{d^2}^{\otimes k-1}/k$. Then, from the operator monotonicity of the $\log$,

$$E_{\mathcal{E}_k}(\rho)\leq S(\rho||\mathrm{tr}_{B_2,\ldots,B_n}(\tilde{\rho}))\leq k. \tag{186}$$

As the upper bound above is independent of n, we then find

$$E^\infty_{\mathcal{E}_k}(\rho)=\lim_{n\to\infty}\frac{1}{n}E^\infty_{\mathcal{E}_k}(\rho^{\otimes n})\leq\lim_{n\to\infty}\frac{k}{n}=0. \tag{187}$$

Note that as $\mathcal{E}_1$ is contained in the set of one-way undistillable states $\mathcal{C}_{\text{one-way}}$, the same is true for $E^\infty_{\mathcal{C}_{\text{one-way}}}$, i.e. it is identically zero. It is interesting that an one-way distillable state cannot be distinguished with an exponential decreasing probability of error from one-way undistillable states if we allow these to be correlated among several copies, while any entangled state can be distinguished from arbitrary sequences of separable states with exponentially accuracy. Moreover, as the set of states with a positive partial transpose (PPT) satisfy conditions 1-5 and is closed under SLOCC, every state with a non-positive partial transpose (NPPT) can be exponentially well distinguished from a sequence of PPT states. It is an intriguing open question if the same holds for distinguishing a two-way distillable state from a sequence of two-way undistillable states. Due to the conjecture existence of NPPT bound (undistillable) entanglement [55–58], property 4 might fail and therefore we do not know what happens in this case.

V. PROOF OF COROLLARY II.3

Proof (Corollary II.3)

The proof is a simple application of the well-known idea of bounding the rate of asymptotic entanglement transformations by entanglement measures (see e.g. [22, 23]). Suppose we can transform ρ into σ asymptotically, where σ is entangled. Then, for every $\epsilon > 0$ there is a sequence of LOCC maps $\{\Lambda_n\}_{n\in\mathbb{N}}$ and a sequence of integers $\{k_n\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} ||\Lambda_n(\rho^{\otimes k_n}) - \sigma^{\otimes n}||_1 = 0. \tag{188}$$

and

$$\limsup_{n\to\infty} \frac{k_n}{n} \leq R(\rho \to \sigma) + \epsilon. \tag{189}$$

From the monotonicity of the relative entropy of entanglement under LOCC [27] and its asymptotically continuity (see Lemma C.3), we find

$$\begin{aligned} E_R^\infty(\sigma) &= \limsup_{n\to\infty} \frac{1}{n} E_R(\sigma^{\otimes n}) \\ &= \limsup_{n\to\infty} \frac{1}{n} E_R(\Lambda_n(\rho^{\otimes k_n})) \\ &\leq \limsup_{n\to\infty} \frac{1}{n} E_R(\rho^{\otimes k_n}) \\ &= \limsup_{n\to\infty} \frac{k_n}{n} \limsup_{n\to\infty} \frac{1}{k_n} E_R(\rho^{\otimes k_n}) \\ &\leq (R(\rho \to \sigma) + \epsilon) E_R^\infty(\rho). \end{aligned} \tag{190}$$

As, from Corollary II.2, $E_R^\infty(\sigma) > 0$ and $\epsilon > 0$ is arbitrary, we find that indeed $R(\rho \to \sigma) > 0$. □

VI. ACKNOWLEDGMENTS

We gratefully thank Koenraad Audenaert, Nilanjana Datta, Jens Eisert, Andrzej Grudka, Masahito Hayashi, Michał and Ryszard Horodecki, Renato Renner, Shashank Virmani, Reinhard Werner, Andreas Winter and the participants in the 2009 McGill-Bellairs workshop for many interesting discussions, and an anonymous referee for filling in gaps in the proofs of Lemma III.6 and Proposition III.1, for pointing out that our main result could be extended to cover the original quantum Stein's Lemma and for many other extremely useful comments on the manuscript. This work is part of the QIP-IRC supported by EPSRC (GR/S82176/0) as well as the Integrated Project Qubit Applications (QAP) supported by the IST directorate as Contract Number 015848′ and was supported by the Brazilian agency Fundao de Amparo Pesquisa do Estado de Minas Gerais (FAPEMIG), an EPSRC Postdoctoral Fellowship for Theoretical Physics and a Royal Society Wolfson Research Merit Award.

