

## Distilling entanglement from arbitrary resources

Francesco Buscemi<sup>1,a)</sup> and Nilanjana Datta<sup>2,b)</sup>

<sup>1</sup>*Institute for Advanced Research, University of Nagoya, Chikusa-ku, Nagoya 464-8601, Japan*

<sup>2</sup>*Statistical Laboratory, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, United Kingdom*

(Received 9 June 2010; accepted 5 August 2010; published online 11 October 2010)

We obtain the general formula for the optimal rate at which singlets can be distilled from any given noisy and arbitrarily correlated entanglement resource by means of local operations and classical communication (LOCC). Our formula, obtained by employing the quantum information spectrum method, reduces to that derived by Devetak and Winter [Proc. R. Soc. London, Ser. A **461**, 207 (2005)], in the special case of an independent and identically distributed resource. The proofs rely on a one-shot version of the so-called “hashing bound,” which, in turn, provides bounds on the one-shot distillable entanglement under general LOCC. © 2010 American Institute of Physics. [doi:10.1063/1.3483717]

### I. INTRODUCTION

A fundamental problem in entanglement theory is to determine how to optimally convert entanglement, shared between two distant parties, Alice and Bob, from one form to another. Entanglement manipulation is the process by which Alice and Bob convert an initial bipartite state  $\rho_{AB}$ , which they share, into a required target state  $\sigma_{AB}$  using local operations and classical communication (LOCC). If the target state  $\sigma_{AB}$  is a maximally entangled state, then the protocol is called entanglement distillation, whereas if the initial state  $\rho_{AB}$  is a maximally entangled state, then the protocol is called entanglement dilution. Optimal rates of these protocols, referred to as the *distillable entanglement* and *entanglement cost* of the state  $\rho_{AB}$ , respectively, were originally evaluated under (i) the assumption that the entanglement resource accessible to Alice and Bob was independent and identically distributed (i.i.d.), that is, it consisted of multiple, independent, and identical copies, i.e., tensor products  $\rho_{AB}^{\otimes n}$ , of the initial bipartite state, and under (ii) the requirement that the final state of the protocol is equal to  $n$  copies of the desired target state  $\sigma_{AB}^{\otimes n}$  with asymptotically vanishing error in the limit  $n \rightarrow \infty$ . The distillable entanglement and entanglement cost computed in this manner are two asymptotic measures of entanglement of the state  $\rho_{AB}$ . Moreover, in the case in which  $\rho_{AB}$  is pure, these two measures of entanglement coincide and are equal to the von Neumann entropy of the reduced state on any one of the subsystems,  $A$  or  $B$ .

The practical ability to transform entanglement from one form to another is useful for many applications in quantum information theory. However, it is not always justified to assume that the entanglement resource available consists of states that are multiple copies (and hence tensor products) of a given entangled state. More generally, an entanglement resource is characterized by an arbitrary sequence of bipartite states that are not necessarily of the tensor product form. Sequences of bipartite states on  $AB$  are considered to exist on Hilbert spaces  $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$  for  $n \in \{1, 2, 3, \dots\}$ .

The asymptotic entanglement cost of such an arbitrary sequence of pure bipartite states was evaluated in Ref. 1, whereas the corresponding cost for the more general case in which the states in the sequence are allowed to be mixed was evaluated in Ref. 2. As regards entanglement

<sup>a)</sup>Electronic mail: buscemi@iar.nagoya-u.ac.jp.

<sup>b)</sup>Electronic mail: n.datta@statslab.cam.ac.uk.

distillation, Hayashi<sup>3</sup> evaluated the optimal rate of entanglement distillation for an arbitrary sequence of pure states. Moreover, Matsumoto<sup>4</sup> considered the case in which the input mixed states are supported on the symmetric subspace. This was further generalized by Brandão and Eisert, who considered the case of permutationally invariant mixed states.<sup>5</sup>

In this paper, we evaluate the asymptotic distillable entanglement for a sequence of arbitrary states in two different scenarios: (i) under one-way (or forward) LOCC, that is, when classical information can only be sent from Alice to Bob, and (ii) under two-way (or general) LOCC, that is, when both Alice and Bob can send classical information to each other (possibly multiple times). The resulting expressions for the distillable entanglement constitute the main results of this paper.

A useful tool for the study of entanglement manipulation in this general scenario is provided by the information spectrum method.<sup>10</sup> This method was introduced in classical information theory by Verdu and Han and has been extended to quantum information theory first by Ogawa, Nagaoka, and Hayashi. The power of the information spectrum approach comes from the fact that it does not depend on the specific structure of sources, channels, or entanglement resources employed in information theoretical protocols.

An important step on the way to our main result is to obtain bounds on the distillable entanglement in the “one-shot” scenario, in which Alice and Bob aim to convert a *single* copy of a desired target state  $\rho_{AB}$ , which they share, into a maximally entangled state using LOCC. The logarithm of the maximum rank of the maximally entangled state, which can be thus obtained with a fixed, finite accuracy, is defined as the one-shot distillable entanglement of  $\rho_{AB}$ . Our first result in this context is the one-shot analog of the well-known Hashing bound,<sup>23</sup> which provides a lower bound on the distillable entanglement under one-way LOCC. Further, we obtain more stringent lower bounds on the one-shot distillable entanglement (both under one- and two-way LOCC) by first allowing the initial state to be preprocessed by means of a suitable LOCC map and then distilling entanglement from the resultant state. We also obtain upper bounds to the one-shot distillable entanglement both under one- and two-way LOCC. For the case of an arbitrary sequence of bipartite states, the lower and upper bounds, obtained in the one-shot scenario, independently converge to the expression for the distillable entanglement, in the asymptotic limit. Finally, we can retrieve the well-known expression of the distillable entanglement for an i.i.d. resource, given in terms of the regularized coherent information, from our main result by an application of the generalized Stein’s lemma.<sup>22</sup>

This paper is organized as follows. In Sec. II we introduce the necessary definitions and notations. In Sec. III we discuss the protocol of entanglement distillation, and in Sec. IV we obtain bounds on the one-shot distillable entanglement. Finally, in Sec. V we state and prove our main results and show that these reduce to the known results in the i.i.d. scenario. Appendices A and B contain some detailed derivations.

## II. DEFINITIONS AND NOTATIONS

### A. Mathematical preliminaries

Let  $\mathcal{B}(\mathcal{H})$  denote the algebra of linear operators acting on a finite-dimensional Hilbert space  $\mathcal{H}$ , and let  $\mathfrak{S}(\mathcal{H})$  denote the set of positive operators of unit trace (states) acting on  $\mathcal{H}$ . Throughout this paper, we restrict our considerations to finite-dimensional Hilbert spaces, and we take the logarithm to base 2.

For given orthonormal bases  $\{|i^A\rangle\}_{i=1}^d$  and  $\{|i^B\rangle\}_{i=1}^d$  in isomorphic Hilbert spaces  $\mathcal{H}_A \simeq \mathcal{H}_B \simeq \mathcal{H}$  of dimension  $d$ , we define a maximally entangled state (MES) of rank  $M \leq d$  to be

$$|\Psi_M^{AB}\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^M |i^A\rangle \otimes |i^B\rangle. \quad (1)$$

When  $M=d$ , for any given operator  $O \in \mathcal{B}(\mathcal{H})$ , the following relation can be shown by direct inspection:

$$(O \otimes I)|\Psi_d^{AB}\rangle = (I \otimes O^T)|\Psi_d^{AB}\rangle, \quad (2)$$

where  $I$  denotes the identity operator and  $O^T$  denotes the transposition with respect to the basis fixed by Eq. (1). Moreover, for any given pure state  $|\phi\rangle$ , we denote the projector  $|\phi\rangle\langle\phi|$  simply as  $\phi$ .

The trace distance between two operators  $A$  and  $B$  is given by

$$\|A - B\|_1 := \text{Tr}[\{A \geq B\}(A - B)] - \text{Tr}[\{A < B\}(A - B)],$$

where  $\{A \geq B\}$  denotes the projector on the subspace where the operator  $(A - B)$  is non-negative and  $\{A < B\} := I - \{A \geq B\}$ . The fidelity of two states  $\rho$  and  $\sigma$  is defined as

$$F(\rho, \sigma) := \text{Tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}] = \|\sqrt{\rho}\sqrt{\sigma}\|_1. \quad (3)$$

The trace distance between two states  $\rho$  and  $\sigma$  is related to the fidelity  $F(\rho, \sigma)$  as follows (see, e.g., Ref. 9):

$$1 - F(\rho, \sigma) \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F^2(\rho, \sigma)}, \quad (4)$$

where we use the notation  $F^2(\rho, \sigma) = (F(\rho, \sigma))^2$ .

## B. Relative entropies and coherent information

Our results on the distillable entanglement are expressed in terms of the following entropic quantities.

For any  $\rho, \sigma \geq 0$ , any  $0 \leq P \leq I$ , and any  $\alpha \in (0, \infty) \setminus \{1\}$ , we define the following entropic function (related to the quasientropies introduced by Petz in Ref. 17):

$$S_\alpha^P(\rho \parallel \sigma) := \frac{1}{\alpha - 1} \log \text{Tr}[\sqrt{P}\rho^\alpha\sqrt{P}\sigma^{1-\alpha}]. \quad (5)$$

Notice that for  $P=I$ , the function defined above reduces to the well-known Rényi relative entropy of order  $\alpha$ .

In this paper, in particular,

$$S_0^P(\rho \parallel \sigma) := \lim_{\alpha \searrow 0} S_\alpha^P(\rho \parallel \sigma) \quad (6)$$

plays an important role. Note that

$$S_0^P(\rho \parallel \sigma) = -\log \text{Tr}[\sqrt{P}\Pi_\rho\sqrt{P}\sigma], \quad (7)$$

where  $\Pi_\rho$  denotes the projector onto the support of  $\rho$ . Further,

$$S_0^I(\rho \parallel \sigma) = S_0(\rho \parallel \sigma) := -\log(\text{Tr} \Pi_\rho \sigma), \quad (8)$$

which is the relative Rényi entropy of order zero.

