Non-Asymptotic Entanglement Distillation

Kun Fang[ID], Xin Wang[ID], Marco Tomamichel[ID], *Senior Member, IEEE*, and Runyao Duan

Abstract— **Entanglement distillation, an essential quantum information processing task, refers to the conversion from multiple copies of noisy entangled states to a smaller number of highly entangled states. In this paper, we study the non-asymptotic fundamental limits for entanglement distillation. We investigate the optimal tradeoff between the distillation rate, the number of prepared states, and the error tolerance. First, we derive the one-shot distillable entanglement under completely positive partial transpose preserving operations as a semidefinite program and demonstrate an exact characterization via the quantum hypothesis testing relative entropy. Second, we establish efficiently computable second-order estimations of the distillation rate for general quantum states. In particular, we provide explicit as well as approximate evaluations for various quantum states of practical interest, including pure states, mixture of Bell states, maximally correlated states, and isotropic states.**

Index Terms— **Entanglement distillation, semidefinite programming, finite resources, second-order analysis.**

I. INTRODUCTION

QUANTUM entanglement is a striking feature of quantum mechanics and a key ingredient in many quantum information processing tasks [1], including teleportation [2], superdense coding [3], and quantum cryptography [3]–[5]. All these protocols necessarily rely on entanglement resources. It is thus of great importance to transform less entangled states into more suitable ones such as maximally entangled states Ψ_k.

K. Fang was with the Centre for Quantum Software and Information, University of Technology Sydney, Ultimo, NSW 2007, Australia. He is now with the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, U.K. (e-mail: kf383@cam.ac.uk).

X. Wang was with the Centre for Quantum Software and Information, University of Technology Sydney, Ultimo, NSW 2007, Australia. He is now with the Joint Center for Quantum Information and Computer Science, University of Maryland at College Park, College Park, MD 20742 USA (e-mail: wangxinfelix@gmail.com).

M. Tomamichel is with the Centre for Quantum Software and Information, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: marco.tomamichel@uts.edu.au).

R. Duan is with the Institute for Quantum Computing, Baidu Research, Beijing 100193, China, and also with the Centre for Quantum Software and Information, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: duanrunyao@baidu.com).

Communicated by M. M. Wilde, Associate Editor for Quantum Information Theory.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2019.2914688

This procedure is known as *entanglement distillation* or *entanglement concentration* [6].

The task of entanglement distillation allows two parties (Alice and Bob) to perform a set of free operations Ω, for example, local operations and classical communication (LOCC). The *distillable entanglement* characterizes the optimal rate at which one can asymptotically obtain maximally entangled states from a collection of identically and independently distributed (i.i.d) prepared entangled states [7]–[9]. The concise definition of distillable entanglement $E_{D,\Omega}(\rho_{AB})$ by the class of operation Ω can be given by

$$\sup \left\{ r \ \middle| \ \lim_{n \to \infty} \left(\inf_{\Pi \in \Omega} \|\Pi(\rho_{AB}^{\otimes n}) - \Psi_2^{\otimes r \cdot n}\|_1 \right) = 0 \right\}. \quad (1)$$

Entanglement distillation from non-i.i.d prepared states has also been considered recently in [10]. Distillable entanglement is a fundamental entanglement measure which captures the resource character of quantum entanglement. Up to now, it remains unknown how to compute the distillable entanglement for general quantum states and various approaches [11]–[19] have been developed to evaluate this important quantity. In particular, the Rains bound [15] as well as the squashed entanglement [17] are arguably the best general upper bounds while the hashing bound [20] is the best lower bound for the distillable entanglement.

The conventional approach to studying entanglement distillation is to consider the asymptotic limit (first-order asymptotics), assuming our access to an unbounded number of i.i.d copies of a quantum state. In a realistic setting, however, the resources are finite and the number of i.i.d prepared states is necessarily limited. More importantly, it is very difficult to perform coherent state manipulations over large numbers of systems. Therefore, it becomes crucial to characterize how well we can faithfully distill maximally entangled states from a finite number of copies of the prepared states. Since the first-order asymptotics are insufficient to give a precise estimation when n is finite, it is necessary to consider higher order asymptotics. Specifically, we consider estimating the optimal distillation rate r_n for n copies of the state to the order $\sqrt{n}$, for example, $r_n = a + b/\sqrt{n} + O(\log n/n)$, where a and b are the so-called first and second-order asymptotics respectively. The first-order term a determines the asymptotic rate of r_n while the second-order term b indicates how fast the rate r_n converges to a. The second-order estimation is especially accurate for large blocklength n where the higher-order term $O(\log n/n)$ is negligible. In particular, for practical use, it is desirable to find efficiently computable coefficients a, b.

The study of such non-asymptotic scenarios has recently garnered great interest in classical information theory (e.g., [21]–[23]) as well as in quantum information theory (e.g., [24]–[36]). Here we study the setting of entanglement distillation. A non-asymptotic analysis of entanglement distillation will help us better exploit the power of entanglement in a realistic setting. Previously, the one-shot distillable entanglement was studied in [37] and [38]. But these bounds are not known to be computable in general, which makes it difficult to apply them as experimental benchmarks. These one-shot bounds are not suitable to establish second-order estimations either. Datta and Leditzky studied the second-order estimation of distillable entanglement under LOCC operations for pure states [39]. Here, we go beyond their results by considering more general operations and states.

The remainder of this paper is organized as follows. In Section III we study the one-shot entanglement distillation under completely positive partial transpose preserving operations and find that the one-shot rate is efficiently computable via a semidefinite program (SDP). Based on this SDP, we present an exact characterization of the one-shot rate via the quantum hypothesis testing relative entropy, which can be seen as a one-shot analog of the Rains bound. In Section IV we investigate the entanglement distillation for *n*-fold tensor product states and provide second-order estimations for general quantum states. In Section V we apply our second-order estimations to various quantum states of practical interest, including pure states, mixture of Bell states, maximally correlated states and isotropic states.

II. PRELIMINARIES

In the following, we will frequently use symbols such as A (or A') and B (or B') to denote finite-dimensional Hilbert spaces associated with Alice and Bob, respectively. A quantum state on system A is a positive operator ρ_A with unit trace. The set of quantum states is denoted as $S(A) := \{\rho_A \geq 0 \,|\, \mathrm{Tr}\, \rho_A = 1\}$. The set of subnormalized quantum states is denoted as $S_{\leq}(A) := \{\rho_A \geq 0 \,|\, 0 < \mathrm{Tr}\, \rho_A \leq 1\}$. We call a positive operator separable if it can be written as a convex combination of tensor product positive operators.

A quantum operation is characterized by a completely positive and trace-preserving (CPTP) linear map. There are several different classes of quantum operations we often use. We call a bipartite quantum operation LOCC if it can be realized by local operations and classical communication. If only one-way classical communication is allowed, say, classical information from Alice to Bob, we call it 1-LOCC. While LOCC, including 1-LOCC, emerges as the natural class of operations in many important quantum information tasks, its mathematical structure is complex and difficult to characterize [40]. Therefore we may consider larger but mathematically more tractable classes of operations. The operations most frequently employed beyond LOCC are the so-called completely positive partial transpose preserving (PPT) operations and separable (SEP) operations. A bipartite quantum operation $\Pi_{AB \to A'B'}$ is said to be a PPT (or SEP) operation if its Choi-Jamiołkowski matrix $J_\Pi = \sum_{i,j,m,k} |i_A j_B\rangle\langle m_A k_B| \otimes \Pi(|i_A j_B\rangle\langle m_A k_B|)$ is

positive under partial transpose (or separable) across the bipartition of $AA' : BB'$, where $\{|i_A\rangle\}$ and $\{|j_B\rangle\}$ are orthonormal bases on Hilbert spaces A and B, respectively. In particular, PPT operations can be characterized via semidefinite conditions [15]. The entanglement theory under PPT operations has been studied in the literature (e.g., [41]–[44]) and offers the limitations of LOCC.

A well known fact is that the classes of above introduced operations obey the strict inclusions [1], [45], 1-LOCC $\subsetneq$ LOCC $\subsetneq$ SEP $\subsetneq$ PPT. As a consequence, for any quantum state ρ_{AB} we have the following chain of inequalities,

$$E_{\mathrm{D,1\text{-}LOCC}}(\rho) \leq E_{\mathrm{D,LOCC}}(\rho) \leq E_{\mathrm{D,SEP}}(\rho) \leq E_{\mathrm{D,PPT}}(\rho). \quad (2)$$

This allows us to use $E_{\mathrm{D,1\text{-}LOCC}}$ and $E_{\mathrm{D,PPT}}$ as lower and upper bounds, respectively, for entanglement distillation under LOCC operations.

Quantum hypothesis testing is the task of distinguishing two possible states of a system, ρ_0 and ρ_1. Two hypotheses are studied: the null hypothesis H_0 is that the state is ρ_0; the alternative hypothesis H_1 is that the state is ρ_1. We are allowed to perform a measurement presented by the POVM $\{M, \mathbb{1} - M\}$ with corresponding classical outcomes 0 and 1. If the outcome is 0, we accept the null hypothesis. Otherwise, we accept the alternative one. The probabilities of *type-I* and *type-II* error are given by $\mathrm{Tr}(\mathbb{1} - M)\rho_0$ and $\mathrm{Tr}\, M\rho_1$, respectively. Quantum hypothesis testing relative entropy considers minimizing the type-II error while keeping type-I error within a given error tolerance. Specifically, it is defined as

$$D_H^\varepsilon(\rho_0\|\rho_1) := -\log\min$$
$$\left\{ \mathrm{Tr}\, M\rho_1 \,\big|\, 0 \leq M \leq \mathbb{1}, \ 1 - \mathrm{Tr}\, M\rho_0 \leq \varepsilon \right\}. \quad (3)$$

Throughout the paper we take the logarithm to be base two unless stated otherwise. Note that D_H^ε is a fundamental quantity in quantum theory [46]–[49] and can be solved by SDP—a powerful tool in quantum information theory with a plethora of applications (e.g., [50]–[55]).

For the convenience of the following discussion, we consider an extension of the quantum hypothesis testing relative entropy where the second argument ρ_1 is only restricted to be Hermitian (and not necessarily a positive semi-definite operator). We will also use the convention that $\log x = -\infty$ for $x \leq 0$ in case the optimal value $\mathrm{Tr}\, M\rho_1 \leq 0$. Then Eq. (3) is still a well-defined SDP. This extension is essential to obtain tight characterizations in this work as well as a related work on coherence distillation [56].