Appendix A: Informationally Complete POVMs

An informationally complete POVM in $\mathcal{B}(\mathbb{C}^m)$ is defined as a set of positive semi-definite operators A_i forming a resolution of the identity and such that $\{A_i\}$ forms a basis for $\mathcal{B}(\mathbb{C}^m)$. Informationally complete POVMs can be explicitly constructed in every dimension (see e.g. [59]).

We say that a family $\{M_i\}$ of elements from $\mathcal{B}(\mathbb{C}^m)$ is a dual of the a family $\{M_i^*\}$ if for all $X \in \mathcal{B}(\mathbb{C}^m)$,

$$X = \sum_i \text{tr}[M_i X] M_i^*. \tag{A1}$$

The above equation implies in particular that the operator X is fully determined by the expectations values $\text{tr}[M_i X]$. Another useful property is that for every informationally complete POVM in $\mathcal{B}(\mathbb{C}^m)$ there is a real number K_m such that for every two states ρ and σ,

$$||\rho - \sigma||_1 \leq K_m ||p_\rho - p_\sigma||_1, \tag{A2}$$

with $p_\rho = \text{tr}(M_i \rho)_i$ and $p_\sigma = \text{tr}(M_i \sigma)_i$. For example, in the family of informationally complete POVM constructed in Ref. [59], $K_m \leq m^4$.

Appendix B: Exponential Quantum de Finetti Theorem

There have been several interesting recent developments on quantum versions [20, 21, 59, 60] of the seminal result by Bruno de Finetti on the characterization of exchangeable probability distributions [61]. Here we state an exponential version of the theorem for quantum states, recently proved by Renner [20, 21].

Theorem II *[20, 21, 62] For any state $|\psi_{n+k}\rangle \in Sym(\mathcal{H}^{\otimes n+k})$ there exists a measure μ over $\mathcal{H}$ and for each pure state $|\theta\rangle \in \mathcal{H}$ another pure state $|\psi_n^\theta\rangle \in |\theta\rangle^{[\otimes,n,r]}$ such that*

$$\left\| tr_{1,\ldots,k}(|\psi_{n+k}\rangle\langle\psi_{n+k}|) - \int \mu(d|\theta\rangle) |\psi_n^\theta\rangle\langle\psi_n^\theta| \right\|_1 \leq n^{\dim(\mathcal{H})} 2^{-\frac{k(r+1)}{2(n+k)}}. \tag{B1}$$

The generalization of Theorem II to permutation-symmetric mixed states goes as follows. First, we use the fact that every permutation-symmetric mixed state ρ_{n+k}^S acting on $\mathcal{H}_S^{\otimes n+k}$ has a symmetric purification $|\psi\rangle_{n+k}^{SE} \in (\mathcal{H}_S \otimes \mathcal{H}_E)^{\otimes n+k}$, with $\dim(\mathcal{H}_E) = \dim(\mathcal{H}_S)$ (see e.g. Lemma 4.2.2 of Ref. [20]). Then we apply Theorem II to $|\psi\rangle_{n+k}^{SE}$ and use the contractiveness of the trace norm under the partial trace to find

$$\left\| \text{tr}_{1,\ldots,k}(\rho_{n+k}) - \int \mu(d\sigma) \rho_\sigma \right\|_1 \leq n^{\dim(\mathcal{H})^2} 2^{-\frac{k(r+1)}{2(n+k)}} \tag{B2}$$

where

$$\rho_\sigma := \text{tr}_E(|\psi_n^{|\theta\rangle}\rangle\langle\psi_n^{|\theta\rangle}|), \tag{B3}$$

with $\sigma := \text{tr}_E(|\theta\rangle\langle\theta|)$ and

$$\mu(d\sigma) := \int_{|\theta\rangle \supset \sigma} \mu(d|\theta\rangle). \tag{B4}$$