In the following, we obtain bounds on the distillable entanglement in terms of two “smoothed” quantities, which are derived from (6), for any  $\delta \geq 0$ , as

$$I_{0,\delta}^{A \rightarrow B}(\rho^{AB}) := \max_{\bar{\rho}^{AB} \in \mathfrak{b}(\rho^{AB}, \delta)} \min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S_0(\bar{\rho}^{AB} \parallel I_A \otimes \sigma^B) \quad (9)$$

and

$$\tilde{I}_{0,\delta}^{A \rightarrow B}(\rho^{AB}) := \max_{P \in \mathfrak{p}(\rho^{AB}, \delta)} \min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S_0^P(\rho^{AB} \parallel I_A \otimes \sigma^B), \quad (10)$$

where

$$b(\rho; \delta) := \{\sigma: \sigma \geq 0, \text{Tr}[\sigma] \leq 1, F^2(\rho, \sigma) \geq 1 - \delta^2\} \tag{11}$$

and

$$p(\rho; \delta) := \{P: 0 \leq P \leq I, \text{Tr}[P\rho] \geq 1 - \delta\}. \tag{12}$$

Note that in (11), the definition of fidelity (3) has been naturally extended to subnormalized density operators. Such smoothed quantities are needed in order to allow for a finite accuracy (i.e., nonzero error) in the protocol, which is a natural requirement in the one-shot regime.

For any given  $\delta > 0$ , we refer to  $I_{0,\delta}^{A \rightarrow B}(\rho^{AB})$  and  $\tilde{I}_{0,\delta}^{A \rightarrow B}(\rho^{AB})$  as smoothed zero-coherent information. These nomenclatures are justified by analogy with the coherent information as follows. For  $\delta = 0$ , both the above quantities reduce to

$$I_0^{A \rightarrow B}(\rho^{AB}) := \min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S_0(\rho^{AB} \| I_A \otimes \sigma^B), \tag{13}$$

where  $S_0(\rho \| \sigma)$  denotes the relative Rényi entropy of order zero of  $\rho$  with respect to  $\sigma$ . By replacing the relative Rényi entropy of order zero with the quantum relative entropy,

$$S(\rho \| \sigma) := \begin{cases} \text{Tr}[\rho \log \rho - \rho \log \sigma] & \text{if } \text{supp } \rho \subseteq \text{supp } \sigma \\ +\infty & \text{otherwise,} \end{cases} \tag{14}$$

we, in fact, obtain the usual coherent information  $I^{A \rightarrow B}(\rho^{AB})$ ,

$$\min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S(\rho^{AB} \| I_A \otimes \sigma^B) = S(\rho^{AB} \| I_A \otimes \rho^B) = S(\rho^B) - S(\rho^{AB}) := I^{A \rightarrow B}(\rho^{AB}). \tag{15}$$

Note, in particular, that for a MES of rank  $M$ , as defined by (1), (13) yields

$$I_0^{A \rightarrow B}(\Psi_M^{AB}) = I^{A \rightarrow B}(\Psi_M^{AB}) = \log M. \tag{16}$$

Further, given an  $\alpha$ -relative Rényi entropy  $S_\alpha(\rho \| \sigma)$  for a bipartite  $\rho = \rho^{AB}$ , we define the corresponding  $\alpha$ -conditional entropy as

$$H_\alpha(\rho^{AB} | \sigma^B) := -S_\alpha(\rho^{AB} \| I_A \otimes \sigma^B) \tag{17}$$

and

$$H_\alpha(\rho^{AB} | B) := \max_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} H_\alpha(\rho^{AB} | \sigma^B) = - \min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S_\alpha(\rho^{AB} \| I_A \otimes \sigma^B). \tag{18}$$

Then for any  $\delta > 0$ , the corresponding smoothed  $\alpha$ -conditional entropies  $H_\alpha^\delta(\rho^{AB} | B)$  are defined as follows:

$$H_\alpha^\delta(\rho^{AB} | B) := \begin{cases} \min_{\tilde{\rho}^{AB} \in b(\rho^{AB}; \delta)} H_\alpha(\tilde{\rho}^{AB} | B) & \text{for } 0 \leq \alpha < 1 \\ \max_{\tilde{\rho}^{AB} \in b(\rho^{AB}; \delta)} H_\alpha(\tilde{\rho}^{AB} | B) & \text{for } 1 < \alpha, \end{cases} \tag{19}$$

and the corresponding smoothed  $\alpha$ -coherent information is defined as

$$I_{\alpha,\delta}^{A \rightarrow B}(\rho^{AB}) := -H_\alpha^\delta(\rho^{AB} | B). \tag{20}$$

For  $\alpha = 0$ , this is identical to definition (9).

The following lemma, proven in Ref. 20, will play a central role in our proof.

*Lemma 1: [Quantum data-processing inequality (Ref. 20)] For any bipartite state  $\rho^{AB}$ , any completely positive, trace-preserving map  $\Phi: B \rightarrow C$ , and any  $\delta \geq 0$ , we have*

$$\tilde{I}_{0,2\sqrt{\delta}}^{A \rightarrow B}(\rho^{AB}) \geq \tilde{I}_{0,\delta}^{A \rightarrow C}((\text{id} \otimes \Phi)(\rho^{AB})).$$

### III. ENTANGLEMENT DISTILLATION: THE “ONE-SHOT” CASE

Let Alice and Bob, who are in two different locations, share a single copy of an arbitrary state  $\rho^{AB}$ . Their aim is to distill entanglement from this shared state (i.e., convert the state into a maximally entangled state) using local operation and classical communication (LOCC) only. If Alice is allowed to send classical information to Bob but not allowed to receive any from him, then the LOCC transformation is said to be one-way (or forward) and is denoted by the symbol  $\Lambda^{\rightarrow}(\rho^{AB})$ . More generally, LOCC operations in which Alice and Bob are both allowed to send classical information to each other are referred to as two-way LOCC and denoted by the symbol  $\Lambda^{\leftrightarrow}(\rho^{AB})$ . We refer to the corresponding protocols as *one-shot entanglement distillation* (under one- and two-way LOCC, respectively). Note that in a two-way LOCC, Alice and Bob are allowed to communicate with each other classically and perform local operations multiple times.

For the sake of generality, we consider the situation where for any given  $\varepsilon \geq 0$ , the final state of the protocol is  $\varepsilon$ -close to a maximally entangled state, with respect to a suitable distance measure. More precisely, we require the fidelity (3) of the final state of the protocol and a maximally entangled state to be  $\geq 1 - \varepsilon$ .

*Definition 1: ( $\varepsilon$ -achievable distillation rates)* For any given  $\varepsilon \geq 0$ , a real number  $R \geq 0$  is said to be an  $\varepsilon$ -achievable rate for two-way entanglement distillation if there exists an integer  $M \geq 2^R$  and a maximally entangled state  $\Psi_M^{A'B'}$  such that

$$F(\Lambda^{\leftrightarrow}(\rho^{AB}), \Psi_M^{A'B'}) \geq 1 - \varepsilon \quad (21)$$

for some two-way LOCC operation  $\Lambda^{\leftrightarrow}: AB \mapsto A'B'$ . An analogous definition holds for one-way LOCC.

*Definition 2: (One-shot distillable entanglement)* For any given  $\varepsilon \geq 0$ , the one-shot distillable entanglement,  $E_D^{\leftrightarrow}(\rho_{AB}; \varepsilon)$ , under two-way LOCC is the maximum of all  $\varepsilon$ -achievable two-way entanglement distillation rates. An analogous definition holds for the one-shot distillable entanglement under one-way LOCC, which is denoted by the symbol  $E_D^{\rightarrow}(\rho_{AB}; \varepsilon)$ .

*Definition 3: (One-way entanglement distillation fidelity)* Given a bipartite state  $\rho^{AB}$ , for any  $m \in \mathbb{N}$ , we define the one-way entanglement distillation fidelity as follows:

$$F_D^{\rightarrow}(\rho^{AB}; m) := \max_{\Lambda^{\rightarrow}} F(\Lambda^{\rightarrow}(\rho^{AB}), \Psi_m^{A'B'}), \quad (22)$$

where the maximization is over all one-way LOCC maps  $\Lambda^{\rightarrow}: AB \mapsto A'B'$ , and  $\Psi_m^{A'B'}$  is some MES of rank  $m$ . An analogous definition holds for the two-way entanglement distillation fidelity  $F_D^{\leftrightarrow}(\rho^{AB}; m)$ .

*Definition 4: (Completely positive instruments)* A completely positive (CP) instrument<sup>11</sup> is a family of CP maps  $\{\mathcal{E}_x\}_{x \in \mathcal{X}}$ , labeled by the parameter  $x$ , which sums up to a trace-preserving (TP) map.

Roughly speaking, we can think of an instrument as a quantum operation with both classical and quantum outputs. We denote an instrument acting on a quantum system  $A$  by the symbol  $\mathcal{I}_A: A \rightarrow A' \mathcal{X}$ , where  $A'$  and  $\mathcal{X}$  denote the systems corresponding to the quantum and classical outputs, respectively. Without loss of generality, the action of an instrument  $\mathcal{I}_A$  on the state  $\rho$  can be represented as  $\mathcal{I}_A(\rho) = \sum_{x \in \mathcal{X}} \mathcal{E}_x(\rho) \otimes |x\rangle\langle x|_{\mathcal{X}}$ , where  $|x\rangle$  are orthonormal states representing the classical register  $\mathcal{X}$  storing the measurement outcome  $x$ .

The most general one-way entanglement distillation protocol consists of Alice using a CP instrument on her part of the shared bipartite state, communicating the classical output to Bob, and Bob performing a completely positive, trace-preserving (CPTP) map on his part of the shared state accordingly.