III. ONE-SHOT ENTANGLEMENT DISTILLATION

In this section, we consider distilling a maximally entangled state from a single copy of the resource state and study the tradeoff between the one-shot distillation rate and the fidelity of distillation. Since the distillation process cannot always be accomplished perfectly, we use the fidelity of distillation to characterize the performance of a given distillation task. Then the one-shot distillable entanglement is defined as the logarithm of the maximal dimension of the maximally entangled state that we can obtain while keeping the infidelity of the distillation process within a given tolerance.

Definition 1 *For any bipartite quantum state ρ_{AB}, the fidelity of distillation under the operation class Ω is defined as* [15]

$$F_\Omega(\rho_{AB}, k) := \max_{\Pi \in \Omega} \operatorname{Tr} \Pi_{AB \to A'B'}(\rho_{AB})\Psi_k, \qquad (4)$$

where $\Psi_k := (1/k)\sum_{i,j=0}^{k-1} |ii\rangle\langle jj|$ is the k-dimensional maximally entangled state and the maximization is taken over all possible operations Π in the set Ω.

Definition 2 *For any bipartite quantum state ρ_{AB}, the one-shot ε-error distillable entanglement under the operation class Ω is defined as*

$$E_{\mathrm{D},\Omega}^{(1),\varepsilon}(\rho_{AB}) := \log \max \left\{ k \in \mathbb{N} \mid F_\Omega(\rho_{AB}, k) \ge 1 - \varepsilon \right\}. \quad (5)$$

The asymptotic distillable entanglement is then given by the regularization:

$$E_{\mathrm{D},\Omega}(\rho_{AB}) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} E_{\mathrm{D},\Omega}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}). \qquad (6)$$

Due to the linear characterization of PPT operations [15], the one-shot distillable entanglement under PPT operations can be easily computed via the following optimization. The main ingredients of the proof use the symmetry of the maximally entangled state and the spectral decomposition of the swap operator, which are standard techniques used in literatures, e.g., [15], [41], [43]. We present the detailed proof here for the sake of completeness.

Proposition 3 *For any bipartite quantum state ρ_{AB} and error tolerance ε, the one-shot distillable entanglement under PPT operations is given by*

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = \log \max \lfloor 1/\eta \rfloor \quad \text{s.t.}$$
$$0 \le M_{AB} \le \mathbb{1}_{AB},$$
$$\operatorname{Tr} \rho_{AB} M_{AB} \ge 1 - \varepsilon,$$
$$-\eta \mathbb{1}_{AB} \le M_{AB}^{\mathsf{T}_B} \le \eta \mathbb{1}_{AB}. \qquad (7)$$

Proof: From the definition of the one-shot distillable entanglement, we have

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = \log \max \, k \in \mathbb{N} \quad \text{s.t.}$$
$$\operatorname{Tr} \Pi_{AB \to A'B}(\rho_{AB})\Psi_k \ge 1 - \varepsilon, \ \Pi \in \mathrm{PPT}. \quad (8)$$

According to the Choi-Jamiołkowski representation of quantum operations [57], [58], we can represent the output state of operation $\Pi_{AB \to A'B'}$ via its Choi-Jamiołkowski matrix J_Π as

$$\Pi_{AB \to A'B'}(\rho_{AB}) = \operatorname{Tr}_{AB}\left[J_\Pi \cdot \rho_{AB}^{\mathsf{T}} \otimes \mathbb{1}_{A'B'} \right]. \qquad (9)$$

By straightforward calculations, we have

$$\operatorname{Tr} \Pi_{AB \to A'B'}(\rho_{AB})\Psi_k$$
$$= \operatorname{Tr}\left[\operatorname{Tr}_{AB}(J_\Pi \cdot \rho_{AB}^{\mathsf{T}} \otimes \mathbb{1}_{A'B'}) \right]\Psi_k \qquad (10)$$
$$= \operatorname{Tr} J_\Pi \cdot (\rho_{AB}^{\mathsf{T}} \otimes \mathbb{1}_{A'B'})(\mathbb{1}_{AB} \otimes \Psi_k) \qquad (11)$$
$$= \operatorname{Tr} J_\Pi \cdot (\mathbb{1}_{AB} \otimes \Psi_k)(\rho_{AB}^{\mathsf{T}} \otimes \mathbb{1}_{A'B'}) \qquad (12)$$
$$= \operatorname{Tr}\left[\operatorname{Tr}_{A'B'} J_\Pi \cdot (\mathbb{1}_{AB} \otimes \Psi_k) \right]\rho_{AB}^{\mathsf{T}}. \qquad (13)$$

Recall that Π is a PPT operation if and only if its Choi-Jamiołkowski matrix J_Π satisfies [15]

$$J_\Pi \ge 0, \quad \operatorname{Tr}_{A'B'} J_\Pi = \mathbb{1}_{AB}, \quad J_\Pi^{\mathsf{T}_{BB'}} \ge 0. \qquad (14)$$

Combining Eqs. (8), (13), (14) we have the optimization

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = \log \max \, k \in \mathbb{N} \quad \text{s.t.}$$
$$\operatorname{Tr}\left[\operatorname{Tr}_{A'B'} J_\Pi \cdot (\mathbb{1}_{AB} \otimes \Psi_k) \right]\rho_{AB}^{\mathsf{T}} \ge 1 - \varepsilon, \qquad (15)$$
$$J_\Pi \ge 0, \ \operatorname{Tr}_{A'B'} J_\Pi = \mathbb{1}_{AB}, \ J_\Pi^{\mathsf{T}_{BB'}} \ge 0.$$

Suppose one optimal solution in optimization (15) is given by $\widetilde{J}_\Pi$. Since Ψ_k is invariant under any local unitary $U_{A'} \otimes \overline{U}_{B'}$, i.e., $(U_{A'} \otimes \overline{U}_{B'})\Psi_k(U_{A'} \otimes \overline{U}_{B'})^\dagger = \Psi_k$, we can verify that $(U_{A'} \otimes \overline{U}_{B'})\widetilde{J}_\Pi(U_{A'} \otimes \overline{U}_{B'})^\dagger$ is also optimal. Since any convex combination of optimal solutions remains optimal, we know that

$$\int dU (U_{A'} \otimes \overline{U}_{B'})\widetilde{J}_\Pi(U_{A'} \otimes \overline{U}_{B'})^\dagger \qquad (16)$$

is optimal, where dU is the Haar measure. According to Schur's lemma, the result of the above integral gives an optimal solution admitting the structure of $W_{AB} \otimes \Psi_k + Q_{AB} \otimes (\mathbb{1} - \Psi_k)$ with certain linear operators W_{AB} and Q_{AB}. Thus without loss of generality, we can restrict our consideration of the optimal Choi-Jamiołkowski matrix in (15) as

$$J_\Pi = W_{AB} \otimes \Psi_k + Q_{AB} \otimes (\mathbb{1} - \Psi_k). \qquad (17)$$

In the following we take Eq. (17) into the optimization (15) and perform some further simplifications. Denote P_+ and P_- as the symmetric and anti-symmetric projections respectively. From the spectral decomposition, we know that $\Psi_k^{\mathsf{T}_{B'}} = (P_+ - P_-)/k$ and

$$J_\Pi^{\mathsf{T}_{BB'}} = W_{AB}^{\mathsf{T}_B} \otimes \Psi_k^{\mathsf{T}_{B'}} + Q_{AB}^{\mathsf{T}_B} \otimes (\mathbb{1} - \Psi_k)^{\mathsf{T}_{B'}} \qquad (18)$$
$$= W_{AB}^{\mathsf{T}_B} \otimes (P_+ - P_-)/k$$
$$\quad + Q_{AB}^{\mathsf{T}_B} \otimes [(k-1)P_+ + (k+1)P_-]/k \qquad (19)$$
$$= \left[W_{AB}^{\mathsf{T}_B} + (k-1)Q_{AB}^{\mathsf{T}_B} \right] \otimes P_+/k$$
$$\quad + \left[-W_{AB}^{\mathsf{T}_B} + (k+1)Q_{AB}^{\mathsf{T}_B} \right] \otimes P_-/k. \qquad (20)$$

Since P_+ and P_- are positive and orthogonal to each other, we have $J_\Pi^{\mathsf{T}_{BB'}} \ge 0$ if and only if

$$W_{AB}^{\mathsf{T}_B} + (k-1)Q_{AB}^{\mathsf{T}_B} \ge 0, \qquad (21)$$
$$-W_{AB}^{\mathsf{T}_B} + (k+1)Q_{AB}^{\mathsf{T}_B} \ge 0. \qquad (22)$$

Note that $\operatorname{Tr}\left[\operatorname{Tr}_{A'B'} J_\Pi \cdot (\mathbb{1}_{AB} \otimes \Psi_k) \right]\rho_{AB}^{\mathsf{T}} = \operatorname{Tr} W_{AB}\rho_{AB}^{\mathsf{T}}$. We can simplify the optimization (15) as

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho) = \log \max \, k \in \mathbb{N} \quad \text{s.t.}$$
$$\operatorname{Tr} W_{AB}\rho_{AB}^{\mathsf{T}} \ge 1 - \varepsilon, \ W_{AB} \ge 0,$$
$$W_{AB} + (k^2 - 1)Q_{AB} = \mathbb{1}_{AB}, \ Q_{AB} \ge 0 \qquad (23)$$
$$(1-k)Q_{AB}^{\mathsf{T}_B} \le W_{AB}^{\mathsf{T}_B} \le (1+k)Q_{AB}^{\mathsf{T}_B}.$$

Eliminating the variable Q_{AB} via the condition $W_{AB} + (k^2 - 1)Q_{AB} = \mathbb{1}_{AB}$ and taking $M_{AB} = W_{AB}^{\mathsf{T}}$, $\eta = 1/k$, we obtain the desired result. ∎

Remark The optimization in the above proposition is not exactly an SDP due to the nonlinear objective function. However, to compute the one-shot distillable entanglement, we can first implement the SDP:

$$\min \ \eta \quad \text{s.t.}$$
$$0 \le M_{AB} \le \mathbb{1}_{AB},$$
$$\mathrm{Tr}\, \rho_{AB} M_{AB} \ge 1 - \varepsilon, \tag{24}$$
$$-\eta \mathbb{1}_{AB} \le M_{AB}^{\mathsf{T}_B} \le \eta \mathbb{1}_{AB}.$$

and obtain the optimal value η_0. Then the one-shot distillable entanglement is given by $\log\lfloor 1/\eta_0 \rfloor$. Since the second step is trivial, we will also call the Eq. (3) an SDP characterization.