In the equation above $|\theta\rangle \supset \sigma$ means that the integration is taken with respect to the purifying system E and runs over all purifications of σ.

a. Chernoff-Hoeffding Bound for Almost Power States

The states $\mathrm{tr}_E(|\psi_n^\theta\rangle\langle\psi_n^\theta|)$ behave like $\mathrm{tr}_E(|\theta\rangle\langle\theta|)^{\otimes n}$ in many respects. One example is the case where the same POVM is measured on all the n copies.

Let $\{M_\omega\}_{\omega\in\mathcal{W}}$ be a POVM on $\mathcal{H}$ and define its induced probability distribution on $|\theta\rangle$ by $P_M(|\theta\rangle\langle\theta|) = \{\langle\theta|M_\omega|\theta\rangle\}_{\omega\in\mathcal{W}}$. Theorems 4.5.2 of Ref. [20] and its reformulation as Lemma 2 of Ref. [47] show the following.

Lemma B.1 *[20, 47] Let $|\Psi_n\rangle$ be a vector from $|\theta\rangle^{[\otimes,n,r]}$ with $0 \le r \le \frac{n}{2}$ and $\{M_\omega\}_{\omega\in\mathcal{W}}$ be a POVM on $\mathcal{H}$.*

$$Pr\left(\|P_M(|\theta\rangle\langle\theta|) - P_M(|\Psi_n\rangle\langle\Psi_n|)\|_1 > \delta\right) \le 2^{-n\left(\frac{\delta^2}{4} - h\left(\frac{r}{n}\right)\right) + |\mathcal{W}|\log(\frac{n}{2}+1)} \tag{B5}$$

where $P_M(|\Psi_n\rangle\langle\Psi_n|)$ is the frequency distribution of outcomes of $M^{\otimes n}$ applied to $|\Psi_n\rangle\langle\Psi_n|$, and the probability is taken over those outcomes.

This Lemma shows that apart from the factor $h(r/n)$, which in an usual application of Lemma B.1 is taken to be vanishing small, the statistics of the frequency distribution obtained by measuring an almost power state along $|\theta\rangle$ is the same as if we had $|\theta\rangle^{\otimes n}$.

1. Post-selected states

The next lemma, due to König and Renner, appeared in [59] as Theorem A.1 and is used in the proof of Corollary II.2.

Lemma B.2 *[59] Let $\rho_{m+1} \in \mathcal{D}(\mathcal{H}^{\otimes m+1})$ be a permutation-symmetric state and $\mathcal{M} := \{M_k\}$ an informationally complete POVM in $\mathcal{H}$. Consider the probability distribution*

$$p(i_1, ..., i_m) := tr(\mathbb{I}\otimes M_{i_1}\otimes M_{i_2}\otimes ...\otimes M_{i_m}\rho_{m+1}), \tag{B6}$$

associated to the measurement of $\mathcal{M}$ in m of the subsystems of ρ_{m+1}. Define the post-selected states

$$\pi_{i_1,...,i_m} := \frac{tr_{\setminus 1}(\mathbb{I}\otimes M_{i_1}\otimes M_{i_2}\otimes ...\otimes M_{i_m}\rho_{m+1})}{tr(\mathbb{I}\otimes M_{i_1}\otimes M_{i_2}\otimes ...\otimes M_{i_m}\rho_{m+1})} \tag{B7}$$

and let $L_m^{i_1,...,i_m}$ be the estimated state when the sequence of outcome $\{i_1, ..., i_m\}$ is obtained. Define $\mathcal{R}$ as the set of all outcome sequences such that

$$||L_m^{i_1,...,i_m} - \pi_{i_1,...,i_m}||_1 \ge \delta. \tag{B8}$$

Then there is a $M > 0$ (only depending on the dimension of $\mathcal{H}$ and on the POVM $\mathcal{M}$) such that

$$\sum_{(i_1,...,i_m)\in\mathcal{R}} p(i_1, ..., i_m) \le 2^{-Mm\delta^2}. \tag{B9}$$