**IV. ONE-SHOT BOUNDS ON DISTILLABLE ENTANGLEMENT**

**A. Lower bounds (direct parts)**

*Lemma 2: (One-shot hashing bound) For any given bipartite state  $\rho^{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$  and any  $\varepsilon \geq 0$ , the one-shot distillable entanglement via one-way (forward) LOCC transformations is bounded as follows:*

$$E_D^{\rightarrow}(\rho^{AB}; \varepsilon) \geq I_{0, \varepsilon/8}^{A \rightarrow B}(\rho^{AB}) + \log \left[ \frac{1}{d_A} + \frac{\varepsilon^2}{4} \right] - \Delta, \tag{23}$$

where  $I_{0, \varepsilon/8}^{A \rightarrow B}$  denotes the smoothed zero-coherent information defined by (9),  $d_A = \dim \mathcal{H}_A$ , and  $\Delta \in [0, 1]$  is a constant included to ensure that the right hand side of (23) is equal to the logarithm of an integer number.

In order to prove the above lemma, we need the following additional lemma, which is proved in Appendix A.

*Lemma 3: Given a state  $\rho^{AB}$ , for any  $\delta \geq 0$  and any positive integer  $m \leq d_A$ ,*

$$F_D^{\rightarrow}(\rho^{AB}; m) \geq 1 - 4\delta - \sqrt{m \left\{ 2^{I_{2, \delta}^{A \rightarrow E}(\rho^{AE})} - \frac{1}{d_A} \right\}}, \tag{24}$$

where  $\rho^{AE}$  is the reduced state  $\text{Tr}_B[\Omega^{ABE}]$  of any arbitrary purification  $|\Omega^{ABE}\rangle$  of  $\rho^{AB}$  and  $I_{2, \delta}^{A \rightarrow E}(\rho^{AE})$  is given by (20) for  $\alpha=2$ .

*Proof of Lemma 2:* For any fixed  $\varepsilon \geq 0$ , a positive real number  $R = \log m$  is an  $\varepsilon$ -achievable rate for one-way distillation if  $F_D^{\rightarrow}(\rho^{AB}, m) \geq 1 - \varepsilon$ . Due to Eq. (24), we know that  $R = \log m$  is achievable if

$$4\delta + \sqrt{m \left\{ 2^{I_{2, \delta}^{A \rightarrow E}(\rho^{AE})} - \frac{1}{d_A} \right\}} \leq \varepsilon. \tag{25}$$

For  $0 \leq \delta \leq \varepsilon/4$ ,  $\log m$  is achievable if, in particular,

$$m 2^{I_{2, \delta}^{A \rightarrow E}(\rho^{AE})} \leq (\varepsilon - 4\delta)^2 + \frac{1}{d_A} \tag{26}$$

since  $m/d_A \geq 1/d_A$ . Since  $E_D^{\rightarrow}(\rho^{AB}; \varepsilon)$  is defined as the maximum over all achievable rates, Eq. (26) implies that for all  $\delta \in [0, \varepsilon/4]$ ,

$$E_D^{\rightarrow}(\rho^{AB}; \varepsilon) \geq \log \left[ (\varepsilon - 4\delta)^2 + \frac{1}{d_A} \right] - I_{2, \delta}^{A \rightarrow E}(\rho^{AE}) - \Delta, \tag{27}$$

where  $\Delta$  is a positive number, less than or equal to 1, subtracted in order to make the right hand side of the above equation equal to the logarithm of an integer number.

The last ingredient needed to complete the proof of Lemma 2 is the fact that for the reduced states  $\rho^{AB}$  and  $\rho^{AE}$  of the same pure state  $|\Omega^{ABE}\rangle$ ,

$$-I_{2, \delta}^{A \rightarrow E}(\rho^{AE}) \geq I_{0, \delta}^{A \rightarrow B}(\rho^{AB}). \tag{28}$$

This inequality is stated as Lemma 12 of Appendix B, where it is proved using duality arguments along the lines following Ref. 16. The statement of Lemma 2 is finally obtained from (27) and (28) for  $\delta = \varepsilon/8$ . ■

Due to the one-shot hashing bound, Lemma 2, we know that the zero-coherent information is an achievable rate for one-way entanglement distillation. Since the zero-coherent information can, in general, increase under the action of a LOCC transformation, we can think of preprocessing the

initial state by means of a suitable LOCC map and distilling entanglement out of the preprocessed state, instead of the initial given one. This procedure leads us to the following achievable rates for one- and two-way entanglement distillation.

*Corollary 1: (Lower bounds)* Let  $\mathcal{I}_A: A \rightarrow A' \mathcal{X}$  denote an instrument on  $A$ , and let  $\Lambda_{AB}^{\rightarrow}: AB \rightarrow A' B'$  denote a one-way (from  $A$  to  $B$ ) LOCC transformation. Then,

$$\begin{aligned} E_D^{\rightarrow}(\rho^{AB}; \varepsilon) &\geq \max_{\Lambda_{AB}^{\rightarrow}} I_{0, \varepsilon/8}^{A' \rightarrow B'}(\sigma^{A' B'}) + \log \left[ \frac{1}{d_{A'}} + \frac{\varepsilon^2}{4} \right] - \Delta \\ &\geq \max_{\mathcal{I}_A} I_{0, \varepsilon/8}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}) + \log \left[ \frac{1}{d_{A'}} + \frac{\varepsilon^2}{4} \right] - \Delta', \end{aligned} \quad (29)$$

where  $\sigma^{A' B'} = \Lambda_{AB}^{\rightarrow}(\rho^{AB})$ ,  $\sigma^{A' B \mathcal{X}} = (\mathcal{I}_A \otimes \text{id}_B)(\rho^{AB})$ , and  $\Delta, \Delta' \in [0, 1]$  are included to ensure that the lower bounds (29) are each equal to the logarithm of a positive integer.

Analogously, let  $\Lambda_{AB}^{\leftrightarrow}: AB \rightarrow A' B'$  be a two-way LOCC transformation. Then,

$$E_D^{\leftrightarrow}(\rho^{AB}; \varepsilon) \geq \max_{\Lambda_{AB}^{\leftrightarrow}} I_{0, \varepsilon/8}^{A' \rightarrow B'}(\sigma^{A' B'}) + \log \left[ \frac{1}{d_{A'}} + \frac{\varepsilon^2}{4} \right] - \Delta'', \quad (30)$$

where  $\Delta'' \in [0, 1]$  is included to ensure that the lower bound is equal to the logarithm of a positive integer.

## B. Upper bounds (converse parts)

When distilling entanglement with one-way LOCC protocols, there is no need to employ a full one-way LOCC transformation when preprocessing the initial state: the following lemma shows that, in fact, an instrument on Alice's side only, followed by the communication of the outcome to Bob, suffices.

*Lemma 4: (One-way weak converse)* For any given bipartite state  $\rho^{AB}$  and any  $\varepsilon \geq 0$ ,

$$E_D^{\rightarrow}(\rho^{AB}; \varepsilon) \leq \max_{\mathcal{I}_A} \tilde{I}_{0, 4\sqrt{\varepsilon}}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}), \quad (31)$$

where the maximization is done over instruments  $\mathcal{I}_A := \{\mathcal{E}^m\}_{m \in \mathcal{X}}$  where each  $\mathcal{E}^m$  maps  $A$  to  $A'$ , and  $\sigma^{A' B \mathcal{X}} := \sum_m (\mathcal{E}^m \otimes \text{id}_B)(\rho^{AB}) \otimes |m\rangle\langle m|^{\mathcal{X}}$ .

*Proof:* Suppose that  $R$  is a one-way  $\varepsilon$ -achievable rate, i.e., there exists an integer  $M$  with  $\log M \geq R$  and a one-way forward LOCC operation  $\Lambda^{\rightarrow}: AB \mapsto A' B'$  such that

$$\langle \Psi_M^{A' B'} | \Lambda^{\rightarrow}(\rho^{AB}) | \Psi_M^{A' B'} \rangle \geq (1 - \varepsilon)^2. \quad (32)$$

The most general one-way forward LOCC operation is constructed as follows: (i) Alice applies a CP instrument on her share, (ii) she communicates the outcome  $m$  to Bob, and (iii) Bob deterministically performs a decoding operation on his share, depending on Alice's outcome. Such a procedure is conveniently represented by writing the following classical-quantum (c-q) state:

$$\tau^{A' B' \mathcal{X}} := \sum_{m \in \mathcal{X}} (\mathcal{E}^m \otimes \mathcal{D}^m)(\rho^{AB}) \otimes |m\rangle\langle m|^{\mathcal{X}}, \quad (33)$$

where  $\mathcal{D}^m$ 's are CPTP maps for all  $m$ , while  $\mathcal{E}^m$ 's are just CP maps normalized so that their sum  $\mathcal{E} := \sum_m \mathcal{E}^m$  is TP. The classical flags  $|m\rangle\langle m|^{\mathcal{X}}$  represent the classical information that Alice sends to Bob. This is the reason why we consider both systems  $B'$  and  $\mathcal{X}$  to be in Bob's hands.