We also note that it is possible to use the SDP of distillation fidelity in [15] to obtain Eq. (24). However, for a general task, the SDP of its optimal fidelity is not sufficient to ensure that the rate can also be characterized by an SDP. A counter-example can be given by the one-shot PPT-assisted entanglement dilution task [43].

The one-shot entanglement distillation under non-entangling operations was studied by Brandão and Datta [38]. Particularly they provided both lower and upper (non-matching) bounds of the one-shot distillable rate via the quantum hypothesis testing relative entropy. Based on the SDP formula in Proposition 3, we are now ready to give a similar but exact characterization for distillation under PPT operations.

Theorem 4 *For any bipartite quantum state ρ_{AB} and the error tolerance $\varepsilon \in (0, 1)$, it holds*

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = \min_{\substack{\|G^{\mathsf{T}_B}\|_1 \le 1 \\ G = G^\dagger}} D_H^\varepsilon(\rho_{AB} \| G_{AB}) - \delta, \tag{25}$$

where $\delta \in [0, 1]$ is the least constant such that the r.h.s. is the logarithm of an integer.

Proof: We first note the fact that $\delta = \log x - \log\lfloor x \rfloor \in [0, 1]$ for any $x \ge 1$. Thus we can use the least constant $\delta \in [0, 1]$ to adjust the r.h.s. of Eq. (3) to be the logarithm of an integer without taking the floor function, i.e.,

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = -\log \big[\min \ \eta \big] - \delta \quad \text{s.t.}$$
$$0 \le M_{AB} \le \mathbb{1}_{AB},$$
$$\mathrm{Tr}\, \rho_{AB} M_{AB} \ge 1 - \varepsilon,$$
$$-\eta \mathbb{1}_{AB} \le M_{AB}^{\mathsf{T}_B} \le \eta \mathbb{1}_{AB}. \tag{26}$$

The main ingredient of the following proof is the norm duality between the trace norm and the operator norm. Denote the set $\mathcal{S}_\rho := \{ M_{AB} \,|\, 0 \le M_{AB} \le \mathbb{1}_{AB}, \ \mathrm{Tr}\, \rho_{AB} M_{AB} \ge 1 - \varepsilon \}$. Then we have

$$E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}) = -\log \min_{M \in \mathcal{S}_\rho} \|M_{AB}^{\mathsf{T}_B}\|_\infty - \delta \tag{27}$$

$$= -\log \min_{M \in \mathcal{S}_\rho} \max_{\substack{\|G\|_1 \le 1 \\ G = G^\dagger}} \mathrm{Tr}\, M_{AB}^{\mathsf{T}_B} G_{AB} - \delta \tag{28}$$

$$= -\log \max_{\substack{\|G\|_1 \le 1 \\ G = G^\dagger}} \min_{M \in \mathcal{S}_\rho} \mathrm{Tr}\, M_{AB}^{\mathsf{T}_B} G_{AB} - \delta \tag{29}$$

$$= -\log \max_{\substack{\|G^{\mathsf{T}_B}\|_1 \le 1 \\ G = G^\dagger}} \min_{M \in \mathcal{S}_\rho} \mathrm{Tr}\, M_{AB} G_{AB} - \delta \tag{30}$$

$$= \min_{\substack{\|G^{\mathsf{T}_B}\|_1 \le 1 \\ G = G^\dagger}} -\log \min_{M \in \mathcal{S}_\rho} \mathrm{Tr}\, M_{AB} G_{AB} - \delta \tag{31}$$

$$= \min_{\substack{\|G^{\mathsf{T}_B}\|_1 \le 1 \\ G = G^\dagger}} D_H^\varepsilon(\rho_{AB} \| G_{AB}) - \delta. \tag{32}$$

The first line follows from Eq. (III) and $\|X\|_\infty = \min\{ \eta \,|\, -\eta\mathbb{1} \le X \le \eta\mathbb{1} \}$. The second line follows from the norm duality between the trace norm and the operator norm, i.e., $\|X\|_\infty = \max_{\|Y\|_1 \le 1, Y = Y^\dagger} \mathrm{Tr}\, XY$. The third line uses the Sion minimax theorem [59] to swap the minimization with the maximization. In the fourth line, we replace G_{AB} with $G_{AB}^{\mathsf{T}_B}$. The last line follows by definition. ∎

Compared to the result by Brandão and Datta [38], Theorem 4 gives an exact and complete characterization for one-shot entanglement distillation under PPT operations. The proof technique of this theorem was later applied in coherence theory, where the one-shot distillable coherence under maximally incoherent operations is also completely characterized by the quantum hypothesis testing relative entropy [56]. It is also worth noting that the second argument G_{AB} in Eq. (25) is not necessarily positive, with an explicit example presented in Appendix A. Thus the extension to general Hermitian operators becomes crucial to obtain the exact characterization instead of non-matching bounds.

In term of the asymptotic distillable entanglement, the best known upper bound was given by the Rains bound [15], [60]. That is, $E_{\mathrm{D,PPT}}(\rho_{AB}) \le R(\rho_{AB})$ with

$$R(\rho_{AB}) = \min_{\substack{\|\sigma^{\mathsf{T}_B}\|_1 \le 1 \\ \sigma_{AB} \ge 0}} D(\rho_{AB} \| \sigma_{AB}), \tag{33}$$

where the quantum relative entropy is given by $D(\rho \| \sigma) := \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$ if $\mathrm{supp}\, \rho \subseteq \mathrm{supp}\, \sigma$ and $+\infty$ otherwise. Theorem 4 can be seen as a one-shot analog of this result, since we can quickly recover the Rains bound through the quantum Stein's lemma [47].

Corollary 5 *For any bipartite quantum state ρ_{AB}, it holds $E_{\mathrm{D,PPT}}(\rho_{AB}) \le R(\rho_{AB})$.*

Proof: Denote a minimizer of the Rains bound as σ_{AB} and then $R(\rho_{AB}) = D(\rho_{AB} \| \sigma_{AB})$. According to Theorem 4, we have

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}) \le \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} D_H^\varepsilon(\rho_{AB}^{\otimes n} \| \sigma_{AB}^{\otimes n}),$$

since $\sigma_{AB}^{\otimes n}$ is a feasible solution. The l.h.s. gives $E_{\mathrm{D,PPT}}(\rho_{AB})$ by definition while the r.h.s. converges to $D(\rho_{AB} \| \sigma_{AB})$ due to the quantum Stein's lemma [47]. ∎

IV. Non-Asymptotic Entanglement Distillation

In this section, we study the estimation of distillable entanglement for given n copies of the resource state. We provide both lower and upper bounds that are efficiently computable for general quantum states.

Before proceeding, we need to introduce some basic notations. The purified distance between two subnormalized quantum states is defined as $P(\rho, \sigma) := \sqrt{1 - F^2(\rho, \sigma)}$ with the generalized fidelity [32]

$$F(\rho, \sigma) := \|\rho^{1/2}\sigma^{1/2}\|_1 + \sqrt{(1 - \mathrm{Tr}\,\rho)(1 - \mathrm{Tr}\,\sigma)}. \quad (34)$$

Denote the ε-ball around ρ_{AB} as

$$\mathcal{B}_\varepsilon(\rho_{AB}) := \{\,\widetilde{\rho}_{AB} \in \mathcal{S}_\le(AB) \mid P(\rho_{AB}, \widetilde{\rho}_{AB}) \le \varepsilon\,\}. \quad (35)$$

The smooth conditional max-entropy is defined as

$$H_{\max}^\varepsilon(A|B)_\rho := \inf_{\widetilde{\rho}_{AB} \in \mathcal{B}_\varepsilon(\rho_{AB})} \sup_{\sigma_B \in \mathcal{S}(B)} \log F(\widetilde{\rho}_{AB}, \mathbb{1}_A \otimes \sigma_B). \quad (36)$$

The following lemma gives the second-order expansion of the quantum hypothesis testing relative entropy and the smooth conditional max-entropy. This lemma is crucial to obtain our second-order bounds.

Lemma 6 *For any quantum states ρ_{AB} and positive operator σ_{AB}, it holds [27], [61]*

$$D_H^\varepsilon\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = nD(\rho\|\sigma) \quad (37)$$
$$+ \sqrt{nV(\rho\|\sigma)}\,\Phi^{-1}(\varepsilon) + O(\log n),$$

$$H_{\max}^\varepsilon(A^n|B^n)_{\rho^{\otimes n}} = -nI(A\rangle B)_\rho \quad (38)$$
$$- \sqrt{nV(A\rangle B)_\rho}\,\Phi^{-1}(\varepsilon^2) + O(\log n),$$

where $V(\rho\|\sigma) := \mathrm{Tr}\,\rho(\log\rho - \log\sigma)^2 - D(\rho\|\sigma)^2$ is the information variance, $I(A\rangle B)_\rho := D(\rho_{AB}\|\mathbb{1}_A \otimes \rho_B)$ is the coherent information, $V(A\rangle B)_\rho := V(\rho_{AB}\|\mathbb{1}_A \otimes \rho_B)$ is the coherent information variance and Φ^{-1} is the inverse of cumulative normal distribution function.

Theorem 7 *For any bipartite quantum state ρ_{AB}, the number of prepared states n, the error tolerance $\varepsilon \in (0, 1)$, and the operation class $\Omega \in \{1\text{-LOCC}, \text{LOCC}, \text{SEP}, \text{PPT}\}$, it holds*

$$f(\rho, n, \varepsilon) + O(\log n) \le E_{D,\Omega}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}) \le g(\rho, n, \varepsilon) + O(\log n), \quad (39)$$

where $f(\rho, n, \varepsilon)$ and $g(\rho, n, \varepsilon)$ are efficiently computable functions given by

$$f(\rho, n, \varepsilon) = nI(A\rangle B)_\rho + \sqrt{nV(A\rangle B)_\rho}\,\Phi^{-1}(\varepsilon), \quad (40)$$
$$g(\rho, n, \varepsilon) = nR(\rho_{AB}) + \sqrt{nV(\rho_{AB}\|\sigma_{AB})}\,\Phi^{-1}(\varepsilon), \quad (41)$$

and σ_{AB} is any minimizer of the Rains bound, Φ^{-1} is the inverse of cumulative normal distribution function.

Proof: Due to the inclusion relations of the operation classes, we only need to show the upper bound for PPT operations and lower bound for 1-LOCC operations. Each second-order bound can be obtained by applying the corresponding one-shot bound to the n-fold tensor product state $\rho^{\otimes n}$ and using the second-order expansion of the related entropies.

