Quantum Channel Simulation and the Channel's Smooth Max-Information

Kun Fang, Xin Wang, Marco Tomamichel, *Senior Member, IEEE*, and Mario Berta

Abstract—We study the general framework of quantum channel simulation, that is, the ability of a quantum channel to simulate another one using different classes of codes. First, we show that the minimum error of simulation and the one-shot quantum simulation cost under no-signalling assisted codes are given by semidefinite programs. Second, we introduce the channel's smooth max-information, which can be seen as a one-shot generalization of the mutual information of a quantum channel. We provide an exact operational interpretation of the channel's smooth max-information as the one-shot quantum simulation cost under no-signalling assisted codes, which significantly simplifies the study of channel simulation and provides insights and bounds for the case under entanglement-assisted codes. Third, we derive the asymptotic equipartition property of the channel's smooth max-information; i.e., it converges to the quantum mutual information of the channel in the independent and identically distributed asymptotic limit. This implies the quantum reverse Shannon theorem in the presence of no-signalling correlations. Finally, we explore the simulation cost of various quantum channels.

Index Terms—Quantum channel simulation, quantum channel coding, no-signalling, max-information, semidefinite programming.

I. INTRODUCTION

CHANNEL simulation is a fundamental problem in information theory. It asks how to use a (noisy) channel from Alice (A) to Bob (B) to simulate another (noisy) channel also from A to B [1], [2]. Depending on the different resources available between A and B, this simulation problem has many variants.

K. Fang was with the Centre for Quantum Software and Information, University of Technology Sydney, Sydney, 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, Sydney, NSW 2007, Australia. He is now with the Institute for Quantum Computing, Baidu Research, Beijing 100193, China (e-mail: wangxinfelix@gmail.com).

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

M. Berta is with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: m.berta@imperial.ac.uk).

Communicated by M. Mosonyi, 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.2019.2943858

For classical channels, Shannon's noisy channel coding theorem determines the capability of noisy classical channels to simulate noiseless ones [3]. Dual to this famous coding theorem, the 'reverse Shannon theorem' concerns the use of noiseless channels to simulate noisy ones [2]. Specifically, every channel can be simulated using an amount of classical communication equal to the capacity of the channel when there is free shared randomness between A and B in the asymptotic setting [2]. For quantum channels, the case when A and B share an unlimited amount of entanglement has been completely solved by the quantum reverse Shannon theorem (QRST) [4], [5], which states that the rate to optimally simulate a quantum channel in the asymptotic setting is determined by its entanglement-assisted classical capacity. In the zero-error scenario [6], using one channel to simulate another exactly with the aid of no-signalling correlations has been studied recently in [7]–[9], while the simulation with free quantum operations that completely preserve positivity of the partial transpose has been studied in [10]. The problem of quantum channel simulations via other quantum resources has also been investigated in [11], [12].

In realistic settings, the number of channel uses is necessarily limited, and it is not easy to perform encoding and decoding circuits coherently over a large number of qubits in the near future. Therefore, it is important to characterize how well we can simulate a quantum channel from another with finite resources. The first step in this direction is to consider the one-shot setting. One-shot analysis has recently attracted great interest in classical information theory (see, e.g., [13], [14]) and quantum information theory (see, e.g., [15]–[23]). In one-shot information theory, the smooth max-information of a quantum state [5] and its generalizations [24] are all basic and useful quantities, which have various applications in quantum rate distortion theory as well as the physics of quantum many-body systems.

In this work, we focus on quantum channel simulation in both the one-shot and asymptotic regimes. The central quantity we introduce is the channel's smooth max-information. Our results can be summarized as follows. In Section II, we introduce the task of channel simulation and its related quantities. In Section III, we develop a framework for quantum channel simulation assisted with different codes in the one-shot regime, where one has access only to a single use of the quantum channel. In particular, we characterize the minimum error of channel simulation under the so-called no-signalling (NS) codes [7], [8], which allow the encoder and decoder to share non-local quantum correlations. Such codes are
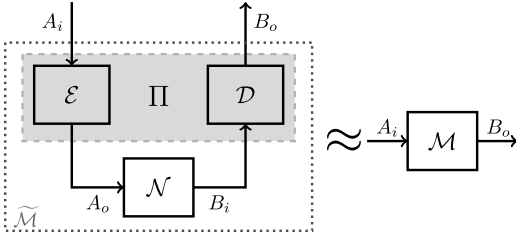
Fig. 1.   General framework of quantum channel simulation.

no-signalling from the sender to the receiver and vice versa, representing the ultimate limit of quantum codes obeying quantum mechanics and providing converse bounds for entanglement-assisted codes. The cost of approximately simulating a channel via noiseless quantum channels under NS-assisted codes can be characterized as an semidefinite program (SDP) [25]. In Section IV, we introduce the channel's smooth max-information, which can be seen as a one-shot generalization of the mutual information of a quantum channel. Notably, this newly introduced entropy has the exact operational interpretation as the one-shot quantum simulation cost under NS-assisted codes. Then we prove its asymptotic equipartition property which directly implies the quantum reverse Shannon theorem in the presence of no-signalling correlations.