By the quantum data-processing inequality (Lemma 1), we have that

$$\tilde{I}_{0,\delta}^{A' \rightarrow B' \mathcal{X}}(\tau^{A' B' \mathcal{X}}) \leq \tilde{I}_{0,2\sqrt{\delta}}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}), \tag{34}$$

where  $\sigma^{A' B \mathcal{X}} := \sum_m (\mathcal{E}^m \otimes \text{id}_B)(\rho^{AB}) \otimes |m\rangle\langle m|^{\mathcal{X}}$ . Moreover, since we assumed that  $\langle \Psi_M^{A' B'} | \tau^{A' B'} | \Psi_M^{A' B'} \rangle \geq (1-\varepsilon)^2$ , the operator

$$P := \sum_m |\Psi_M\rangle\langle \Psi_M|^{A' B'} \otimes |m\rangle\langle m|^{\mathcal{X}}$$

is such that  $\text{Tr}[P \tau^{A' B' \mathcal{X}}] \geq (1-\varepsilon)^2 \geq 1-2\varepsilon$ . Then, starting from (34) and recalling the definition (10),

$$\begin{aligned} \tilde{I}_{0,4\sqrt{\varepsilon}}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}) &\geq \tilde{I}_{0,2\varepsilon}^{A' \rightarrow B' \mathcal{X}}(\tau^{A' B' \mathcal{X}}) \geq -\max_{\omega^{B' \mathcal{X}}} \log \text{Tr}[\sqrt{P} \Pi_{\tau^{A' B' \mathcal{X}}} \sqrt{P} (\mathbb{I}_{A'} \otimes \omega^{B' \mathcal{X}})] \\ &\geq -\max_{\omega^{B' \mathcal{X}}} \log \text{Tr}[P (\mathbb{I}_{A'} \otimes \omega^{B' \mathcal{X}})] = I_0^{A' \rightarrow B' \mathcal{X}} \left( \sum_m q_m |\Psi_M\rangle\langle \Psi_M|^{A' B'} \otimes |m\rangle\langle m|^{\mathcal{X}} \right) \end{aligned} \tag{35}$$

for any probability distribution  $q_m > 0, \sum_m q_m = 1$ . Finally, since the quantum data-processing inequality also holds when the smoothing parameter is equal to zero, we have

$$I_0^{A' \rightarrow B' \mathcal{X}} \left( \sum_m q_m |\Psi_M\rangle\langle \Psi_M|^{A' B'} \otimes |m\rangle\langle m|^{\mathcal{X}} \right) \geq I_0^{A' \rightarrow B'}(\Psi_M^{A' B'}) = \log M \geq R. \tag{36}$$

Hence, we have proved that if  $R$  is an  $\varepsilon$ -achievable rate, there always exists an instrument  $\mathcal{I} = \{\mathcal{E}^m\}_{m \in \mathcal{X}}$  on  $A$  such that

$$\tilde{I}_{0,4\sqrt{\varepsilon}}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}) \geq R. \tag{37}$$

This, in turn, implies that

$$R \leq \max_{\mathcal{I}_A} \tilde{I}_{0,4\sqrt{\varepsilon}}^{A' \rightarrow B \mathcal{X}}(\sigma^{A' B \mathcal{X}}).$$

Then (31) is obtained by taking the maximum over all  $\varepsilon$ -achievable rates. ■

While for the one-way distillation scenario the preprocessing can be reduced, without loss of generality, to an instrument at Alice’s side only, in the two-way scenario we have to keep the preprocessing as general as possible. In particular, for any  $\varepsilon \geq 0$ , we obtain the following upper bound to the one-shot distillable entanglement under two-way LOCC transformations.

*Lemma 5: For any given bipartite state  $\rho^{AB}$  and any  $\varepsilon \geq 0$ , the distillable entanglement under two-way LOCC satisfies the following bound:*

$$E_D^{\leftrightarrow}(\rho^{AB}; \varepsilon) \leq \max_{\Lambda_{AB}^{\leftrightarrow}} \tilde{I}_{0,2\varepsilon}^{A' \rightarrow B'}(\omega^{A' B'}), \tag{38}$$

where the maximization is done over two-way LOCC transformations  $\Lambda_{AB}^{\leftrightarrow}$  mapping  $AB$  to  $A' B'$  and  $\omega^{A' B'} := \Lambda_{AB}^{\leftrightarrow}(\rho^{AB})$ .

*Proof:* Let  $\Lambda_{AB}^{\leftrightarrow}$  be a two-way LOCC transformation whose action on the state  $\rho^{AB}$  yields the state  $\omega^{A' B'} := \Lambda_{AB}^{\leftrightarrow}(\rho^{AB})$  such that  $F(\omega^{A' B'}, \Psi_M^{A' B'}) \geq 1-\varepsilon$  and  $R = \log M$ .

By definitions (10) and (13) of the zero-coherent information, we have

$$\begin{aligned}
R = \log M &= \min_{\sigma^{B'} \in \mathfrak{S}(\mathcal{H}_B)} [-\log \text{Tr}(\Psi_M^{A'B'}(\mathbb{I}_{A'} \otimes \sigma^{B'}))] \\
&\leq \min_{\sigma^{B'} \in \mathfrak{S}(\mathcal{H}_{B'})} [-\log \text{Tr}(\Psi_M^{A'B'} \Pi_{\omega^{A'B'}} \Psi_M^{A'B'}(\mathbb{I}_{A'} \otimes \sigma^{B'}))] \\
&\leq \max_{P \in \mathfrak{p}(\omega^{A'B'}; 2\varepsilon)} \min_{\sigma^{B'} \in \mathfrak{S}(\mathcal{H}_{B'})} [-\log \text{Tr}(\sqrt{P} \Pi_{\omega^{A'B'}} \sqrt{P}(\mathbb{I}_{A'} \otimes \sigma^{B'}))] = \tilde{I}_{0,2\varepsilon}^{A' \rightarrow B'}(\omega^{A'B'}) \\
&\leq \max_{\Lambda_{A'B'}^{\leftrightarrow}} \tilde{I}_{0,2\varepsilon}^{A' \rightarrow B'}(\omega^{A'B'}), \tag{39}
\end{aligned}$$

where the second identity follows from definition (13) of the zero-coherent information and the fact that  $\Pi_{\Psi_M^{A'B'}} = \Psi_M^{A'B'}$ ; the first inequality follows from  $\Pi_{\omega^{A'B'}} \leq \mathbb{I}$ , and the second inequality follows from the fact that  $\Psi_M^{A'B'} \in \mathfrak{p}(\omega^{A'B'}; 2\varepsilon)$  [see definition (12)] since  $\text{Tr}[\Psi_M^{A'B'} \omega^{A'B'}] \geq 1 - 2\varepsilon$ , which, in turn, follows from the fact that  $F^2(\omega^{A'B'}, \Psi_M^{A'B'}) \geq (1 - \varepsilon)^2$ . ■

## V. MAIN RESULT: EXACT ASYMPTOTIC FORMULAS FOR ARBITRARY RESOURCES

In this section, we consider entanglement distillation from arbitrary resources, comprising an arbitrary sequence of bipartite states  $\hat{\rho}_{AB} := \{\rho_{AB}^n\}_{n=1}^{\infty}$ , where  $\rho_{AB}^n \in \mathfrak{S}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n})$ . The one-way distillable entanglement rate for such a sequence is defined as

$$E_{D,\infty}^{\rightarrow}(\hat{\rho}_{AB}) := \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} E_D^{\rightarrow}(\rho_{AB}^n; \varepsilon), \tag{40}$$

and the two-way distillable entanglement  $E_{D,\infty}^{\leftrightarrow}(\hat{\rho}_{AB})$  is defined analogously.

To evaluate the distillable entanglement of such a sequence of states, we employ the well-known quantum information spectrum method.<sup>10,12</sup> Two fundamental quantities used in this approach are the *quantum spectral sup-* and *inf-divergence rates*, defined as follows.

*Definition 5: (Spectral divergence rates)* Given a sequence of states  $\hat{\rho} = \{\rho_n\}_{n=1}^{\infty}$  and a sequence of positive operators  $\hat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$ , the quantum spectral sup-(inf-)divergence rates are defined in terms of the difference operators  $\Pi_n(\gamma) = \rho_n - 2^{-n\gamma} \sigma_n$  as

$$\bar{D}(\hat{\rho} \parallel \hat{\sigma}) := \inf\{\gamma : \limsup_{n \rightarrow \infty} \text{Tr}\{\{\Pi_n(\gamma) \geq 0\} \Pi_n(\gamma)\} = 0\}, \tag{41}$$

$$\underline{D}(\hat{\rho} \parallel \hat{\sigma}) := \sup\{\gamma : \liminf_{n \rightarrow \infty} \text{Tr}\{\{\Pi_n(\gamma) \geq 0\} \Pi_n(\gamma)\} = 1\}, \tag{42}$$

respectively.

It is known that (see, e.g., Ref. 13)

$$\bar{D}(\hat{\rho} \parallel \hat{\sigma}) \geq \lim_{n \rightarrow \infty} \frac{1}{n} S(\rho_n \parallel \sigma_n) \geq \underline{D}(\hat{\rho} \parallel \hat{\sigma}). \tag{43}$$

In analogy with the usual definition of the coherent information (15), we moreover define the *spectral sup-* and *inf-coherent information rates*, respectively, as follows:

$$\bar{I}^{A \rightarrow B}(\hat{\rho}_{AB}) := \min_{\hat{\sigma}_B} \bar{D}(\hat{\rho}_{AB} \parallel \hat{\mathbb{I}}_A \otimes \hat{\sigma}_B), \tag{44}$$

$$I_{\hat{\sigma}_B}^{A \rightarrow B}(\hat{\rho}_{AB}) := \min_{\hat{\sigma}_B} D(\hat{\rho}_{AB} \| \hat{I}_A \otimes \hat{\sigma}_B), \tag{45}$$

where  $\hat{\rho}_{AB} := \{\rho_{AB}^n \in \mathfrak{S}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n})\}_{n=1}^\infty$ ,  $\hat{\sigma}_B := \{\sigma_B^n \in \mathfrak{S}(\mathcal{H}_B^{\otimes n})\}_{n=1}^\infty$ , and  $\hat{I}_A := \{I_A^{\otimes n}\}_{n=1}^\infty$ . Inequality (43) ensures that

$$\bar{I}^{A \rightarrow B}(\hat{\rho}_{AB}) \geq \lim_{n \rightarrow \infty} \frac{1}{n} I^{A \rightarrow B}(\rho_{AB}^n) \geq I_{\hat{\sigma}_B}^{A \rightarrow B}(\hat{\rho}_{AB}). \tag{46}$$

Note that in Eqs. (44) and (45) we could write the minimum instead of infimum due to Lemma 1 of Ref. 12.

Let  $\hat{\mathcal{I}}_A := \{\mathcal{I}_A^n\}_{n=1}^\infty$  denote a sequence of instruments  $\mathcal{I}_A^n: A_n \rightarrow A'_n \mathcal{X}_n$ . Our main results on the distillable entanglement for an arbitrary sequence of states  $\hat{\rho}_{AB}$  under both one- and two-way LOCC are given by the following theorem.