For the second-order upper bound, the proof steps are similar to Corollary 5. Denote σ_{AB} as a minimizer of the Rains bound. We first have $E_{D,\text{PPT}}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}) \le D_H^\varepsilon(\rho_{AB}^{\otimes n}\|\sigma_{AB}^{\otimes n}) - \delta$ by taking a feasible solution $\sigma_{AB}^{\otimes n}$ in Theorem 4 and $\delta \in [0, 1]$. Instead of applying the quantum Stein's lemma here, we use

the second-order expansion of the quantum hypothesis testing relative entropy in Lemma 6 and obtain

$$E_{D,\text{PPT}}^{(1),\varepsilon}\left(\rho_{AB}^{\otimes n}\right) \le nD(\rho_{AB}\|\sigma_{AB}) \quad (42)$$
$$+ \sqrt{nV(\rho_{AB}\|\sigma_{AB})}\,\Phi^{-1}(\varepsilon) + O(\log n).$$

For the second-order lower bound, we adopt the one-shot hashing bound [62] that

$$E_{D,1\text{-LOCC}}^{(1),\varepsilon}(\rho_{AB}) \ge -H_{\max}^{\sqrt{\varepsilon}-\eta}(A|B)_\rho + 4\log\eta, \quad (43)$$

with $\eta \in [0, \sqrt{\varepsilon})$. For the state $\rho_{AB}^{\otimes n}$, we choose $\eta = 1/\sqrt{n}$ and have the following result which holds for $n > 1/\varepsilon$,

$$E_{D,1\text{-LOCC}}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}) \ge -H_{\max}^{\sqrt{\varepsilon}-1/\sqrt{n}}(A^n|B^n)_{\rho^{\otimes n}} + 4\log(1/\sqrt{n}). \quad (44)$$

Using the second-order expansion of the smooth conditional max-entropy in Lemma 6, we have

$$E_{D,1\text{-LOCC}}^{(1),\varepsilon}(\rho_{AB}^{\otimes n}) \ge nI(A\rangle B)_\rho$$
$$+ \sqrt{nV(A\rangle B)_\rho}\,\Phi^{-1}\left((\sqrt{\varepsilon} - 1/\sqrt{n})^2\right) + O(\log n). \quad (45)$$

Note that Φ^{-1} is continuously differentiable around $\varepsilon > 0$. Thus $\Phi^{-1}\left((\sqrt{\varepsilon} - 1/\sqrt{n})^2\right) = \Phi^{-1}(\varepsilon) + O(1/\sqrt{n})$ and we have the desired result. ■

Note that the second-order upper bound works for any minimizer σ_{AB} of the Rains bound regardless of its uniqueness. Thus we can choose the one that gives the tightest result. Due to the higher-order term $O(\log n)$, the estimations in Theorem 7 will work better for large blocklength n where the logarithmic term is negligible.

Since the Rains bound in Eq. (33) is given by convex optimization, there are various methods to solve it numerically. We provide an algorithm in Appendix B which can be used to efficiently compute the Rains bound and output the minimizer operator. It is worth noting that our bounds are similar to the second-order bounds on quantum capacity in [33]. But those bounds in [33] are not easy to compute in general.

The difficulty to obtain good second-order estimations is to find suitable one-shot lower and upper bounds which lead to the same ε dependence in the term $\Phi^{-1}(\varepsilon)$ after the second-order expansion. This is necessary to show the tightness of the second-order estimation for specific states. Our result in Theorem 4 and the one-shot lower bound in [62] coordinate well in this sense. There are other one-shot lower bounds [37], [63] which can be used to establish a second-order estimation. But they do not provide matching ε dependence with our second-order upper bound. For certain pure states, there exists a better one-shot lower bound in [63]. But note that our bounds are already tight for general pure states up to the second-order terms (see Proposition 8).

V. EXAMPLES

In this section, we apply our second-order bounds to estimate the non-asymptotic distillable entanglement of some important classes of states, including pure states, mixtures of Bell states, maximally correlated states and isotropic states.

Proposition 8 (Pure states) *For any bipartite quantum pure state ψ_{AB}, the number of prepared states n, the error tolerance $\varepsilon \in (0, 1)$, and the operation class $\Omega \in \{1\text{-LOCC, LOCC, SEP, PPT}\}$, it holds*

$$E_{\mathrm{D},\Omega}^{(1),\varepsilon}(\psi_{AB}^{\otimes n}) = nS(\rho_A)$$
$$+ \sqrt{n(\mathrm{Tr}\,\rho_A(\log\rho_A)^2 - S(\rho_A)^2)}\,\Phi^{-1}(\varepsilon) + O(\log n), \quad (46)$$

where $S(\rho) := -\mathrm{Tr}\,\rho\log\rho$ is the von Neumann entropy and $\rho_A = \mathrm{Tr}_B\,\psi_{AB}$ is the reduced state.

Proof: Since all the quantities concerned are invariant under local unitaries, we only need to consider a pure state ψ_{AB} with the Schmidt decomposition $|\psi_{AB}\rangle = \sum_i \sqrt{p_i}|i_A i_B\rangle$. Then $\rho_A = \sum_i p_i|i_A\rangle\langle i_A|$ and $\rho_B := \mathrm{Tr}_A\,\psi_{AB} = \sum_i p_i|i_B\rangle\langle i_B|$. Let $\sigma_{AB} = \sum_i p_i|i_A i_B\rangle\langle i_A i_B|$. The following equalities are straightforward by calculation,

$$D(\psi_{AB}\|\sigma_{AB}) = I(A\rangle B)_\psi = S(\rho_A), \qquad (47)$$
$$V(\psi_{AB}\|\sigma_{AB}) = V(A\rangle B)_\psi = \mathrm{Tr}\,\rho_A(\log\rho_A)^2 - S(\rho_A)^2.$$

We can first check that σ_{AB} is a feasible solution for the Rains bound and thus $R(\psi_{AB}) \leq D(\psi_{AB}\|\sigma_{AB})$. Note that $I(A\rangle B)_\rho \leq R(\rho_{AB})$ holds for any quantum state ρ_{AB}. Thus we have $D(\psi_{AB}\|\sigma_{AB}) = I(A\rangle B)_\rho \leq R(\rho_{AB}) \leq D(\psi_{AB}\|\sigma_{AB})$, which implies that σ_{AB} is a minimizer of the Rains bound. Finally applying Theorem 7, we have the desired result. ∎

The second-order estimation for pure states has been given by Datta and Leditzky [39], which was only for LOCC operations. It is known that the asymptotic distillable entanglement of a pure state coincides with the von Neumann entropy of its reduced state under 1-LOCC, LOCC, SEP or PPT operations [64]. Proposition 8 shows that not only are the asymptotic distillable entanglement (first-order asymptotics) the same under these four sets of operations but also their convergence speeds (second-order asymptotics).

In laboratories, we usually obtain mixed states due to the imperfection of operations and decoherence. A common case is a noise dominated by one type of Pauli error [65], [66]. Without loss of generality, we consider the phase noise, which results in the mixture of Bell states

$$\rho_{\mathrm{Bell}} = p|v_1\rangle\langle v_1| + (1-p)|v_2\rangle\langle v_2|, \quad 0 < p < 1, \quad (48)$$

where $|v_1\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ and $|v_2\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Let $\sigma_{AB} = (|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|)/2$. Following similar proof steps as Proposition 8, we can check that σ_{AB} is a minimizer of the Rains bound for ρ_{Bell}. Due to Theorem 7, we have the following result.

Proposition 9 (Mixture of Bell states) *For given quantum state ρ_{Bell}, the number of prepared states n, the error tolerance $\varepsilon \in (0, 1)$, and the operation class $\Omega \in \{1\text{-LOCC, LOCC, SEP, PPT}\}$, it holds*

$$E_{\mathrm{D},\Omega}^{(1),\varepsilon}(\rho_{\mathrm{Bell}}^{\otimes n}) = n(1 - h_2(p))$$
$$+ \sqrt{np(1-p)\left(\log\frac{1-p}{p}\right)^2}\,\Phi^{-1}(\varepsilon) + O(\log n), \quad (49)$$

where $h_2(p) := -p\log p - (1-p)\log(1-p)$ is the binary entropy.

Besides the mixture of Bell states presented above, we can also show the tightness of our second-order bounds for a broader class of states, i.e., maximally correlated states

$$\rho_{\mathrm{mc}} = \sum_{i,j=0}^{d-1} \widetilde{\rho}_{ij}|i_A i_B\rangle\langle j_A j_B|, \qquad (50)$$

where $\widetilde{\rho}_A = \sum_{i,j=0}^{d-1} \widetilde{\rho}_{ij}|i_A\rangle\langle j_A|$ is a quantum state. Denote $\Delta(\cdot) = \sum_{i,j=0}^{d-1}\langle i_A j_B| \cdot |i_A j_B\rangle|i_A j_B\rangle\langle i_A j_B|$ as the completely dephasing channel on the bipartite systems. Let $\sigma_{AB} = \Delta(\rho_{\mathrm{mc}})$. We can check that σ_{AB} is a minimizer of the Rains bound for ρ_{mc}. Due to Theorem 7, we have the following result.

Proposition 10 (Maximally correlated states) *For any maximally correlated state ρ_{mc}, the number of prepared states n, the error tolerance $\varepsilon \in (0, 1)$, and the operation class $\Omega \in \{1\text{-LOCC, LOCC, SEP, PPT}\}$, it holds*

$$E_{\mathrm{D},\Omega}^{(1),\varepsilon}(\rho_{\mathrm{mc}}^{\otimes n}) = nD(\rho_{\mathrm{mc}}\|\Delta(\rho_{\mathrm{mc}}))$$
$$+ \sqrt{nV(\rho_{\mathrm{mc}}\|\Delta(\rho_{\mathrm{mc}}))}\,\Phi^{-1}(\varepsilon) + O(\log n). \quad (51)$$

Note that if $\widetilde{\rho}_A$ is a pure state, the maximally correlated state ρ_{mc} will also reduce to a bipartite pure states. The class of maximally correlated states also contains the mixture of Bell states ρ_{Bell}. By direct calculations, we also have $D(\rho_{\mathrm{mc}}\|\Delta(\rho_{\mathrm{mc}})) = I(A\rangle B)_{\rho_{\mathrm{mc}}}$ and $V(\rho_{\mathrm{mc}}\|\Delta(\rho_{\mathrm{mc}})) = V(A\rangle B)_{\rho_{\mathrm{mc}}}$.