Appendix C: Useful Results

Defining the fidelity $F(\rho,\sigma) = \mathrm{tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})$ we find [45]

Lemma C.1 *For every $\rho, \sigma \in \mathcal{D}(\mathcal{H})$,*

$$1 - F(\rho,\sigma) \leq \frac{1}{2}||\rho - \sigma||_1 = tr(\rho - \sigma)_+ \leq \sqrt{1 - F(\rho,\sigma)^2}. \tag{C1}$$

Lemma C.2 *For A, B positive semidefinite and Λ a trace-preserving completely positive map,*

$$||\Lambda(A)||_1 \leq ||A||_1, \quad tr(\Lambda(A))_+ \leq tr(A)_+, \quad F(\Lambda(A),\Lambda(B)) \geq F(A,B). \tag{C2}$$

Let $E : \mathcal{D}(\mathcal{H}) \to \mathbb{R}_+$. We say E is *asymptotically continuous* if for every $\rho, \sigma \in \mathcal{D}(\mathcal{H})$,

$$|E(\rho) - E(\sigma)| \leq \log(\dim(\mathcal{H}))f(||\rho - \sigma||_1), \tag{C3}$$

for a real-valued function $f : \mathbb{R}_+ \to \mathbb{R}_+$ *independent* of $\dim(\mathcal{H})$ and such that $\lim_{x\to 0} f(x) = 0$. Although not strictly needed, we will also demand that f is monotonic increasing, in order to simplify some of the proofs.

The next Lemma is due to Synak-Radtke and Horodecki [63] and Christandl [64].

Lemma C.3 *[63, 64] For every family of sets $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ satisfying properties 1-4, $E_{\mathcal{M}_n}$ and $E^\infty_{\mathcal{M}}$, given by Eqs. (6) and (10), respectively, are asymptotically continuous.*

In Ref. [63] it was shown that the minimum relative entropy over any convex set that includes the maximal mixed state is asymptotically continuous. It is simple to check that their proof goes through if instead of the maximally mixed state, the set contains $\sigma^{\otimes n}$, for a full rank state σ. For $E_{\mathcal{M}_n}$ the lemma then follows from properties 1 and 2. In Proposition 3.23 of Ref. [64], in turn, it was proven that E^∞_R is asymptotically continuous. It is straightforward to note that the proof actually applies to the regularized minimum relative entropy over any family of sets satisfying properties 1-4. Moreover, the functions f used in [63] and [64] turn out to be monotonic increasing.

The next two lemmata will play an important role in the proof of Proposition II.1. The first, due to Ogawa and Nagaoka, appeared in Ref. [6] as Theorem 1 and was the key element for establishing the strong converse of quantum Stein's Lemma.

Lemma C.4 *[6] Given two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ such that $supp(\rho) \subseteq supp(\sigma)$ and a real number λ,*

$$tr(\rho^{\otimes n} - 2^{\lambda n}\sigma^{\otimes n})_+ \leq 2^{-n(\lambda s - \psi(s))}, \tag{C4}$$

for every $s \in [0,1]$. The function $\psi(s)$ is defined as

$$\psi(s) := \log(tr(\rho^{1+s}\sigma^{-s})). \tag{C5}$$

Note that $\psi(0) = 0$ and $\psi'(0) = S(\rho||\sigma)$. Hence, if $\lambda > S(\rho||\sigma)$, $\mathrm{tr}(\rho^{\otimes n} - 2^{\lambda n}\sigma^{\otimes n})_+$ goes to zero exponentially fast in n.

The next Lemma, due to Datta and Renner [65], appeared in Ref. [65] as Lemma 5 and is used in the proofs of Propositions II.1 and III.1.