In the setting of the entanglement-assisted one-shot capacity of quantum channels, Matthews and Wehner gave a converse bound in terms of the channel's hypothesis testing relative entropy [18]. Moreover, a subset of us recently showed that the activated NS-assisted one-shot capacity is exactly given by the channel's hypothesis testing relative entropy [26] – generalizing the corresponding classical results [13], [27]. This suggests that the operational min- and max-type one-shot analogues of the channel's mutual information are the channel's hypothesis testing relative entropy and the channel's smooth max-information, respectively.

In Section V, as applications, we evaluate the cost of simulating fundamental quantum channels with finite resources. In particular, we derive a linear program to evaluate the finite blocklength simulation cost of quantum depolarizing channels.

II. CHANNEL SIMULATION AND CODES

Let us now formally introduce the task of channel simulation and some basic notations. A quantum channel (quantum operation) $\mathcal{N}_{A \to B}$ is a completely positive (CP) and trace-preserving (TP) linear map from operators acting on a finite-dimensional Hilbert space $\mathcal{H}_A$ to operators acting on a finite-dimensional Hilbert space $\mathcal{H}_B$. For any CPTP map $\mathcal{N}_{A \to B}$, we will frequently use its Choi-Jamiołkowski matrix, which is defined as $J_{\mathcal{N}} := \sum_{i,j=0}^{|A|-1} |i\rangle\langle j| \otimes \mathcal{N}_{A \to B}(|i\rangle\langle j|)$ where $\{|i\rangle\}$ are orthonormal basis in $\mathcal{H}_A$.

The general framework of quantum channel simulation is shown in Fig. 1. Here is how the simulation works. First, Alice performs some pre-processing via an encoder $\mathcal{E}$ on the input system A_i and outputs a quantum system A_o. Then she sends the output system to the shared quantum channel $\mathcal{N}_{A_o \to B_i}$. At Bob's side, he receives a quantum system B_i from the

channel and performs his post-precessing via a decoder $\mathcal{D}$. Finally, Bob outputs a quantum system B_o. The whole process can be considered as a quantum channel from Alice's input A_i to Bob's output B_o, which we denote as $\widetilde{M}_{A_i \to B_o}$. The aim of simulation is to choose the best coding scheme to make the effective channel $\widetilde{M}_{A_i \to B_o}$ as similar to the given target channel $\mathcal{M}$. If we remove the channel $\mathcal{N}$ from Fig. 1, we are left with a map with two inputs A_i, B_i and two outputs A_o, B_o. We denote this map as Π, which generalizes the usual encoding scheme $\mathcal{E}$ and decoding scheme $\mathcal{D}$. Then the effective channel can be written as $\widetilde{M}_{A_i \to B_o} = \Pi_{A_i B_i \to A_o B_o} \circ \mathcal{N}_{A_o \to B_i}$. Note that Π is also known as a bipartite quantum supermap in some literatures [28]–[30] since it maps a quantum channel to another quantum channel.

In particular, a bipartite quantum supermap $\Pi_{A_i B_i \to A_o B_o}$ is A to B no-signalling if A cannot send classical information to B by using Π. That is, for any quantum states ρ_{1,A_i}, ρ_{2,A_i} and linear operator σ_{B_i}, we have $\text{Tr}_{A_o} \Pi(\rho_{1,A_i} \otimes \sigma_{B_i}) = \text{Tr}_{A_o} \Pi(\rho_{2,A_i} \otimes \sigma_{B_i})$, which is equivalent to the semidefinite condition [8] $\text{Tr}_{A_o} J_{\Pi} = \mathbb{1}_{A_i}/|A_i| \otimes \text{Tr}_{A_o A_i} J_{\Pi}$ with J_{Π} being the Choi-Jamiołkowski matrix of Π. Similarly, B to A no-signalling can be characterized by $\text{Tr}_{B_o} J_{\Pi} = \mathbb{1}_{B_i}/|B_i| \otimes \text{Tr}_{B_i B_o} J_{\Pi}$. Note that the bipartite quantum supermap Π in this work is required at least to be B to A no-signalling, which ensures the composition of Π and $\mathcal{N}$ leads to another quantum channel [8], [28]. It is also worth mentioning that there is an isomorphism from bipartite quantum supermaps to bipartite quantum operations, and that the corresponding bipartite quantum operation is required to be no-signalling.