Furthermore, the coherence theory is closely related to the entanglement theory due to the one-to-one correspondence between $\rho_{\mathrm{mc}} = \sum_{i,j=0}^{d-1} \widetilde{\rho}_{ij}|i_A i_B\rangle\langle j_A j_B|$ and $\widetilde{\rho}_A = \sum_{i,j=0}^{d-1} \widetilde{\rho}_{ij}|i_A\rangle\langle j_A|$. An interesting conjecture with a plethora of evidence is that any incoherent operation acting on a state $\widetilde{\rho}_A$ is equivalent to a LOCC operation acting on the associated maximally correlated state ρ_{mc} [67], [68]. If this conjecture holds, Proposition 10 will also give the second-order estimation for non-asymptotic coherence distillation.

Another common noise, in practice, is the so-called depolarizing noise [66], [69], which results in an isotropic state,

$$\rho_F = F \cdot \Psi_d + (1-F)\frac{\mathbb{1} - \Psi_d}{d^2 - 1}, \quad 0 \leq F \leq 1, \qquad (52)$$

where d is the local dimension of the maximally entangled state Ψ_d. The isotropic state is also the Choi-Jamiołkowski state of the depolarizing channel $\mathcal{N}(\rho) = p\rho + (1-p)\mathbb{1}/d$. Its 1-LOCC distillable entanglement is equal to the quantum capacity of the depolarizing channel [6], [18], [70], the determination of which is still a big open problem in quantum information theory. Here we study the non-asymptotic distillable entanglement of this particular class of states. For small blocklength n (e.g. $n \leq 100$), we can compute the exact distillation rate via a linear program. For large blocklength n (e.g $n > 100$), we need to employ the second-order estimation in Theorem 7.

Isotropic states possess the same symmetry as the maximally entangled states, which are invariant under any local unitary $U_A \otimes \overline{U}_B$. Exploiting such symmetry, we can simplify the PPT-assisted distillable entanglement for the n-fold isotropic state as a linear program. We note that the optimal fidelity for n-fold isotropic states can also be simplified to a linear program, which has been studied by Rains in [15]. Here, we follow Rains' approach of utilizing the structure of isotropic states and focus on the distillable rate of n-fold isotropic states under a given infidelity tolerance.

Proposition 11 (Isotropic states) *For any n-fold isotropic state $\rho_F^{\otimes n}$ with integer n and the error tolerance ε, its one-shot distillable entanglement under PPT operations is given by*

$$E_{D,PPT}^{(1),\varepsilon}(\rho_F^{\otimes n}) = \log \max \lfloor 1/\eta \rfloor \quad s.t.$$

$$0 \le m_i \le 1, \ \forall i = 0, 1, \cdots, n,$$

$$\sum_{i=0}^{n} \binom{n}{i} F^i (1-F)^{n-i} m_i \ge 1 - \varepsilon, \quad (53)$$

$$-\eta \le \sum_{i=0}^{n} x_{i,k} m_i \le \eta, \ \forall k = 0, 1, \cdots, n,$$

where the coefficients

$$x_{i,k} = \frac{1}{d^n} \sum_{m=\max\{0, i+k-n\}}^{\min\{i,k\}} \binom{k}{m}\binom{n-k}{i-m}$$
$$\cdot (-1)^{i-m} (d-1)^{k-m} (d+1)^{n-k+m-i}. \quad (54)$$

Proof: The technique is very similar to the one we use in the proof of Proposition 3. Consider the n-fold isotropic state

$$\rho_F^{\otimes n} = \sum_{i=0}^{n} f_i P_i^n(\Psi_d, \Psi_d^\perp), \quad \text{with} \quad (55)$$

$$f_i = F^i \left(\frac{1-F}{d^2-1}\right)^{n-i}, \Psi_d^\perp = \mathbb{1} - \Psi_d. \quad (56)$$

Here, $P_i^n(\Psi_d, \Psi_d^\perp)$ represents the sum of those n-fold tensor product terms with exactly i copies of Ψ_d. For example, $P_1^3(\Psi_d, \Psi_d^\perp) = \Psi_d \otimes \Psi_d^\perp \otimes \Psi_d^\perp + \Psi_d^\perp \otimes \Psi_d \otimes \Psi_d^\perp + \Psi_d^\perp \otimes \Psi_d^\perp \otimes \Psi_d$. Suppose M is the optimal solution of the optimization

$$E_{D,PPT}^{(1),\varepsilon}(\rho_F^{\otimes n}) = \log \max \lfloor 1/\eta \rfloor \quad s.t.$$
$$0 \le M_{AB} \le \mathbb{1}_{AB}, \quad (57)$$
$$\text{Tr} \, \rho_F^{\otimes n} M_{AB} \ge 1 - \varepsilon,$$
$$-\eta \mathbb{1}_{AB} \le M_{AB}^{T_B} \le \eta \mathbb{1}_{AB}.$$

Then for any local unitary $U = \bigotimes_{i=1}^{n} \left(U_A^i \otimes \overline{U}_B^i\right)$ where i denotes the i-th copies of corresponding system, $UMU^\dagger$ is a also optimal solution. Convex combinations of optimal solutions are also optimal. So we can take the optimal solution M to be an operator which is invariant under any local unitary $\bigotimes_{i=1}^{n} \left(U_A^i \otimes \overline{U}_B^i\right)$. Moreover, since $\rho_F^{\otimes n}$ is invariant under the symmetric group acting by permuting the tensor factors, we can take the optimal solution M of the form $\sum_{i=0}^{n} m_i P_i^n(\Psi_d, \Psi_d^\perp)$ without loss of generality.

Since $P_i^n(\Psi_d, \Psi_d^\perp)$ are orthogonal projections, the operator M has eigenvalues $\{m_i\}_{i=0}^{n}$ without considering degeneracy. Next, we need to know the eigenvalues of M^{T_B}. Decomposing operators $\Psi_d^{T_B}$ and $\Psi_d^{\perp T_B}$ into orthogonal projections, i.e.,

$$\Psi_d^{T_B} = (P_+ - P_-)/d, \quad (58)$$

$$\Psi_d^{\perp T_B} = (1 - 1/d)P_+ + (1 + 1/d)P_-, \quad (59)$$

where P_+ and P_- are symmetric and anti-symmetric projections respectively and collecting the terms with respect to $P_k^n(P_+, P_-)$, we have

$$M^{T_B} = \sum_{i=0}^{n} m_i P_i^n \left(\Psi_d^{T_B}, \Psi_d^{\perp T_B}\right) \quad (60)$$

$$= \sum_{i=0}^{n} m_i \left(\sum_{k=0}^{n} x_{i,k} P_k^n(P_+, P_-)\right) \quad (61)$$

$$= \sum_{k=0}^{n} \left(\sum_{i=0}^{n} x_{i,k} m_i\right) P_k^n(P_+, P_-). \quad (62)$$

Since $P_k^n(P_+, P_-)$ are also orthogonal projections, M^{T_B} has eigenvalues $\{t_k\}_{k=0}^{n}$ without considering degeneracy, where $t_k = \sum_{i=0}^{n} x_{i,k} m_i$. As for the condition $\text{Tr} \, M\rho_F^{\otimes n} \ge 1 - \varepsilon$, we have

$$\text{Tr} \, M\rho_F^{\otimes n} = \text{Tr} \sum_{i=0}^{n} f_i m_i P_i^n(\Psi_d, \Psi_d^\perp) \quad (63)$$

$$= \sum_{i=0}^{n} f_i m_i \binom{n}{i} (d^2 - 1)^{n-i} \quad (64)$$

$$= \sum_{i=0}^{n} \binom{n}{i} F^i (1-F)^{n-i} m_i. \quad (65)$$

Taking Eq. (65) and the eigenvalues of M, M^{T_B} into Eq. (57), we have the desired result. ∎

This linear program can be solved *exactly* via Mathematica. In Figure 1, we plot the one-shot distillable entanglement for the n-fold isotropic state $\rho_F^{\otimes n}$ with $d = 3$, $F = 0.9$, and error tolerance 0.001. The blocklength n ranges from 1 to 100. We observe that even if we were able to coherently manipulate 100 copies of the states with the broad class of PPT assistance, the maximal distillation rate still could not reach the hashing bound $I(A\rangle B)_{\rho_F}$ which is asymptotically achievable under 1-LOCC operations. This demonstrates that the asymptotic bounds cannot provide helpful estimations in the practical scenario.

For the approximation of large blocklength distillation, we employ the second-order bounds in Theorem 7. In Figure 2, we show the second-order estimation for n-fold isotropic state $\rho_F^{\otimes n}$ with $d = 3$, $F = 0.9$, and error tolerance 0.001. In this figure we focus on the large blocklength ($n \ge 100$) regime and use a logarithmic scale for the horizontal axis. The second-order bounds in Theorem 7 are not tight, as expected, for isotropic states. But they provide a more refined estimation than the known asymptotic bounds. In Figure 2, the finite blocklength distillation rate lies between the two dashed lines, while the asymptotic rate lies between the two solid lines.
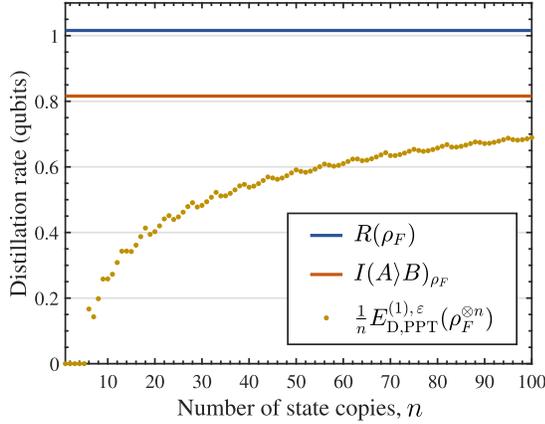
Fig. 1. The dotted line shows the exact value of distillation rate for n-fold isotropic state $\rho_F^{\otimes n}$ with $F = 0.9$, local dimension $d = 3$. The error tolerance is taken at $\varepsilon = 0.001$ and the blocklength n ranges from 1 to 100. The solid line below is the hashing bound, while the solid line above is the Rains bound.

Fig. 3. The dash-dotted line is the fitting curve of exact values of distillation rate for n-fold isotropic state $\rho_F^{\otimes n}$ with $F = 0.9$, local dimension $d = 3$. The error tolerance is taken at $\varepsilon = 0.001$. The dashed line is the second-order upper bound in Eq. (41) and the solid line is the Rains bound.

Fig. 2. The two dashed lines show the second-order lower bound in Eq. (40) and upper bound in Eq. (41) for n-fold isotropic state $\rho_F^{\otimes n}$ with $F = 0.9$, local dimension $d = 3$. The error tolerance is taken at $\varepsilon = 0.001$ and the blocklength n ranges from 10^2 to 10^7. The solid line below is the hashing bound, while the solid line above is the Rains bound.