**Theorem 1:** *Given a sequence of bipartite states  $\hat{\rho}_{AB} := \{\rho_{AB}^n\}_{n=1}^\infty$ ,*

$$E_{D,\infty}^{\rightarrow}(\hat{\rho}_{AB}) = \max_{\hat{\mathcal{I}}_A} I_{\hat{\sigma}_{A' \mathcal{X}B}}^{A' \rightarrow B \mathcal{X}}(\hat{\rho}_{AB}) \tag{47}$$

and

$$E_{D,\infty}^{\leftrightarrow}(\hat{\rho}_{AB}) = \max_{\hat{\Lambda}_{AB}^{\leftrightarrow}} I_{\hat{\nu}_{A' B'}}^{A' \rightarrow B'}(\hat{\rho}_{AB}), \tag{48}$$

where the maximization in (47) is over all sequences of instruments,  $\hat{\mathcal{I}}_A := \{\mathcal{I}_A^n\}_{n=1}^\infty$ , and  $\hat{\sigma}_{A' \mathcal{X}B} := \{\mathcal{I}_A^n(\rho_{AB}^n)\}_{n=1}^\infty$ ; the maximization in (48) is over all sequences of two-way LOCC operations,  $\hat{\Lambda}_{AB}^{\leftrightarrow} := \{\Lambda_{AB}^n\}_{n=1}^\infty$ , and  $\hat{\nu}_{A' B'} := \{\Lambda_{AB}^n(\rho_{AB}^n)\}_{n=1}^\infty$ .

From Corollary 1 and Lemma 4, we have that, for any  $\varepsilon > 0$  and any  $n \geq 1$ ,

$$\frac{1}{n} \max_{\mathcal{I}_A^n} \tilde{I}_{0,4\sqrt{\varepsilon}}^{A' \rightarrow B \mathcal{X}}(\sigma_{A' B \mathcal{X}}^n) \geq \frac{1}{n} E_D^{\rightarrow}(\rho_{AB}^n; \varepsilon) \geq \frac{1}{n} \max_{\mathcal{I}_A^n} I_{0,\varepsilon/8}^{A' \rightarrow B \mathcal{X}}(\sigma_{A' B \mathcal{X}}^n) + \frac{1}{n} \log \left[ \frac{1}{d_{A'_n}} + \frac{\varepsilon^2}{4} \right] - \frac{\Delta'}{n}, \tag{49}$$

where  $\sigma_{A' B \mathcal{X}}^n = (\mathcal{I}_A^n \otimes \text{id}_B)(\rho_{AB}^n)$ . In the case of two-way entanglement distillation, again Corollary 1 and Lemma 5 yield

$$\frac{1}{n} \max_{\Lambda_{AB}^n} \tilde{I}_{0,2\varepsilon}^{A' \rightarrow B'}(\nu_{A' B'}^n) \geq \frac{1}{n} E_D^{\leftrightarrow}(\rho_{AB}^n; \varepsilon) \geq \frac{1}{n} \max_{\Lambda_{AB}^n} I_{0,\varepsilon/8}^{A' \rightarrow B'}(\nu_{A' B'}^n) + \frac{1}{n} \log \left[ \frac{1}{d_{A'_n}} + \frac{\varepsilon^2}{4} \right] - \frac{\Delta''}{n}. \tag{50}$$

Theorem 1 then follows rather straightforwardly by taking the limits  $\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty}$  on either sides of inequalities (49) and (50), and applying the following two lemmas (which were proved in Ref. 20).

*Lemma 6: [Direct part (Ref. 20)] Given a sequence of bipartite states  $\hat{\rho}_{AB}$ ,*

$$\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \max_{\substack{\bar{\rho}_{AB}^n \in \mathfrak{b}(\rho_{AB}^n; \delta) \\ \sigma_B^n}} \min_{\hat{\sigma}_B} \frac{1}{n} S_0(\bar{\rho}_{AB}^n \| I_A^{\otimes n} \otimes \sigma_B^n) \geq \min_{\hat{\sigma}_B} D(\hat{\rho}_{AB} \| \hat{I}_A \otimes \hat{\sigma}_B),$$

or, equivalently,  $\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} I_{0,\delta}^{A \rightarrow B}(\rho_{AB}^n) \geq I_{\hat{\sigma}_B}^{A \rightarrow B}(\hat{\rho}_{AB})$ .

*Lemma 7: [Weak converse (Ref. 20)] Given a sequence of bipartite states  $\hat{\rho}_{AB}$ ,*

$$\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \max_{P_n \in \mathfrak{p}(\rho_{AB}^n; \delta)} \min_{\hat{\sigma}_B} \frac{1}{n} S_0^P(\rho_{AB}^n \| I_A^{\otimes n} \otimes \sigma_B^n) \leq \min_{\hat{\sigma}_B} D(\hat{\rho}_{AB} \| \hat{I}_A \otimes \hat{\sigma}_B),$$

or, equivalently,  $\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \tilde{I}_{0,\delta}^{A \rightarrow B}(\rho_{AB}^n) \leq I^{A \rightarrow B}(\hat{\rho}_{AB})$ .

### A. The special case of i.i.d. resources

Let us now consider the case in which Alice and Bob share multiple, independent, and identical copies of a given bipartite state  $\rho_{AB}$ . The entanglement resource is in this case characterized by the sequence  $\hat{\rho}_{AB} := \{\rho_{AB}^{\otimes n}\}_{n=1}^{\infty}$ . The asymptotic distillable entanglement of the state  $\rho_{AB}$  can be obtained from Theorem 1 by employing the following lemma, which was proved in Ref. 21 by using the generalized Stein's lemma.<sup>22</sup>

*Lemma 8: For any given bipartite state  $\rho_{AB}$ ,*

$$\min_{\hat{\sigma}_B} D(\hat{\rho}_{AB} \| \hat{I}_A \otimes \hat{\sigma}_B) = S(\rho_{AB} \| I_A \otimes \rho_B), \quad (51)$$

where  $\hat{\rho}_{AB} = \{\rho_{AB}^{\otimes n}\}_{n=1}^{\infty}$ ,  $\hat{\sigma}_B := \{\sigma_B^n \in \mathfrak{S}(\mathcal{H}_B^{\otimes n})\}_{n=1}^{\infty}$ , and  $\hat{I}_A := \{I_A^{\otimes n}\}_{n=1}^{\infty}$ . Notice that the optimizing sequence  $\hat{\sigma}_B$  is not i.i.d. in general.

We can then retrieve the expressions for the asymptotic distillable entanglement of any arbitrary bipartite state  $\rho^{AB}$ , obtained in Ref. 23, as a corollary of our Theorem 1.

*Corollary 2: (Reference 23) For any bipartite state  $\rho^{AB}$ , the one-way distillable entanglement rate is given by*

$$E_{D,\infty}^{\rightarrow}(\rho^{AB}) = \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\mathcal{I}_A^n} I^{A' \rightarrow B\mathcal{X}}(\sigma_{A'B\mathcal{X}}^n), \quad (52)$$

where  $\sigma_{A'B\mathcal{X}}^n = (\mathcal{I}_A^n \otimes \text{id}_B)(\rho_{AB}^{\otimes n})$ . The two-way distillable entanglement rate is given by

$$E_{D,\infty}^{\leftrightarrow}(\rho^{AB}) = \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\Lambda_{AB}^n} I^{A' \rightarrow B'}(\nu_{A'B'}^n), \quad (53)$$

where  $\nu_{A'B'}^n = \Lambda_{AB}^n(\rho_{AB}^{\otimes n})$ .

*Proof:* Let  $\hat{\rho}_{AB}$  be the i.i.d. sequence  $\{\rho_{AB}^{\otimes n}\}_{n=1}^{\infty}$ . From Theorem 1, we have that

$$E_{D,\infty}^{\rightarrow}(\hat{\rho}_{AB}) \geq \max_{\mathcal{I}_A} I^{A' \rightarrow B\mathcal{X}}(\hat{\sigma}_{A'B\mathcal{X}}), \quad (54)$$

where  $\hat{\sigma}_{A'B\mathcal{X}} = (\mathcal{I}_A^{\otimes n} \otimes \text{id}_B^{\otimes n})(\rho_{AB}^{\otimes n})$ , i.e., the sequence  $\hat{\sigma}_{A'B\mathcal{X}}$  is i.i.d. Due to Lemma 8, then

$$E_{D,\infty}^{\rightarrow}(\hat{\rho}_{AB}) \geq \max_{\mathcal{I}_A} I^{A' \rightarrow B\mathcal{X}}(\sigma_{A'B\mathcal{X}}). \quad (55)$$

By a standard blocking argument, it follows that, in particular, for any  $m \geq 1$ ,

$$E_{D,\infty}^{\rightarrow}(\rho^{AB}) \geq \frac{1}{m} \max_{\mathcal{I}_A^m} I^{A' \rightarrow B\mathcal{X}}(\sigma_{A'B\mathcal{X}}^m), \quad (56)$$

where  $\sigma_{A'B\mathcal{X}}^m = (\mathcal{I}_A^m \otimes \text{id}_B^{\otimes m})(\rho_{AB}^{\otimes m})$ . By taking the limit  $m \rightarrow \infty$ , we obtain

$$E_{D,\infty}^{\rightarrow}(\rho^{AB}) \geq \lim_{m \rightarrow \infty} \frac{1}{m} \max_{\mathcal{I}_A^m} I^{A' \rightarrow B\mathcal{X}}(\sigma_{A'B\mathcal{X}}^m). \quad (57)$$

The converse direction, that is,

$$E_{D,\infty}^{\rightarrow}(\rho^{AB}) \leq \lim_{m \rightarrow \infty} \frac{1}{m} \max_{\mathcal{I}_A^m} I^{A' \rightarrow B\mathcal{X}}(\sigma_{A'B\mathcal{X}}^m), \quad (58)$$

simply comes from the fact that the  $\underline{D}(\hat{\rho} \parallel \hat{\sigma}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} S(\rho^n \parallel \sigma^n)$ .