Lemma C.5 *[65] Let $\rho \in \mathcal{D}(\mathcal{H})$ and Y, Δ be positive semidefinite operators such that $\rho \leq Y + \Delta$ and $tr(\Delta) < 1$. Then there exists a state $\tilde{\rho} \in \mathcal{D}(\mathcal{H})$ such that*

$$\tilde{\rho} \leq (1 - tr(\Delta))^{-1}Y, \tag{C6}$$

and

$$F(\rho,\tilde{\rho}) \geq 1 - tr(\Delta), \quad ||\rho - \tilde{\rho}||_1 \leq 4\sqrt{tr(\Delta)}. \tag{C7}$$

Proof Let $T := Y^{1/2}(Y+\Delta)^{-1/2}$, $\rho' := T\rho T^\dagger$ and set $\tilde{\rho} := \rho'/\mathrm{tr}(\rho')$. As $\rho \leq Y + \Delta$, we find

$$\rho' = T\rho T^\dagger \leq Y \tag{C8}$$

and hence

$$\tilde{\rho} = \mathrm{tr}(\rho')\rho' \leq \mathrm{tr}(T^\dagger T\rho)Y. \tag{C9}$$

Let us show that

$$\mathrm{tr}(T^\dagger T\rho) \geq 1 - \mathrm{tr}(\Delta). \tag{C10}$$

Eq. (C6) then follows from Eqs. (C9,C10). Note that

$$T^\dagger T = (Y+\Delta)^{-1/2}Y(Y+\Delta)^{-1/2} \leq \mathbb{I}. \tag{C11}$$

Then, using the inequality $\rho \leq Y + \Delta$,

$$\mathrm{tr}((\mathbb{I} - T^\dagger T)\rho) \leq \mathrm{tr}(Y+\Delta) - \mathrm{tr}((Y+\Delta)T^\dagger T) = \mathrm{tr}(\Delta), \tag{C12}$$

from which Eq. (C10) follows.

In the proof of Lemma 5 of Ref. [65] it is proven that $F(\rho,\rho') \geq 1 - \mathrm{tr}(\Delta)$. Hence

$$F(\rho,\tilde{\rho}) = \mathrm{tr}(\rho')^{-1/2}F(\rho,\rho') \geq F(\rho,\rho') \geq 1 - \mathrm{tr}(\Delta), \tag{C13}$$

where we used that $\mathrm{tr}(\rho') = \mathrm{tr}(T^\dagger T\rho) \leq 1$, which follows from $T^\dagger T \leq \mathbb{I}$. The inequality for the trace norm follows from Eq. (C.1). □

We also make use of the following simple lemma.

Lemma C.6 *Let $|\Psi\rangle \in \mathcal{H}$ be such that $|\Psi\rangle := \sum_{k\in\mathcal{X}} |\psi_k\rangle$. Then*

$$|\Psi\rangle\langle\Psi| \leq |\mathcal{X}| \sum_{k\in\mathcal{X}} |\psi_k\rangle\langle\psi_k| \tag{C14}$$

Proof For every $|\theta\rangle \in \mathcal{H}$, $|\langle\theta|(|\psi_k\rangle\langle\psi'_k|)|\theta\rangle| = |\langle\theta|\psi_k\rangle||\langle\theta|\psi'_k\rangle|$. Then,

$$\begin{aligned}
\langle\theta|\left(|\Psi\rangle\langle\Psi|\right)|\theta\rangle &= \left|\sum_{k,k'}\langle\theta|\left(|\psi_k\rangle\langle\psi'_k|\right)|\theta\rangle\right| \\
&\leq |\mathcal{X}|^2 \sum_{k,k'} \frac{1}{|\mathcal{X}|^2}\sqrt{\langle\theta|\left(|\psi_k\rangle\langle\psi_k|\right)|\theta\rangle\langle\theta|\left(|\psi'_k\rangle\langle\psi'_k|\right)|\theta\rangle} \\
&\leq |\mathcal{X}|^2 \sqrt{\sum_{k,k'} \frac{1}{|\mathcal{X}|^2}\langle\theta|\left(|\psi_k\rangle\langle\psi_k|\right)|\theta\rangle\langle\theta|\left(|\psi'_k\rangle\langle\psi'_k|\right)|\theta\rangle} \\
&= |\mathcal{X}|\langle\theta|\left(\sum_{k\in\mathcal{X}}|\psi_k\rangle\langle\psi_k|\right)|\theta\rangle,
\end{aligned} \tag{C15}$$