An interesting observation is made when we present Figure 1 and Figure 2 in a single plot. The linear program in Figure 1 is only implemented for n less than 100 due to the limited computational power. But we can use the curve fitting via least-squares method and construct an ansatz curve

$$c_1 + c_2 \frac{1}{\sqrt{n}} + c_3 \frac{\log n}{n} + c_4 \frac{1}{n}, \qquad (66)$$

which has the best fit to the series of points $\frac{1}{n} E_{\mathrm{D,PPT}}^{(1),\varepsilon}(\rho_F^{\otimes n})$ ($1 \leq n \leq 100$) in Figure 1. Combining with the second-order upper bound in Figure 2, we get Figure 3. It shows that for small n, the second-order upper bound does not give an accurate estimation since we ignore the term $O(\log n / n)$. But for large n ($\geq 10^2$), the fitting curve almost coincides with the second-order upper bound, demonstrating that the second-order upper bound works better for large blocklength. The convergence of the fitting curve indicates that $E_{\mathrm{D,PPT}}(\rho_F) = R(\rho_F)$ for isotropic states ρ_F. It

would be of great interest to find an analytical proof to this conjecture.

VI. DISCUSSIONS

We have provided both theoretical and numerical[1] results for the entanglement distillation in the non-asymptotic regimes. Since entanglement distillation has become a central building block of quantum network proposals [65], [69], [71], [72], our finite blocklength estimations could be applied as useful benchmarks for experimentalists to build a reliable quantum network in the future. Theoretically, we have obtained an exact characterization of the one-shot entanglement distillation under PPT operations in terms of the hypothesis testing relative entropy. This result not only leads to an improved understanding of the resource theory of entanglement, but also provides a potential approach to resolve the distillable entanglement under PPT operations or improve the Rains bound by taking other forms of feasible solution, for example, non-i.i.d. operators.

APPENDIX A
AN EXAMPLE FOR THEOREM 4

In this section, we give an explicit example to show that the optimal solution in the optimization

$$\min \left\{ D_H^\varepsilon(\rho_{AB} \| G_{AB}) \mid \| G^{\mathsf{T}_B} \|_1 \leq 1, G = G^\dagger \right\} \qquad (67)$$

is not taken at any positive operator G_{AB}. Specifically, we show a strict difference between the following two optimizations:

$$\mathrm{OPT1} = \min_{\substack{\| G^{\mathsf{T}_B} \|_1 \leq 1 \\ G = G^\dagger}} D_H^\varepsilon(\rho_{AB} \| G_{AB}), \qquad (68)$$

$$\mathrm{OPT2} = \min_{\substack{\| G^{\mathsf{T}_B} \|_1 \leq 1 \\ G \geq 0}} D_H^\varepsilon(\rho_{AB} \| G_{AB}). \qquad (69)$$

[1]The codes for numerical calculations are available at https://github.com/fangkunfred/entanglement-distillation.
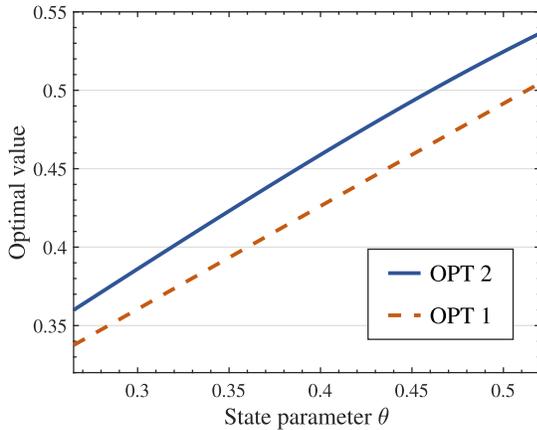
Fig. 4.   This figure demonstrates the difference of optimal value in OPT1 and OPT2 with respect to the state ρ_θ. The solid line depicts the optimal value of OPT1 while the dashed line depicts the optimal value of OPT2. The parameter θ ranges from $\pi/12$ to $\pi/6$ and error tolerance is taken at $\varepsilon = 1 - \sqrt{3}/2$.

Note that the dual SDP of the quantum hypothesis testing relative entropy is given by

$$D_H^\varepsilon(\rho_{AB}\|G_{AB}) = -\log \max \ \mathrm{Tr}\,X + t(1-\varepsilon) \quad \text{s.t.}$$
$$G - X - t\rho \geq 0, \ X \leq 0, \ t \geq 0. \tag{70}$$

Thus we have the following SDPs:

$$\mathrm{OPT1} = -\log \max \ \mathrm{Tr}\,X + t(1-\varepsilon) \quad \text{s.t.}$$
$$G - X - t\rho \geq 0, \ X \leq 0, \ t \geq 0, \tag{71}$$
$$\|G^{\mathsf{T}_B}\|_1 \leq 1, \ G = G^\dagger,$$

and

$$\mathrm{OPT2} = -\log \max \ \mathrm{Tr}\,X + t(1-\varepsilon) \quad \text{s.t.}$$
$$G - X - t\rho \geq 0, \ X \leq 0, \ t \geq 0, \tag{72}$$
$$\|G^{\mathsf{T}_B}\|_1 \leq 1, \ G \geq 0.$$

We implement these two SDPs for the quantum state $\rho_\theta = (3|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2|)/4$ with $|\varphi_1\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$ and $|\varphi_2\rangle = |10\rangle$. The difference between OPT1 and OPT2 is shown in Figure 4. The numerics is run via the solver SDPT3 which can be solved to a very high (near-machine) precision. The maximal gap in the plot is approximately 3.4×10^{-2}.

APPENDIX B
NUMERICAL ESTIMATION OF RAINS BOUND

In this section, we provide an algorithm to numerically compute the Rains bound with high accuracy. In particular, the calculation of upper and lower bounds of the Rains bound can have near-machine precision while the final result of Rains bound itself is within error tolerance 10^{-6} by default. This algorithm closely follows the approach in [74] and [75] which intends to compute the PPT-relative entropy of entanglement.

Note that the only difference between the Rains bound and the PPT-relative entropy of entanglement is the feasible set. Due to the similarity between these two quantities, we can have a similar algorithm for the Rains bound. For the sake of completeness, we will restate the main idea of this algorithm

and clarify that our adjustment will work to compute the Rains bound. In the following discussion, we will consider the natural logarithm, denoted as ln, for convenience.

The key idea for this algorithm is based on the cutting-plane method combined with semidefinite programming. Clearly, calculating the Rains bound is equivalent to the optimization problem

$$\min_{\sigma_{AB}\in\mathrm{PPT}'} (-\mathrm{Tr}\,\rho_{AB}\ln\sigma_{AB}), \quad \text{with} \tag{73}$$

$$\mathrm{PPT}' = \big\{ \sigma_{AB} \geq 0 \,\big|\, \|\sigma_{AB}^{\mathsf{T}_B}\|_1 \leq 1 \big\}. \tag{74}$$

If we relax the minimization over all quantum states, the optimal solution is taken at $\sigma = \rho$. Thus $-\mathrm{Tr}\,\rho\ln\rho$ provides a trivial lower bound on (73). Since the objective function is convex with respect to σ over the Rains set (PPT'), its epigraph is supported by tangent hyperplanes at every interior point $\sigma^{(i)} \in \mathrm{int}\,\mathrm{PPT}'$. Thus we can construct a successively refined sequence of approximations to the epigraph of the objective function restricted to the interior of the Rains set.

Specifically, for an arbitrary positive definite operator X, we have a spectral decomposition $X = U_X\mathrm{diag}(\lambda_X)U_X^\dagger$ with unitary matrix U_X and diagonal matrix $\mathrm{diag}(\lambda_X)$ formed by the eigenvalues λ_X. Then we have the first-order expansion

$$\ln(X + \Delta) = \ln X \tag{75}$$
$$+ U_X\left[D(\lambda_X) \circ U_X^\dagger \Delta U_X\right]U_X^\dagger + O(\|\Delta\|^2),$$

where $D(\lambda)$ is the Hermitian matrix given by

$$D(\lambda)_{i,j} = \begin{cases} (\ln\lambda_i - \ln\lambda_j)/(\lambda_i - \lambda_j), & \lambda_i \neq \lambda_j, \\ 1/\lambda_i, & \lambda_i = \lambda_j. \end{cases} \tag{76}$$

For any given set of feasible points $\{\sigma^{(i)}\}_{i=0}^N \subset \mathrm{int}\,\mathrm{PPT}'$, we have spectral decompositions $\sigma^{(i)} = U_{(i)}\mathrm{diag}(\lambda^{(i)})U_{(i)}^\dagger$. Then $\mathrm{epi}(-\mathrm{Tr}\,\rho\ln\sigma)|_{\mathrm{int}\,\mathrm{PPT}'}$ is a subset of all $(\sigma, t) \in \mathrm{int}\,\mathrm{PPT}' \times \mathbb{R}$ satisfying

$$-\mathrm{Tr}\,\rho\left\{\ln\sigma^{(i)}\right. \tag{77}$$
$$\left. + U_{(i)}\left[D(\lambda^{(i)}) \circ U_{(i)}^\dagger(\sigma - \sigma^{(i)})U_{(i)}\right]U_{(i)}^\dagger\right\} \leq t,$$

for all $i = 0, \cdots, N$. Equivalently, we can introduce slack variable s_i on the l.h.s. of Eq. (77) and have

$$\mathrm{Tr}\,E^{(i)}\sigma + t - s_i = -\mathrm{Tr}\,\rho\ln\sigma^{(i)} + \mathrm{Tr}\,E^{(i)}\sigma^{(i)}, \tag{78}$$

$$s_i \geq 0, \ E^{(i)} = U_{(i)}\left[D(\lambda^{(i)}) \circ U_{(i)}^\dagger\rho U_{(i)}\right]U_{(i)}^\dagger, \tag{79}$$

for all $i = 0, \cdots, N$. So the optimal value of optimization

$$\min \ t \quad \text{s.t.}$$
$$\mathrm{Tr}\,E^{(i)}\sigma + t - s_i = -\mathrm{Tr}\,\rho\ln\sigma^{(i)} + \mathrm{Tr}\,E^{(i)}\sigma^{(i)}, \tag{80}$$
$$s_i \geq 0, \ i = 0, \cdots, N, \ \sigma \in \mathrm{PPT}'$$

provides a lower bound on (73). For any feasible point $\sigma^* \in \mathrm{PPT}'$, $-\mathrm{Tr}\,\rho\ln\sigma^*$ provides an upper bound on (73). For each iteration of the algorithm, we add a interior point $\sigma^{(N+1)}$ of the Rains set to the set $\{\sigma^{(i)}\}_{i=0}^N$, which may lead to a tighter lower bound and update the feasible point σ^* if $\sigma^{(N+1)}$ provides a tighter upper bound. We use the variables $\overline{R}$ and $\underline{R}$ to store the upper and lower bounds. Since $\underline{R}$ and

$\overline{R}$ are nondecreasing and nonincreasing, respectively, at each iteration, we can terminate the algorithm when $\underline{R}$ and $\overline{R}$ are close enough, for example, less than given tolerance ε. The full algorithm is presented in Algorithm 1.