In the task of channel simulation, we say Π is an Ω-assisted code if it can be implemented by local operations with Ω-assistance. In the following, we eliminate Ω for the case of unassisted codes. We write $\Omega = \text{NS}$ and $\Omega = \text{PPT}$ for NS-assisted and positive-partial-transpose-preserving-assisted (PPT-assisted) codes, respectively [31], [32]. In particular,

- an unassisted code reduces to the product of encoder and decoder, i.e., $\Pi = \mathcal{D}_{B_i \to B_o} \mathcal{E}_{A_i \to A_o}$;
- a NS-assisted code corresponds to a bipartite quantum supermap which is no-signalling from Alice to Bob and vice-versa;
- a PPT-assisted code corresponds to a bipartite quantum supermap whose Choi-Jamiołkowski matrix is positive under partial transpose over systems $B_i B_o$.

For any two quantum channels $\mathcal{N}$ and $\mathcal{M}$, the minimum error of simulation from $\mathcal{N}$ to $\mathcal{M}$ under Ω-assisted codes is defined as

$$\omega_{\Omega}(\mathcal{N}, \mathcal{M}) := \frac{1}{2} \inf_{\Pi \in \Omega} \|\Pi \circ \mathcal{N} - \mathcal{M}\|_{\diamond}, \quad (1)$$

where $\|\mathcal{F}\|_{\diamond} := \sup_{k \in \mathbb{N}} \sup_{\|X\|_1 \le 1} \|(\mathcal{F} \otimes \text{id}_k)(X)\|_1$ denotes the diamond norm and $\|X\|_1 := \text{Tr}\sqrt{X^{\dagger}X}$ denotes the Schatten 1-norm. The channel simulation rate from $\mathcal{N}$ to $\mathcal{M}$ under Ω-assisted codes is defined as

$$S_{\Omega}(\mathcal{N}, \mathcal{M}) := \lim_{\varepsilon \to 0} \inf\left\{ \frac{n}{m} \,\Big|\, \omega_{\Omega}(\mathcal{N}^{\otimes n}, \mathcal{M}^{\otimes m}) \le \varepsilon \right\}, \quad (2)$$

where the infimum is taken over ratios $\frac{n}{m}$ with $n, m \in \mathbb{N}$. In this framework of channel simulation, the classical capacity

$C(\mathcal{N})$ and the quantum capacity $Q(\mathcal{N})$ of the channel $\mathcal{N}$ are respectively given by

$$C(\mathcal{N}) = S(\mathcal{N}, \widehat{\mathrm{id}}_2)^{-1} \text{ and } Q(\mathcal{N}) = S(\mathcal{N}, \mathrm{id}_2)^{-1}, \quad (3)$$

where $\widehat{\mathrm{id}}_2$ is the one-bit noiseless channel and id_2 is the one-qubit noiseless channel.

If we consider simulating the given channel $\mathcal{N}$ via an m-dimensional noiseless quantum channel id_m, then the one-shot ε-error quantum simulation cost under Ω-assisted codes is defined as

$$S^{(1)}_{\Omega,\varepsilon}(\mathcal{N}) := \log\inf\left\{m \in \mathbb{N} \mid \omega_\Omega(\mathrm{id}_m, \mathcal{N}) \leq \varepsilon\right\}, \quad (4)$$

where the logarithms in this work are taken in the base two. The asymptotic quantum simulation cost is given by the regularization

$$S_\Omega(\mathcal{N}) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} S^{(1)}_{\Omega,\varepsilon}(\mathcal{N}^{\otimes n}). \quad (5)$$

III. CHANNEL SIMULATION VIA NOISY QUANTUM CHANNELS

Based on the definitions in the above section, we show that the minimum error of simulation under NS-assisted (and PPT-assisted) codes can be given by SDPs. The one-shot ε-error quantum simulation cost under NS-assisted codes can also be given by an SDP. These SDPs can be easily implemented for small blocklength and they also lay the foundation of analysis in the following sections.

Proposition 1 *For any two quantum channels $\mathcal{N}$ and $\mathcal{M}$ with corresponding Choi-Jamiołkowski matrices $J_\mathcal{N}$ and $J_\mathcal{M}$, the minimum error of simulation from $\mathcal{N}$ to $\mathcal{M}$ under NS-assisted codes $\omega_{\mathrm{NS}}(\mathcal{N}, \mathcal{M})$ is given by the following SDP,*