where the inequality in the third line follows from Jensen's inequality. □

The final lemma, adapted from lemma 4.1.2 of [66], is used in the proof of Lemma III.6.

Lemma C.7 *Given two probability distributions $p, q : \{1, ..., n\} \to \mathbb{R}$ and real numbers $0 \leq \lambda_i \leq 1$, $i \in \{1, ..., n\}$, and μ,*

$$\sum_{i=1}^{n} \lambda_i (p(i) - 2^{\mu} q(i)) \leq \Pr_{\{p\}} \left(i : \log \frac{p(i)}{q(i)} \geq \mu \right). \tag{C16}$$

Proof The lemma can be proved by the following chain of inequalities

$$\begin{aligned} \Pr_{\{p\}} \left(i : \log \frac{p(i)}{q(i)} \geq \mu \right) &= \sum_{i : p(i) \geq 2^{\mu} q(i)} p(i) \\ &\geq \sum_{i : p(i) \geq 2^{\mu} q(i)} \lambda_i p(i) \\ &\geq \sum_{i : p(i) \geq 2^{\mu} q(i)} \lambda_i (p(i) - 2^{\mu} q(i)) \\ &\geq \sum_{i} \lambda_i (p(i) - 2^{\mu} q(i)). \end{aligned} \tag{C17}$$

In the first inequality we used that $0 \leq \lambda_i \leq 1$, in the second that $q(i) \geq 0$, and in the last that we add negative terms corresponding to the i's for which $p(i) < 2^{\mu} q(i)$. □

[1] T.M. Cover and J.A. Thomas. Elements of Information Theory. Series in Telecommunication. John Wiley and Sons, New York, 1991.

[2] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Stat. **23**, 493 (1952).

[3] I. Csiszár and G. Longo. On the error exponent for source coding and for testing. simple statistical hypotheses. Studia Sci. Math. Hungarica **6**, 181 (1971).

[4] R.E. Blahut. Hypothesis testing and information theory. IEEE Trans. Inf. Theo. **20**, 405 (1974).

[5] F. Hiai and D. Petz. The proper formula for the relative entropy an its asymptotics in quantum probability. Comm. Math. Phys. **143**, 99 (1991).

[6] T. Ogawa and H. Nagaoka. Strong Converse and Stein's Lemma in the Quantum Hypothesis Testing. IEEE Trans. Inf. Theo. **46**, 2428 (2000).

[7] M. Hayashi. Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing. J. Phys. A: Math. Gen. **35**, 10759 (2002).

[8] T. Ogawa and M. Hayashi. On error exponents in quantum hypothesis testing. IEEE Trans. Inf. Theo. **50**, 1368 (2004).

[9] M. Nussbaum and A. Szkola. The Chernoff lower bound for symmetric quantum hypothesis testing. Ann. Stat. **37**, 1040 (2009).

[10] K.M.R. Audenaert, J. Calsamiglia, Ll. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, F. Verstraete. The Quantum Chernoff Bound. Phys. Rev. Lett. **98**, 160501 (2007).

[11] H. Nagaoka. The Converse Part of The Theorem for Quantum Hoeffding Bound. quant-ph/0611289.

[12] H. Nagaoka and M. Hayashi. An Information-Spectrum Approach to Classical and Quantum Hypothesis Testing for Simple Hypotheses. IEEE Trans. Inf. Theo. **53**, 534 (2007).

[13] K.M.R. Audenaert, M. Nussbaum, A. Szkola, F. Verstraete. Asymptotic Error Rates in Quantum Hypothesis Testing. Comm. Math. Phys. **279**, 251 (2008).