Algorithm 1 Rains Bound Algorithm

1: **Input:** bipartite state ρ_{AB} and dimensions d_A, d_B

2: **Output:** Upper bound $\overline{R}$, lower bound $\underline{R}$

3: **if** $\rho \in \mathrm{PPT}'$ **then**

4:     **return** $\underline{R} = \overline{R} = 0$

5: **else**

6:     **initialize** $\varepsilon = 10^{-6}$, $N = 0$, $\sigma^* = \sigma^{(0)} = \mathbb{1}_{AB}/(d_A d_B)$, $\underline{R} = -\mathrm{Tr}\,\rho \ln \rho$, $\overline{R} = -\mathrm{Tr}\,\rho \ln \sigma^*$

7:     **while** $\overline{R} - \underline{R} \geq \varepsilon$ **do**

8:         **solve** $\min \left\{ t \mid (78\text{-}79), t \geq \underline{R}, \sigma \in \mathrm{PPT}' \right\}$

9:         store optimal solution $(\underline{t}, \underline{\sigma})$, update $\underline{R} = \underline{t}$

10:        **if** $\overline{R} - \underline{R} \leq \varepsilon$ **then**

11:            **return** $\underline{R}$, $\overline{R}$

12:        **else**

13:            add point $\sigma^{(N+1)}$, and **set** $N = N + 1$

14:            **if** $-\mathrm{Tr}\,\rho \ln \sigma^{(N)} \leq -\mathrm{Tr}\,\rho \ln \sigma^*$ **then**

15:                **update** $\sigma^* = \sigma^{(N)}$, $\overline{R} = -\mathrm{Tr}\,\rho \ln \sigma^*$

The following lemma ensures that the condition $\sigma \in \mathrm{PPT}'$ is SDP representable.

Lemma 12 *The condition $\sigma \in \mathrm{PPT}'$ holds if and only if $\sigma \geq 0$ and there exist operators σ_+, $\sigma_- \geq 0$ such that $\sigma^{T_B} = \sigma_+ - \sigma_-$ and $\mathrm{Tr}(\sigma_+ + \sigma_-) \leq 1$.*

Proof: If $\sigma \in \mathrm{PPT}'$, then $\sigma \geq 0$. Use the spectral decomposition $\sigma^{T_B} = \sigma_+ - \sigma_-$, where σ_+ and σ_- are positive operator with orthogonal support. Then $|\sigma^{T_B}| = \sigma_+ + \sigma_-$ and $\mathrm{Tr}(\sigma_+ + \sigma_-) = \|\sigma^{T_B}\|_1 \leq 1$. On the other hand, if there exist positive operators σ_+ and σ_- such that $\sigma^{T_B} = \sigma_+ - \sigma_-$ and $\mathrm{Tr}(\sigma_+ + \sigma_-) \leq 1$, then $\|\sigma^{T_B}\|_1 = \|\sigma_+ - \sigma_-\|_1 \leq \|\sigma_+\|_1 + \|\sigma_-\|_1 = \mathrm{Tr}(\sigma_+ + \sigma_-) \leq 1$. Thus $\sigma \in \mathrm{PPT}'$. ∎

For given $\left\{ \sigma^{(i)} \right\}_{i=0}^{N}$, Step 8 in Algorithm 1 is an SDP which can be explicitly written as

$$\min \; t \quad \text{s.t.}$$
$$(78\text{-}79), i = 0, \cdots, N,$$
$$t \geq \underline{R}, \; \sigma \geq 0, \; \sigma_+ \geq 0, \; \sigma_- \geq 0, \tag{81}$$
$$\sigma^{T_B} = \sigma_+ - \sigma_-, \; \mathrm{Tr}(\sigma_+ + \sigma_-) \leq 1.$$

As for Step 13, the variable $\sigma^{(N+1)}$ can be given by

$$\sigma^{(N+1)} = \arg \min \; -\mathrm{Tr}\,\rho \ln \sigma \quad \text{s.t.}$$
$$\sigma = \alpha Z + (1 - \alpha)\underline{\sigma}, \; \alpha \in [0, 1], \tag{82}$$

where Z is some fixed reference point. This one-dimensional minimization can be efficiently performed using the standard derivative-based bisection scheme [73].

As a by-product, the above algorithm can be used to check the nonadditivity of the Rains bound, which has been recently
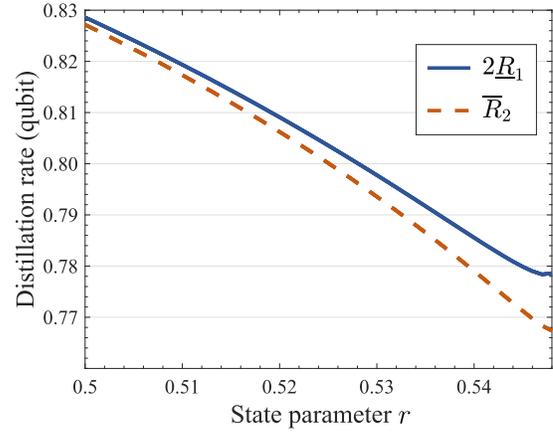
Fig. 5. This figure demonstrates the difference between the lower bound $2\underline{R}_1$ on $2R(\rho_r)$ and the upper bound $\overline{R}_2$ on $R(\rho_r^{\otimes 2})$. The solid line depicts $2\underline{R}_1$ while the dashed line depicts $\overline{R}_2$.

proved in [75]. We also consider the states ρ_r defined in [75]. Denote $\underline{R}_1$ the lower bound computed by our algorithm for $R(\rho_r)$ and $\overline{R}_2$ the upper bound computed by our algorithm for $R(\rho_r^{\otimes 2})$. In Figure 5, we can clearly observe that there is a strict gap between $\overline{R}_2$ and $2\underline{R}_1$, which implies $R(\rho_r^{\otimes 2}) \leq \overline{R}_2 < 2\underline{R}_1 \leq 2R(\rho_r)$. Note that $\underline{R}_1$ and $\overline{R}_2$ only depend on the SDPs in Eqs. (81) and (82), both of which can be solved to a very high (near-machine) precision, while the maximal gap in the plot is approximately 10^{-2}. Thus our algorithm provides direct numerical evidence (not involving any other entanglement measures) for the nonadditivity of the Rains bound.

Remark After the completion of this work, we notice that there is another approach to efficiently calculating the Rains bound in [77] and [78]. In these works, the authors make use of rational (Padé) approximations of the (matrix) logarithm function and then transform the rational functions to SDPs. Without the successive refinement, their algorithm can be much faster with relatively high accuracy. However, our algorithm is efficient enough in low-dimensional cases.

REFERENCES

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, no. 2, pp. 865–942, Jun. 2009.

[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, no. 13, pp. 1895–1899, Mar. 1993.

[3] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, vol. 69, no. 20, pp. 2881–2884, Nov. 1992.

[4] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proc. Int. Conf. Comput., Syst. Signal Process.*, Bangalore, India, Dec. 1984, pp. 175–179.

[5] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, no. 6, pp. 661–663, Aug. 1991.

[6] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 5, pp. 3824–3851, 1996.

[7] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels," *Phys. Rev. Lett.*, vol. 76, no. 5, pp. 722–725, Jan. 1996.

[8] E. M. Rains, "Rigorous treatment of distillable entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 60, no. 1, pp. 173–178, Jul. 1999.

[9] S. Virmani and M. B. Plbnio, "An introduction to entanglement measures," *Quantum Inf. Comput.*, vol. 7, no. 1, pp. 1–51, 2007.

[10] S. Waeldchen, J. Gertis, E. T. Campbell, and J. Eisert, "Renormalizing entanglement distillation," *Phys. Rev. Lett.*, vol. 116, no. 2, Jan. 2016, Art. no. 020502.

[11] S. Popescu and D. Rohrlich, "Thermodynamics and the measure of entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 56, no. 5, pp. R3319–R3321, Nov. 1997.

[12] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A, Gen. Phys.*, vol. 57, no. 3, pp. 1619–1633, Mar. 1998.

[13] E. M. Rains, "Bound on distillable entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 60, no. 1, pp. 179–184, Jul. 1999.

[14] G. Vidal and R. F. Werner, "Computable measure of entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 65, no. 3, Feb. 2002, Art. no. 032314.

[15] E. M. Rains, "A semidefinite program for distillable entanglement," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2921–2933, Nov. 2001.

[16] M. Horodecki, P. Horodecki, and R. Horodecki, "Asymptotic manipulations of entanglement can exhibit genuine irreversibility," *Phys. Rev. Lett.*, vol. 84, no. 19, pp. 4260–4263, May 2000.

[17] M. Christandl and A. Winter, "'Squashed entanglement': An additive entanglement measure," *J. Math. Phys.*, vol. 45, no. 3, pp. 829–840, Mar. 2004.

[18] F. Leditzky, N. Datta, and G. Smith, "Useful states and entanglement distillation," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 4689–4708, Jul. 2018.

[19] X. Wang and R. Duan, "Improved semidefinite programming upper bound on distillable entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 94, no. 5, Nov. 2016, Art. no. 050301.

[20] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 461, pp. 207–235, Jan. 2005.

[21] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[22] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.

[23] V. Y. F. Tan, "Asymptotic estimates in information theory with non-vanishing error probabilities," *Found. Trends Commun. Inf. Theory*, vol. 10, no. 4, pp. 1–184, 2014.

[24] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Phys. Rev. Lett.*, vol. 108, no. 20, May 2012, Art. no. 200501.

[25] J. M. Renes and R. Renner, "Noisy channel coding via privacy amplification and information reconciliation," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7377–7385, Nov. 2011.

[26] M. Berta, M. Christandl, and R. Renner, "A conceptually simple proof of the quantum reverse Shannon theorem," in *Theory of Quantum Computation, Communication, and Cryptography*. Berlin, Germany: Springer, 2011, pp. 131–140. doi: 10.1007/978-3-642-18073-6_11.

[27] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7693–7710, Nov. 2013.