The proof for  $E_{D,\infty}^{\leftrightarrow}(\rho^{AB})$  follows from exactly the same line of arguments.  $\blacksquare$

## ACKNOWLEDGMENTS

F.B. acknowledges support from the Program for Improvement of Research Environment for Young Researchers from Special Coordination Funds for Promoting Science and Technology (SCF) commissioned by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. N.D. acknowledges support from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant No. 213681. This work was completed when F.B. was visiting the Statistical Laboratory of the University of Cambridge.

## APPENDIX A: PROOF OF LEMMA 3

The following lemmas are employed in the proof of Lemma 3.

*Lemma 9: (References 14 and 20) For any self-adjoint operator  $X$  and any positive operator  $\xi > 0$ , we have*

$$\|X\|_1^2 \leq \text{Tr}[\xi] \text{Tr}[X \xi^{-1/2} X \xi^{-1/2}] \leq \text{Tr}[\xi] \text{Tr}[X^2 \xi^{-1}]. \quad (A1)$$

*Lemma 10: Given a tripartite pure state  $|\Omega^{A'BE}\rangle \in \mathcal{H}_{A'} \otimes \mathcal{H}_B \otimes \mathcal{H}_E$ , let  $\omega^{A'B}$ ,  $\omega^{A'E}$ , and  $\omega^E$  be its reduced states. Then, for any state  $\chi^E$ ,*

$$\max_{\mathcal{D}} F^2((\text{id}_{A'} \otimes \mathcal{D}_B)(\omega^{A'B}), \Psi_m^{A'B'}) \geq F^2(\omega^{A'E}, \tau^{A'} \otimes \chi^E), \quad (A2)$$

where  $|\Psi_m^{A'B'}\rangle \in \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$  is some fixed maximally entangled state of rank  $m$ ,  $\tau^{A'} = \text{Tr}_{B'}[|\Psi_m^{A'B'}\rangle\langle\Psi_m^{A'B'}|]$ , and  $\mathcal{D}: \mathcal{B}(\mathcal{H}_B) \mapsto \mathcal{B}(\mathcal{H}_{B'})$  denotes a CPTP map. The same holds also if the norm of the vector  $|\Omega^{A'BE}\rangle$  is not normalized to 1.

*Proof:* Fix some purification  $|\chi^{RE}\rangle \in \mathcal{H}_R \otimes \mathcal{H}_E$  of  $\chi^E$ . Then, for the fixed purification  $|\Psi_m^{A'B'}\rangle$  of  $\tau^{A'}$ , we have, by Uhlmann's theorem,<sup>8</sup> the monotonicity of the fidelity under partial trace, and Stinespring's dilation theorem,<sup>6</sup>

$$\begin{aligned} F^2(\omega^{A'E}, \tau^{A'} \otimes \chi^E) &= \max_{|\varphi^{A'B'RE}\rangle} F^2(\varphi^{A'B'RE}, \Psi_m^{A'B'} \otimes \chi^{RE}) \\ &\quad \text{Tr}_{B'R}[\varphi^{A'B'RE}] = \omega^{A'E} \\ &= \max_{\substack{V: B \rightarrow B'R \\ V^\dagger V = I_B}} F^2((I^{A'} \otimes V_B \otimes I^E) \Omega^{A'BE} (I^{A'} \otimes V_B^\dagger \otimes I^E), \Psi_m^{A'B'} \otimes \chi^{RE}) \\ &\leq \max_{\mathcal{D}} F^2((\text{id}_{A'} \otimes \mathcal{D}_B)(\omega^{A'B}), \Psi_m^{A'B'}), \end{aligned} \quad (A3)$$

where  $\mathcal{D}: \mathcal{B}(\mathcal{H}_B) \mapsto \mathcal{B}(\mathcal{H}_{B'})$  denotes a CPTP map. In the second equality of (A3) we used the fact that all possible purifications of a given mixed state ( $\omega^{A'E}$ , in our case) are related by some local isometry acting on the purifying system only (i.e., subsystem  $B$ ).  $\blacksquare$

*Lemma 11: For any  $P, Q \geq 0$ ,*

$$F(P, Q) := \|\sqrt{P}\sqrt{Q}\|_1 \geq \frac{\text{Tr } P + \text{Tr } Q}{2} - \frac{1}{2} \|P - Q\|_1. \quad (A4)$$

*Proof:* By adapting the proof in, e.g., Ref. 9, we see that

$$F(P, Q) = \min_{\{E_m\}: \text{POVM}} \sum_m \sqrt{p_m} \sqrt{q_m}, \tag{A5}$$

where  $p_m := \text{Tr}[E_m P]$  and  $q_m := \text{Tr}[E_m Q]$ . Also,

$$\|P - Q\|_1 = \max_{\{E_m\}: \text{POVM}} \sum_m |p_m - q_m|. \tag{A6}$$

Again according to Ref. 9, let  $\{\bar{E}_m\}$  be the positive operator valued measure (POVM) achieving  $F(P, Q)$ , and let  $\bar{p}_m$  and  $\bar{q}_m$  be the corresponding coefficients. Then,

$$\begin{aligned} \|P - Q\|_1 &\geq \sum_m |\bar{p}_m - \bar{q}_m| = \sum_m |\sqrt{\bar{p}_m} - \sqrt{\bar{q}_m}| \cdot |\sqrt{\bar{p}_m} + \sqrt{\bar{q}_m}| \geq \sum_m (\sqrt{\bar{p}_m} - \sqrt{\bar{q}_m})^2 \\ &= \sum_m \bar{p}_m + \sum_m \bar{q}_m - 2 \sum_m \sqrt{\bar{p}_m} \sqrt{\bar{q}_m} = \text{Tr } P + \text{Tr } Q - 2F(P, Q). \end{aligned} \tag{A7}$$

■

*Proof of Lemma 3:* The most general transformation composed of local operations and forward classical communication (one-way LOCC) can be written as

$$\Lambda^\rightarrow(\rho^{AB}) = \int (\mathcal{E}_\mu \otimes \mathcal{D}_\mu)(\rho^{AB}) d\mu, \tag{A8}$$

where  $d\mu$  is an appropriate measure, the  $\mathcal{D}_\mu: B \rightarrow B'$  are CPTP maps for all  $\mu$ , while the  $\mathcal{E}_\mu: A \rightarrow A'$  are CP maps normalized so that  $\mathcal{E} := \int \mathcal{E}_\mu d\mu$  is TP. The physical interpretation of such a transformation is that (i) Alice performs a measurement on her share, (ii) she communicates the outcome  $\mu$  to Bob, and (iii) Bob deterministically performs a decoding operation on his share, depending on Alice's outcome.

In the following, we will construct one particular one-way LOCC and evaluate how good that is for distilling entanglement. Let us fix the value of the positive integer  $m \leq d_A$  and define

$$\mathcal{E}_g(\rho^A) := \frac{d_A}{m} P_m^A U_g^A \rho^A (U_g^A)^\dagger (P_m^A)^\dagger, \tag{A9}$$

where  $U_g^A$  is a unitary representation of the element  $g$  of the group  $\text{SU}(d_A)$  and

$$P_m^A = \sum_{i=1}^m |i^A\rangle\langle i^A|, \tag{A10}$$

the vectors  $|i^A\rangle$ ,  $i = 1, \dots, d_A$ , being the same as in Eq. (1). Then, by introducing the Haar measure  $dg$  on  $\text{SU}(d_A)$ , it is a standard calculation to check that

$$\int \mathcal{E}_g(\rho^A) dg = \frac{d_A}{m} P_m^A \left( \int U_g^A \rho^A (U_g^A)^\dagger dg \right) (P_m^A)^\dagger = \frac{d_A}{m} P_m^A \frac{I^A}{d_A} (P_m^A)^\dagger = \frac{P_m^A (P_m^A)^\dagger}{m} \tag{A11}$$

for all states  $\rho^A$ , i.e., the average map is trace-preserving.

For later convenience, starting from a fixed pure state  $|\Omega^{ABE}\rangle$  purifying  $\rho^{AB}$ , let us define the unnormalized state

$$|\Omega_{m,g}^{A'BE}\rangle := \sqrt{\frac{d_A}{m}}(P_m^A U_g^A \otimes I_B \otimes I_E)|\Omega^{ABE}\rangle.$$

The reduced unnormalized states  $\text{Tr}_E[\Omega_{m,g}^{A'BE}]$  and  $\text{Tr}_B[\Omega_{m,g}^{A'BE}]$  will be correspondingly denoted as  $\omega_{m,g}^{A'B}$  and  $\omega_{m,g}^{A'E}$  (and so on).

By definition, the one-way distillation fidelity  $F_D^{\rightarrow}(\rho^{AB}, m)$  satisfies the bound,

$$\begin{aligned} F_D^{\rightarrow}(\rho^{AB}, m) &\geq F\left(\int_{\mathcal{D}} \max(\text{id}_{A'} \otimes \mathcal{D}_B)(\omega_{m,g}^{A'B}) dg, \Psi_m^{A'B'}\right) \\ &\geq \int dg p(m,g) \max_{\mathcal{D}} F((\text{id}_{A'} \otimes \mathcal{D}_B)(\tilde{\omega}_{m,g}^{A'B}), \Psi_m^{A'B'}), \end{aligned} \tag{A12}$$

where the second line comes from concavity of the fidelity, and  $p(m, g) := \text{Tr } \omega_{m,g}^{A'B}$ . In the last line,  $\tilde{\omega}_{m,g}^{A'B}$  denotes a normalized state. In Eq. (A12),  $|\Psi_m^{A'B'}\rangle$  is any MES of rank  $m$  purifying  $\tau_m^{A'} := \frac{1}{m} P_m^A (P_m^A)^\dagger$ .