[14] M. Hayashi. Error Exponent in Asymmetric Quantum Hypothesis Testing and Its Application to Classical-Quantum Channel coding. Phys. Rev. A, **76**, 062301 (2007).

[15] I. Bjelakovi'c and R. Siegmund-Schultze. An ergodic theorem for quantum relative entropy. Comm. Math. Phys. **247**, 697 (2004).
[16] I. Bjelakovic, J.-D. Deuschel, T. Krueger, R. Seiler, Ra. Siegmund-Schultze and A. Szkola. Typical support and Sanov large deviations of correlated states. Comm. Math. Phys. **279**, 559 (2008).
[17] F. Hiai, M. Mosonyi, and T. Ogawa. Error exponents in hypothesis testing for correlated states on a spin chain. J. Math. Phys. **49**, 032112 (2008).
[18] M. Mosonyi, F. Hiai, T. Ogawa, and M. Fannes. Asymptotic distinguishability measures for shift-invariant quasi-free states of fermionic lattice systems. J. Math. Phys. 49, 032112 (2008).
[19] I. Bjelaković, J.D. Deuschel, T. Krüger, R. Seiler, Ra. Siegmund-Schultze, and A. Szola. A quantum version of Sanov's theorem. Comm. Math. Phys. **260**, 659 (2005).
[20] R. Renner. Security of Quantum Key Distribution. PhD thesis ETH, Zurich 2005.
[21] R. Renner. Symmetry implies independence. Nature Physics **3**, 645 (2007).
[22] M.B. Plenio and S. Virmani. An introduction to entanglement measures. Quant. Inf. Comp. **7**, 1 (2007).
[23] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. Rev. Mod. Phys. **81**, 865 (2009).
[24] R.F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. Phys. Rev. A **40**, 4277 (1989).
[25] N. Datta. Min- and Max- Relative Entropies and a New Entanglement Measure. IEEE Trans. Inf. Theo. bf 55, 2816 (2009).
[26] V. Vedral, M.B. Plenio, M.A. Rippin and P.L. Knight. Quantifying Entanglement. Phys. Rev. Lett. **78**, 2275 (1997).
[27] V. Vedral and M.B. Plenio. Entanglement Measures and Purification Procedures. Phys. Rev. A **57**, 1619 (1998).
[28] G. Vidal and R. Tarrach. Robustness of Entanglement. Phys. Rev. A **59**, 141 (1999).
[29] A.W. Harrow and M.A. Nielsen. How robust is a quantum gate in the presence of noise? Phys. Rev. A **68**, 012308 (2003).
[30] F.G.S.L. Brandão. Quantifying entanglement with witness operators. Phys. Rev. A **72**, 022310 (2005).
[31] N. Datta. Max- Relative Entropy of Entanglement, alias Log Robustness. Int. J. Quant. Inf. **7**, 475 (2009).
[32] C. Mora, M. Piani, H.J. Briegel, Epsilon-measures of entanglement. New J. Phys. **10**, 083027 (2008).
[33] R. Renner and S. Wolf. Smooth Renyi Entropy and Applications. Proceedings of 2004 IEEE Int. Symp. Inf. Theo., 233 (2004).
[34] E.B. Davies. Linear Operators and their Spectra. Cambridge University Press (2007).
[35] K.G.H. Vollbrecht and R.F. Werner. Entanglement measures under symmetry. Phys. Rev. A **64**, 062307 (2001).
[36] V. Vedral, M.B. Plenio, K. Jacobs and P.L. Knight. Statistical Inference, Distinguishability of Quantum States, And Quantum Entanglement. Phys. Rev. A **56**, 4452 (1997).
[37] M. Piani. Relative Entropy and Restricted Measurements. Phys. Rev. Lett. **103**, 160504 (2009).
[38] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke. Irreversibility for all bound entangled states. Phys. Rev. Lett. **95**, 190501 (2005).
[39] F.G.S.L. Brandão and M.B. Plenio. A Reversible Theory of Entanglement and its Relation to the Second Law. Commun. Math. Phys. 295, **829** (2010).
[40] F.G.S.L. Brandão and M.B. Plenio. Entanglement Theory and the Second Law of Thermodynamics. Nature Physics **4**, 873 (2008).
[41] M. Horodecki. Quantum entanglement: Reversible path to thermodynamics. Nature Physics **4**, 833 (2008).
[42] S. Boyd and L. Vandenberghe. Convex optimization. Cambridge University Press, Cambridge, 2000.
[43] R. Bathia. Matrix Analysis (Graduate Texts in Mathematics). Springer, 1996.
[44] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. Locking entanglement measures with a single qubit. Locking entanglement measures with a single qubit. Phys. Rev. Lett. **94**, 200501 (2005).
[45] A. Uhlmann. The "transition probability" in the state space of a *-algebra. Rep. Math. Phys. 9 (1976).
[46] W. Fulton and J. Harris. Representation Theory: A First Course. Springer, New York, 1991.
[47] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim. Quantum key distribution based on private states: unconditional security over untrusted channels with zero quantum capacity. IEEE Trans. Inf. Theory **54**, 2604 (2008).
[48] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim. Unconditional privacy over