[28] D. Leung and W. Matthews, "On the power of PPT-preserving and non-signalling codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4486–4499, Sep. 2015.

[29] N. Datta and M.-H. Hsieh, "One-shot entanglement-assisted quantum and classical communication," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1929–1939, Mar. 2013.

[30] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7317–7329, Nov. 2014.

[31] S. Beigi, N. Datta, and F. Leditzky, "Decoding quantum information via the Petz recovery map," *J. Math. Phys.*, vol. 57, no. 8, Art. no. 082203, Aug. 2016.

[32] M. Tomamichel, *Quantum Information Processing With Finite Resources—Mathematical Foundations* (SpringerBriefs in Mathematical Physics), vol. 5. Springer, 2016. doi: 10.1007/978-3-319-21891-5.

[33] M. Tomamichel, M. Berta, and J. M. Renes, "Quantum coding with finite resources," *Nature Commun.*, vol. 7, May 2016, Art. no. 11419.

[34] X. Wang, W. Xie, and R. Duan, "Semidefinite programming strong converse bounds for classical capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 640–653, Jan. 2018.

[35] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1385–1403, Feb. 2018.

[36] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," *Commun. Math. Phys.*, vol. 355, no. 3, pp. 1283–1315, Nov. 2017.

[37] F. Buscemi and N. Datta, "Distilling entanglement from arbitrary resources," *J. Math. Phys.*, vol. 51, no. 10, Oct. 2010, Art. no. 102201.

[38] F. G. S. L. Brandão and N. Datta, "One-shot rates for entanglement manipulation under non-entangling maps," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1754–1760, Mar. 2011.

[39] N. Datta and F. Leditzky, "Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 582–608, Jan. 2015.

[40] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, "Everything you always wanted to know about LOCC (but were afraid to ask)," *Commun. Math. Phys.*, vol. 328, no. 1, pp. 303–326, May 2014.

[41] K. Audenaert, M. B. Plenio, and J. Eisert, "Entanglement cost under positive-partial-transpose-preserving operations," *Phys. Rev. Lett.*, vol. 90, no. 2, Jan. 2003, Art. no. 027901.

[42] X. Wang and R. Duan, "Irreversibility of asymptotic entanglement manipulation under quantum operations completely preserving positivity of partial transpose," *Phys. Rev. Lett.*, vol. 119, Nov. 2017, Art. no. 180506.

[43] W. Matthews and A. Winter, "Pure-state transformations and catalysis under operations that completely preserve positivity of partial transpose," *Phys. Rev. A, Gen. Phys.*, vol. 78, no. 1, Jul. 2008, Art. no. 012317.

[44] X. Wang and M. M. Wilde. (Sep. 2018). "Exact entanglement cost of quantum states and channels under PPT-preserving operations." [Online]. Available: http://arxiv.org/abs/1809.09592

[45] C. H. Bennett *et al.*, "Quantum nonlocality without entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 59, no. 2, pp. 1070–1091, Feb. 1999.

[46] C. W. Helstrom, *Quantum Detection and Estimation Theory*, vol. 123. New York, NY, USA: Academic, 1976. doi: 10.1016/s0076-5392(08)x6017-5.

[47] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, no. 1, pp. 99–114, 1991.

[48] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2428–2433, Nov. 2000.

[49] M. Hayashi. (Sep. 2017). "Role of hypothesis testing in quantum information." [Online]. Available: http://arxiv.org/abs/1709.07701

[50] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, "Distinguishing separable and entangled states," *Phys. Rev. Lett.*, vol. 88, no. 18, Apr. 2002, Art. no. 187904.

[51] A. W. Harrow, A. Natarajan, and X. Wu, "An improved semidefinite programming hierarchy for testing entanglement," *Commun. Math. Phys.*, vol. 352, no. 3, pp. 881–904, Mar. 2017.

[52] M. Berta, O. Fawzi, and V. B. Scholz, "Quantum bilinear optimization," *SIAM J. Optim.*, vol. 26, no. 3, pp. 1529–1564, Jan. 2016.

[53] X. Wang and R. Duan, "Separation between quantum Lovász number and entanglement-assisted zero-error classical capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1454–1460, Mar. 2018.

[54] I. Kogias, P. Skrzypczyk, D. Cavalcanti, A. Acín, and G. Adesso, "Hierarchy of steering criteria based on moments for all bipartite quantum systems," *Phys. Rev. Lett.*, vol. 115, no. 21, Nov. 2015, Art. no. 210401.

[55] Y. Li, X. Wang, and R. Duan, "Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario," *Phys. Rev. A, Gen. Phys.*, vol. 95, no. 5, May 2017, Art. no. 052346.

[56] B. Regula, K. Fang, X. Wang, and G. Adesso, "One-shot coherence distillation," *Phys. Rev. Lett.*, vol. 121, no. 1, Jul. 2018, Art. no. 010401.

[57] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra Appl.*, vol. 10, no. 3, pp. 285–290, 1975.

[58] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Rep. Math. Phys.*, vol. 3, no. 4, pp. 275–278, 1972.

[59] M. Sion, "On general minimax theorems," *Pacific J. Math.*, vol. 8, no. 1, pp. 171–176, 1958.

[60] K. Audenaert, B. De Moor, K. G. H. Vollbrecht, and R. F. Werner, "Asymptotic relative entropy of entanglement for orthogonally invariant states," *Phys. Rev. A, Gen. Phys.*, vol. 66, no. 3, Sep. 2002, Art. no. 32310.

[61] K. Li, "Second-order asymptotics for quantum hypothesis testing," *Ann. Statist.*, vol. 42, no. 1, pp. 171–189, 2014.

[62] M. M. Wilde, M. Tomamichel, and M. Berta, "Converse bounds for private communication over quantum channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 3, pp. 1792–1817, Mar. 2017.

[63] F. Buscemi and N. Datta, "General theory of environment-assisted entanglement distillation," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1940–1954, Mar. 2013.

[64] H.-K. Lo and S. Popescu, "Concentrating entanglement by local actions: Beyond mean values," *Phys. Rev. A, Gen. Phys.*, vol. 63, no. 2, 2001, Art. no. 022301.

[65] L. Childress, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, "Fault-tolerant quantum communication based on solid-state photon emitters," *Phys. Rev. Lett.*, vol. 96, no. 7, Feb. 2006, Art. no. 070504.

[66] E. T. Campbell, "Distributed quantum-information processing with minimal local resources," *Phys. Rev. A, Gen. Phys.*, vol. 76, no. 4, Oct. 2007, Art. no. 040302.

[67] A. Winter and D. Yang, "Operational resource theory of coherence," *Phys. Rev. Lett.*, vol. 116, no. 12, Mar. 2016, Art. no. 120404.

[68] A. Streltsov, G. Adesso, and M. B. Plenio. (Sep. 2016). "*Colloquium*: Quantum coherence as a resource." [Online]. Available: http://arxiv.org/abs/1609.02439

[69] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, "Freely scalable quantum technologies using cells of 5-to-50 qubits with very lossy and noisy photonic links," *Phys. Rev. X*, vol. 4, no. 4, Dec. 2014, Art. no. 041041.

[70] F. Leditzky, D. Leung, and G. Smith, "Quantum and private capacities of low-noise channels," *Phys. Rev. Lett.*, vol. 120, no. 16, Apr. 2018, Art. no. 160503.

[71] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, "Quantum repeaters based on entanglement purification," *Phys. Rev. A, Gen. Phys.*, vol. 59, no. 1, pp. 169–181, Jan. 1999.

[72] D. Gottesman, T. Jennewein, and S. Croke, "Longer-baseline telescopes using quantum repeaters," *Phys. Rev. Lett.*, vol. 109, no. 7, Aug. 2012, Art. no. 070503.

[73] Y. Zinchenko, S. Friedland, and G. Gour, "Numerical estimation of the relative entropy of entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 82, no. 5, Nov. 2010, Art. no. 52336.

[74] M. W. Girard, Y. Zinchenko, S. Friedland, and G. Gour, "Erratum: Numerical estimation of the relative entropy of entanglement [phys. rev. a 82, 052336 (2010)]," *Phys. Rev. A, Gen. Phys.*, vol. 91, Feb. 2015, Art. no. 029901.

[75] X. Wang and R. Duan, "Nonadditivity of Rains' bound for distillable entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 95, no. 6, Jun. 2017, Art. no. 062322.

[76] H. Fawzi and O. Fawzi, "Efficient optimization of the quantum relative entropy," *J. Phys. A, Math. Theor.*, vol. 51, no. 15, Apr. 2018, Art. no. 154003.

[77] H. Fawzi, J. Saunderson, and P. A. Parrilo, "Semidefinite approximations of the matrix logarithm," *Found. Comput. Math.*, vol. 19, no. 2, pp. 259–296, Mar. 2018.

Kun Fang is currently a Postdoctoral Researcher at the Department of Applied Mathematics and Theoretical Physics, University of Cambridge. He received the Ph.D. degree in quantum information from the University of Technology Sydney in October 2018. He received the B.S. degree from Wuhan University. His research interests include the theory of quantum information and optimization theory.

Xin Wang is currently a Hartree Postdoctoral Fellow at the Joint Center for Quantum Information and Computer Science (QuICS) at the University of Maryland, College Park. He received the Ph.D. degree in quantum information from the University of Technology Sydney in August 2018. His research interests include quantum information, quantum computation, and optimization theory.

Marco Tomamichel (M'13-SM'16) is a Senior Lecturer at the Centre for Quantum Software and Information and School of Software, Faculty of Engineering and Information Technology at the University of Technology Sydney. He received a M.Sc. in Electrical Engineering and Information Technology degree from ETH Zurich (Switzerland) in 2007. He then graduated with a Ph.D. in Physics at the Institute of Theoretical Physics at ETH Zurich in 2012. Before commencing his current position, he has been a Senior Research Fellow at the Centre for Quantum Technologies at the National University of Singapore and a Lecturer at the School of Physics at the University of Sydney. His research interests include classical and quantum information theory with finite resources as well as applications to cryptography.

Runyao Duan is currently the founding director of the Institute for Quantum Computing at Baidu Research. He received the B.S. and Ph.D. degrees in Computer Science from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in July of 2002 and 2006, respectively. In August 2006, he became an Assistant Professor at the same Department. From December 2008 to September 2016, he was with the Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney (UTS), Australia, initially as Senior Lecturer, from August 2010 as Associate Professor, and from July 2012 as (Australian Research Council) Future Fellow and Professor. From 15 September 2016, he became the Director of Centre for Quantum Software and Information, UTS. His research interests include quantum information theory, quantum computation, and theoretical computer science.