Using Lemma 10, we have

$$F_D^{\rightarrow}(\rho^{AB}, m) \geq \int dg p(m,g) F(\tilde{\omega}_{m,g}^{A'E}, \tau_m^{A'} \otimes \tilde{\omega}_{m,g}^E) = \int dg F(\omega_{m,g}^{A'E}, \tau_m^{A'} \otimes \omega_{m,g}^E),$$

where, in the second line, we used the fact that  $F(p\rho, p\sigma) = pF(\rho, \sigma)$ . Further, using Lemma 11, we have that

$$\begin{aligned} F_D^{\rightarrow}(\rho^{AB}, m) &\geq \int dg \frac{\text{Tr } \omega_{m,g}^{A'E} + \text{Tr } \omega_{m,g}^E}{2} - \frac{1}{2} \int dg \|\omega_{m,g}^{A'E} - \tau_m^{A'} \otimes \omega_{m,g}^E\|_1 \\ &= 1 - \frac{1}{2} \int dg \|\omega_{m,g}^{A'E} - \tau_m^{A'} \otimes \omega_{m,g}^E\|_1, \end{aligned}$$

where, in the second line, we used the fact that  $\int \text{Tr } \omega_{m,g}^{A'E} dg = \int \text{Tr } \omega_{m,g}^E dg = 1$ .

Now, for any fixed  $\delta \geq 0$ , let  $\bar{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)$ , where  $\rho^{AE} = \text{Tr}_B[\Omega^{ABE}]$ . Let us, moreover, define  $\bar{\omega}_{m,g}^{A'E} := \frac{d_A}{m} (P_m^A U_g^A \otimes I_E) \bar{\rho}^{AE} (P_m^A U_g^A \otimes I_E)^\dagger$ . By the triangle inequality, we have that

$$\begin{aligned} \|\omega_{m,g}^{A'E} - \tau_m^{A'} \otimes \omega_{m,g}^E\|_1 &\leq \|\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E\|_1 + \|\omega_{m,g}^{A'E} - \bar{\omega}_{m,g}^{A'E}\|_1 + \|\tau_m^{A'} \otimes \bar{\omega}_{m,g}^E - \tau_m^{A'} \otimes \omega_{m,g}^E\|_1 \\ &\leq \|\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E\|_1 + \|\omega_{m,g}^{A'E} - \bar{\omega}_{m,g}^{A'E}\|_1 + \|\bar{\omega}_{m,g}^E - \omega_{m,g}^E\|_1. \end{aligned}$$

Since  $\|\bar{\omega}_{m,g}^E - \omega_{m,g}^E\|_1 \leq \|\bar{\omega}_{m,g}^{AE} - \omega_{m,g}^{AE}\|_1$ , we have that

$$F_D^{\rightarrow}(\rho^{AB}, m) \geq 1 - \int dg \|\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E\|_1 - 2 \int dg \|\omega_{m,g}^{A'E} - \bar{\omega}_{m,g}^{A'E}\|_1$$

for any choice of  $\bar{\rho}^{AE}$  in  $\mathfrak{b}(\rho^{AE}; \delta)$ . Now, thanks to Lemma 3.2 of Ref. 7 and Eq. (4), we know that

$$\int dg \|\omega_{m,g}^{A'E} - \bar{\omega}_{m,g}^{A'E}\|_1 \leq \|\bar{\rho}^{AE} - \rho^{AE}\|_1 \leq 2\delta,$$

which leads us to the estimate

$$F_D^{\rightarrow}(\rho^{AB}, m) \geq 1 - 4\delta - \int dg \|\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E\|_1.$$

Hence, we are left with estimating the last group average.

In order to do so, we exploit a technique used by Renner<sup>14</sup> and Berta<sup>18</sup>: by applying Lemma 9, for any given state  $\sigma^E$  invertible on  $\text{supp } \bar{\rho}^E$ , we obtain the estimate

$$\|\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E\|_1^2 \leq m \text{Tr}[(\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E)X_{m,g}^{A'E}] := m\|\tilde{\rho}_{m,g}^{A'E} - \tau_m^{A'} \otimes \tilde{\rho}_{m,g}^E\|_2^2,$$

where

- (1)  $X_{m,g}^{A'E} := (P_m^{A'} \otimes \sigma^E)^{-1/2}(\bar{\omega}_{m,g}^{A'E} - \tau_m^{A'} \otimes \bar{\omega}_{m,g}^E)(P_m^{A'} \otimes \sigma^E)^{-1/2}$ ,
- (2)  $P_m^{A'} = P_m^A (P_m^A)^\dagger = m \tau_m^{A'}$  [see Eq. (A10)],
- (3)  $\|O\|_2 := \sqrt{\text{Tr}[O^\dagger O]}$  denotes the Hilbert–Schmidt norm,
- (4)  $\tilde{\rho}_{m,g}^{A'E} := (P_m^{A'} \otimes \sigma^E)^{-1/4} \bar{\omega}_{m,g}^{A'E} (P_m^{A'} \otimes \sigma^E)^{-1/4}$ , and finally,
- (5)  $\tilde{\rho}_{m,g}^E := \text{Tr}_{A'}[\tilde{\rho}_{m,g}^{A'E}] = (\sigma^E)^{-1/4} \bar{\omega}_{m,g}^E (\sigma^E)^{-1/4}$ .

It is easy to check that

$$\|\tilde{\rho}_{m,g}^{A'E} - \tau_m^{A'} \otimes \tilde{\rho}_{m,g}^E\|_2^2 = \|\tilde{\rho}_{m,g}^{A'E}\|_2^2 - \frac{1}{m} \|\tilde{\rho}_{m,g}^E\|_2^2.$$

Further, using the concavity of the function  $f(x) = \sqrt{x}$ , we have

$$F_D^{\rightarrow}(\rho^{AB}, m) \geq 1 - 4\delta - \sqrt{\left\{ m \int dg \|\tilde{\rho}_{m,g}^{A'E}\|_2^2 - \int dg \|\tilde{\rho}_{m,g}^E\|_2^2 \right\}}. \tag{A13}$$

Standard calculations, similar to those reported in Refs. 19, 7, and 18, lead to

$$\int dg \|\tilde{\rho}_{m,g}^{A'E}\|_2^2 = \frac{d_A d_A - m}{m d_A^2 - 1} \|\tilde{\rho}^E\|_2^2 + \frac{d_A m d_A - 1}{m d_A^2 - 1} \|\tilde{\rho}^{AE}\|_2^2$$

and

$$\int dg \|\tilde{\rho}_{m,g}^E\|_2^2 = \frac{d_A m d_A - 1}{m d_A^2 - 1} \|\tilde{\rho}^E\|_2^2 + \frac{d_A d_A - m}{m d_A^2 - 1} \|\tilde{\rho}^{AE}\|_2^2,$$

where

$$\tilde{\rho}^{AE} := (\mathbf{I}_A \otimes \sigma^E)^{-1/4} \bar{\rho}^{AE} (\mathbf{I}_A \otimes \sigma^E)^{-1/4},$$

and  $\tilde{\rho}^E := \text{Tr}_A[\bar{\rho}^{AE}]$ . By simple manipulations, we arrive at

$$m \int dg \|\tilde{\rho}_{m,g}^{A'E}\|_2^2 - \int dg \|\tilde{\rho}_{m,g}^E\|_2^2 = \frac{d_A^2(m^2 - 1)}{m(d_A^2 - 1)} \left\{ \|\tilde{\rho}^{AE}\|_2^2 - \frac{1}{d_A} \|\tilde{\rho}^E\|_2^2 \right\}.$$

Since  $m \leq d_A$ ,

$$\frac{d_A^2(m^2 - 1)}{m(d_A^2 - 1)} = m \frac{1 - \frac{1}{m^2}}{1 - \frac{1}{d_A^2}} \leq m$$

so that Eq. (A13) can be rewritten as

$$F_D^{\rightarrow}(\rho^{AB}, m) \geq 1 - 4\delta - \sqrt{m \left\{ \|\tilde{\rho}^{RE}\|_2^2 - \frac{1}{d_A} \|\tilde{\rho}^E\|_2^2 \right\}}$$

for any choice of the states  $\bar{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)$  and  $\sigma^E$  invertible on  $\text{supp } \bar{\rho}^E$ .

Now, notice that

$$\|\tilde{\rho}^{AE}\|_2^2 \leq 2^{S_2(\tilde{\omega}^{AE} \| I_A \otimes \sigma^E)}.$$

This inequality easily follows from (A1), i.e.,

$$\text{Tr}[(\omega^{-1/4} \rho \omega^{-1/4})^2] = \text{Tr}[\omega^{-1/2} \rho \omega^{-1/2} \rho] \leq \text{Tr}[\rho^2 \omega^{-1}] = 2^{S_2(\rho \| \omega)}.$$

Moreover, from Lemma 9,  $\|\tilde{\rho}^E\|_2^2 \geq 1$ . Thus,

$$F_D^-(\rho^{AB}, m) \geq 1 - 4\delta - \sqrt{m \left\{ 2^{S_2(\tilde{\rho}^{AE} \| I_A \otimes \sigma^E)} - \frac{1}{d_A} \right\}}$$

for any choice of states  $\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)$  and  $\sigma^E$ , the latter strictly positive on  $\text{supp } \tilde{\rho}^E$ . In order to tighten the bound, we first optimize (i.e., minimize)  $S_2(\tilde{\rho}^{AE} \| I_A \otimes \sigma^E)$  over  $\sigma^E$  for any  $\tilde{\rho}^{AE}$ , obtaining  $I_2^{A \rightarrow E}(\tilde{\rho}^{AE} | E)$ . We further optimize (i.e., minimize)  $I_2^{A \rightarrow E}(\tilde{\rho}^{AE})$  over  $\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)$ , eventually obtaining  $I_{2,\delta}^{A \rightarrow E}(\rho^{AE})$ . ■

**APPENDIX B: LEMMA 12**

*Lemma 12: For any pure state  $\Psi^{ABE}$  of a tripartite system ABE, for any  $\delta > 0$ , we have that*

$$-I_{2,\delta}^{A \rightarrow E}(\rho^{AE}) \geq I_{0,\delta}^{A \rightarrow B}(\rho^{AB}), \tag{B1}$$

where  $\rho^{AE}$  and  $\rho^{AB}$  denote the corresponding reduced states of the subsystems AE and AB, respectively; the smoothed two-coherent information  $I_{2,\delta}^{A \rightarrow E}(\rho^{AE})$  is defined through (20) for  $\alpha=2$ , and  $I_{0,\delta}^{A \rightarrow B}(\rho^{AB})$  is the smoothed zero-coherent information given by (9).