channels which cannot convey quantum information. Phys. Rev. Lett. **100**, 110502 (2008).
[49] A. Dembo and O. Zeitouni. Large deviations techniques and applications. Springer-Verlag (1998).
[50] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. Lin. Alg. Appl. **26**, 203 (1979).
[51] A. Jencova and M.B. Ruskai. A Unified Treatment of Convexity of Relative Entropy and Related Trace Functions, with Conditions for Equality. arXiv:0903.2895.
[52] M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. Comm. Math. Phys. **269**, 107 (2007).
[53] M. Fannes. A continuity property of the entropy density for spin lattice systems. Comm. Math. Phys. **31**, 291 (1973).
[54] R.M. Dudley. Real Analysis and Probability. Cambridge University Press (2002).
[55] D.P. DiVincenzo, P.W. Shor, J.A. Smolin, B.M. Terhal, A.V. Thapliyal. Evidence for Bound Entangled States with Negative Partial Transpose. Phys. Rev. A **61**, 062312 (2000).
[56] W. Dür, J.I. Cirac, M. Lewenstein, and D. Bruss. Distillability and partial transposition in bipartite systems. Phys. Rev. A **61**, 062313 (2000).
[57] L. Clarisse. Entanglement Distillation; A Discourse on Bound Entanglement in Quantum Information Theory. quant-ph/0612072.
[58] F.G.S.L. Brandão and J. Eisert. Correlated entanglement distillation and the structure of the set of undistillable states. J. Math. Phys. **49**, 042102 (2008).
[59] R. König and R. Renner. A de Finetti representation for finite symmetric quantum states. J. Math. Phys. **46**, 122108 (2005).
[60] M. Christandl, R. König, G. Mitchison, and R. Renner. One-and-a-half quantum de Finetti theorems. Comm. Math. Phys. **273**, 473 (2007).
[61] B. de Finetti. La prévision: ses lois logiques, ses sources subjectives. Ann. Inst. Henri Poincaré **7**, 1 (1937).
[62] R. König and G. Mitchison. A most compendious and facile quantum de Finetti theorem. J. Math. Phys. **50**, 012105 (2009).
[63] B. Synak-Radtke and M. Horodecki. On asymptotic continuity of functions of quantum states. J. Phys. A: Math. Gen. **39**, 423 (2006).
[64] M. Christandl. The Structure of Bipartite Quantum States - Insights from Group Theory and Cryptography. PhD thesis, February 2006, University of Cambridge. quant-ph/0604183.
[65] N. Datta and R. Renner. Smooth Renyi Entropies and the Quantum Information Spectrum. IEEE Trans. Inf. Theory **55**, 2807 (2009).
[66] T.S. Han. Information-spectrum Methods in Information Theory. Springer, 2003.