*Proof:* We make use of the fact that for any  $P, Q \geq 0$ ,

$$D_{\max}(P \| Q) \geq S_2(P \| Q), \tag{B2}$$

where  $S_2(P \| Q)$  is the relative Rényi entropy of order 2, and  $D_{\max}(P \| Q)$  is the max-relative entropy between  $P$  and  $Q$  defined as follows:<sup>15</sup>

$$D_{\max}(P \| Q) := \log \min\{\lambda: P \leq \lambda Q\}. \tag{B3}$$

For any  $\delta \geq 0$ , the smoothed conditional min-entropy is defined as

$$H_{\min}^\delta(\rho^{AE} | E) := \max_{\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} \max_{\sigma^E \in \mathfrak{S}(\mathcal{H}_E)} [-D_{\max}(\tilde{\rho}^{AE} \| I^A \otimes \sigma^E)]. \tag{B4}$$

Moreover,

$$\begin{aligned} H_{\min}^\delta(\rho^{AE} | \rho^E) &:= \max_{\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} H_{\min}(\tilde{\rho}^{AE} | \tilde{\rho}^E) = \max_{\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} [-D_{\max}(\tilde{\rho}^{AE} \| I^A \otimes \tilde{\rho}^E)] \\ &= \max_{\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} \left[ -D_{\max} \left( \tilde{\rho}^{AE} \| I^A \otimes \frac{\tilde{\rho}^E}{\text{Tr } \tilde{\rho}^E} \right) + \log \text{Tr } \tilde{\rho}^E \right] \\ &\leq \max_{\tilde{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} \left[ -D_{\max} \left( \tilde{\rho}^{AE} \| I^A \otimes \frac{\tilde{\rho}^E}{\text{Tr } \tilde{\rho}^E} \right) \right] \leq H_{\min}^\delta(\rho^{AE} | E), \end{aligned} \tag{B5}$$

where in the third equality we used the fact that  $D_{\max}(P \| Q) = D_{\max}(P \| cQ) + \log c$  for any  $c \in \mathbb{R}$  and, in the subsequent inequality, the fact that  $\text{Tr } \tilde{\rho}^E \leq 1$ .

We also need the following duality relation, which was proved in Ref. 18 for two reduced states  $\rho^{AB}$  and  $\rho^{AE}$  of the same tripartite pure state  $\Psi^{ABE}$ , but which can be extended to subnormalized states  $\tilde{\rho}^{AB}$  and  $\tilde{\rho}^{AE}$  coming from  $\tilde{\Psi}^{ABE}$  as well,

$$H_{\min}(\bar{\rho}^{AE}|\bar{\rho}^E) := -D_{\max}(\bar{\rho}^{AE} \| \mathbb{I}_A \otimes \bar{\rho}^E) = H_0(\bar{\rho}^{AB}|B), \quad (\text{B6})$$

where  $H_0(\bar{\rho}^{AB}|B) := -\min_{\sigma^B \in \mathfrak{S}(\mathcal{H}_B)} S_0(\bar{\rho}^{AB} \| \mathbb{I}_A \otimes \sigma^B)$ .

Further (as in the proof of Lemma 3 of Ref. 16), for any  $\delta \geq 0$ , let  $\mathfrak{b}_*(\rho; \delta)$  denote the set of pure states close to a state  $\rho$ , i.e.,

$$\mathfrak{b}_*(\rho; \delta) := \{\psi \in \mathfrak{b}(\rho; \delta) : \text{rank} \psi = 1\} \quad (\text{B7})$$

and let

$$\bar{\mathfrak{b}}(\rho^{AB}; \delta) := \{\text{Tr}_E(\bar{\varphi}^{ABE}) : \bar{\varphi}^{ABE} \in \mathfrak{b}_*(\Psi^{ABE}; \delta)\}, \quad (\text{B8})$$

where  $\Psi^{ABE}$  is any arbitrarily fixed purification of  $\rho^{AB}$ . Hence,  $\bar{\mathfrak{b}}(\rho^{AB}; \delta)$  is the set of states that are  $\delta$ -close to  $\rho^{AB}$  (with respect to the fidelity) on the purified space. It was proved in Ref. 16 that

$$\bar{\mathfrak{b}}(\rho^{AB}; \delta) = \mathfrak{b}(\rho^{AB}; \delta). \quad (\text{B9})$$

This is because, on one hand, the monotonicity of the fidelity under partial trace ensures that  $\bar{\mathfrak{b}}(\rho^{AB}; \delta) \subseteq \mathfrak{b}(\rho^{AB}; \delta)$ . On the other hand, by Uhlmann's theorem,<sup>8</sup> every  $\bar{\rho}^{AB} \in \mathfrak{b}(\rho^{AB}; \delta)$  has a purification  $\bar{\varphi}^{ABE} \in \mathfrak{b}_*(\Psi^{ABE}; \delta)$ , and this implies that  $\mathfrak{b}(\rho^{AB}; \delta) \subseteq \bar{\mathfrak{b}}(\rho^{AB}; \delta)$ .

We now proceed to prove Lemma 12. From definitions (19) and (20) of  $I_{2,\delta}^{A \rightarrow E}(\rho^{AE})$  and the  $\alpha$ -conditional entropies, respectively, we have that

$$\begin{aligned} -I_{2,\delta}^{A \rightarrow E}(\rho^{AE}) &\equiv H_2^\delta(\rho^{AE}|E) = \max_{\bar{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} \max_{\sigma^E \in \mathfrak{S}(\mathcal{H}_E)} [-S_2(\bar{\rho}^{AE} \| \mathbb{I}^A \otimes \sigma^E)] \\ &\geq \max_{\bar{\rho}^{AE} \in \mathfrak{b}(\rho^{AE}; \delta)} \max_{\sigma^E \in \mathfrak{S}(\mathcal{H}_E)} [-D_{\max}(\bar{\rho}^{AE} \| \mathbb{I}^A \otimes \sigma^E)] \equiv H_{\min}^\delta(\rho^{AE}|E) \geq H_{\min}^\delta(\rho^{AE}|\rho^E) \\ &= \max_{\bar{\varphi}^{ABE} \in \mathfrak{b}_*(\Psi^{ABE}; \delta)} H_{\min}(\bar{\rho}^{AE}|\bar{\rho}^E) = \max_{\bar{\varphi}^{ABE} \in \mathfrak{b}_*(\Psi^{ABE}; \delta)} [-H_0(\bar{\rho}^{AB}|B)] = I_{0,\delta}^{A \rightarrow B}(\rho^{AB}), \end{aligned} \quad (\text{B10})$$

where the second inequality follows from (B5), while the first identity on the third line follows from the fact that  $\bar{\mathfrak{b}}(\rho^{AE}; \delta) = \mathfrak{b}(\rho^{AE}; \delta)$ . The subsequent identity follows from (B6), while the last identity follows from (B9) and the definition of the smoothed zero-coherent information (9). ■

<sup>1</sup>G. Bowen and N. Datta, *IEEE Trans. Inf. Theory* **54**, 3677 (2008).

<sup>2</sup>G. Bowen and N. Datta, e-print arXiv:0705.1957.

<sup>3</sup>M. Hayashi, *IEEE Trans. Inf. Theory* **52**, 1904 (2006).

<sup>4</sup>K. Matsumoto, e-print arXiv:0708.3129v1.

<sup>5</sup>F. G. S. L. Brandão and J. Eisert, *J. Math. Phys.* **49**, 042102 (2008).

<sup>6</sup>W. F. Stinespring, *Proc. Am. Math. Soc.* **6**, 211 (1955).

<sup>7</sup>P. Hayden, M. Horodecki, J. Yard, and A. Winter, *Open Syst. Inf. Dyn.* **15**, 7 (2008).

<sup>8</sup>A. Uhlmann, *Rep. Math. Phys.* **9**, 273 (1976).

<sup>9</sup>M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

<sup>10</sup>S. Verdú and T. S. Han, *IEEE Trans. Inf. Theory* **40**, 1147 (1994); T. S. Han, *Information-Spectrum Methods in Information Theory* (Springer-Verlag, Berlin, 2002); T. Ogawa and H. Nagaoka, *IEEE Trans. Inf. Theory* **46**, 2428 (2000); H. Nagaoka and M. Hayashi, *ibid.* **53**, 534 (2007).

<sup>11</sup>M. Ozawa, *J. Math. Phys.* **25**, 79 (1984).

<sup>12</sup>M. Hayashi and H. Nagaoka, *IEEE Trans. Inf. Theory* **49**, 1753 (2003).

<sup>13</sup>G. Bowen and N. Datta, Proceedings of the International Symposium of Information Theory (ISIT), 2006, p. 451; e-print arXiv:quant-ph/0604013.

<sup>14</sup>R. Renner, "Security of quantum key distribution," Ph.D. thesis, ETH Zurich, 2005.

<sup>15</sup>N. Datta, *IEEE Trans. Inf. Theory* **55**, 2816 (2009).

<sup>16</sup>M. Tomamichel, R. Colbeck, and R. Renner, *IEEE Trans. Inf. Theory* **55**, 5840–5847 (2009).

<sup>17</sup>D. Petz, *Rep. Math. Phys.* **23**, 57 (1986).

<sup>18</sup>M. Berta, "Single-shot quantum state merging," Diploma thesis, ETH Zurich, 2008; e-print arXiv:quant-ph/0912.4495v1.

- <sup>19</sup>M. Horodecki, J. Oppenheim, and A. Winter, *Commun. Math. Phys.* **269**, 107 (2007).
- <sup>20</sup>F. Buscemi and N. Datta, *IEEE Trans. Inf. Theory* **56**, 1447 (2010).
- <sup>21</sup>F. Buscemi and N. Datta, e-print arXiv:quant-ph/0906.3698v2.
- <sup>22</sup>F. G. S. L. Brandao and M. B. Plenio, *Commun. Math. Phys.* **295**, 791 (2010).
- <sup>23</sup>I. Devetak and A. Winter, *Proc. R. Soc. London, Ser. A* **461**, 207 (2005).