$$\inf \quad \gamma \quad (6a)$$
$$\text{s.t.} \quad \mathrm{Tr}_{B_o} Y_{A_i B_o} \leq \gamma \mathbb{1}_{A_i}, \quad (6b)$$
$$Y_{A_i B_o} \geq J_{\widetilde{\mathcal{M}}} - J_\mathcal{M}, \quad Y_{A_i B_o} \geq 0, \quad (6c)$$
$$J_{\widetilde{\mathcal{M}}} = \mathrm{Tr}_{A_o B_i}(J_\mathcal{N}^T \otimes \mathbb{1}_{A_i B_o}) J_\Pi, \quad (6d)$$
$$J_\Pi \geq 0, \quad \mathrm{Tr}_{A_o B_o} J_\Pi = \mathbb{1}_{A_i B_i}, \quad (\text{CP, TP}) \quad (6e)$$
$$\mathrm{Tr}_{A_o} J_\Pi = \mathbb{1}_{A_i}/|A_i| \otimes \mathrm{Tr}_{A_o A_i} J_\Pi, \quad (A \nrightarrow B) \quad (6f)$$
$$\mathrm{Tr}_{B_o} J_\Pi = \mathbb{1}_{B_i}/|B_i| \otimes \mathrm{Tr}_{B_i B_o} J_\Pi. \quad (B \nrightarrow A) \quad (6g)$$

To obtain $\omega_{\mathrm{NS} \cap \mathrm{PPT}}(\mathcal{N}, \mathcal{M})$, we only need to add the PPT constraint $J_\Pi^{T_{B_i B_o}} \geq 0$, where $T_{B_i B_o}$ denotes the partial transpose over systems $B_i B_o$.

Proof: Note that for any two quantum channels $\mathcal{N}_1, \mathcal{N}_2$ from A to B, the diamond norm of their difference can be expressed as an SDP of the following form [33]:

$$\frac{1}{2}\|\mathcal{N}_1 - \mathcal{N}_2\|_\diamond = \inf\left\{\gamma \mid \mathrm{Tr}_B Y \leq \gamma \mathbb{1}_A,\right.$$
$$\left. Y \geq J_{\mathcal{N}_1} - J_{\mathcal{N}_2}, \ Y \geq 0\right\}, \quad (7)$$

where $J_{\mathcal{N}_1}$ and $J_{\mathcal{N}_2}$ are the corresponding Choi-Jamiołkowski matrices. We denote the Choi-Jamiołkowski matrix of code Π as J_Π. From Lemma 13 in the Appendix, we know

that the Choi-Jamiołkowski matrix of the effective channel $\widetilde{\mathcal{M}} = \Pi \circ \mathcal{N}$ is given by

$$J_{\widetilde{\mathcal{M}}} = \mathrm{Tr}_{A_o B_i}(J_\mathcal{N}^T \otimes \mathbb{1}_{A_i B_o}) J_\Pi. \quad (8)$$

Together with the constraints of the code Π, we have the resulting SDP (6). The constraints in Eq. (6e) represent the CP and TP conditions of the bipartite supermap Π. The constraints in Eqs. (6f) and (6g) represent the no-signalling conditions that A cannot signal to B and B cannot signal to A, respectively. ∎

Corollary 2 *The minimum error to simulate a quantum channel $\mathcal{N}$ via a noiseless quantum channel under NS-assisted codes $\omega_{\mathrm{NS}}(\mathrm{id}_m, \mathcal{N})$ is given by the following SDP,*

$$\inf \quad \gamma \quad (9a)$$
$$\text{s.t.} \quad \mathrm{Tr}_B Y_{AB} \leq \gamma \mathbb{1}_A, \quad (9b)$$
$$Y_{AB} \geq J_{\widetilde{\mathcal{N}}} - J_\mathcal{N}, \quad Y_{AB} \geq 0, \quad (9c)$$
$$J_{\widetilde{\mathcal{N}}} \geq 0, \quad \mathrm{Tr}_B J_{\widetilde{\mathcal{N}}} = \mathbb{1}_A, \quad (9d)$$
$$J_{\widetilde{\mathcal{N}}} \leq \mathbb{1}_A \otimes V_B, \quad \mathrm{Tr}\, V_B = m^2. \quad (9e)$$

To obtain $\omega_{\mathrm{NS} \cap \mathrm{PPT}}(\mathrm{id}_m, \mathcal{N})$, we only need to add the PPT constraint $-\mathbb{1}_A \otimes V_B^T \leq m J_{\widetilde{\mathcal{N}}}^{T_B} \leq \mathbb{1}_A \otimes V_B^T$.

Proof: Denote the Choi-Jamiołkowski matrix of id_m as $J_m = \sum_{i,j=0}^{m-1} |ii\rangle\langle jj|_{A_o B_i}$. Then $J_m^T = J_m$ and the SDP for the minimum error $\omega_{\mathrm{NS}}(\mathrm{id}_m, \mathcal{N})$ can be restated as

$$\inf \quad \gamma \quad (10a)$$
$$\text{s.t.} \quad \mathrm{Tr}_{B_o} Y_{A_i B_o} \leq \gamma \mathbb{1}_{A_i}, \quad (10b)$$
$$Y_{A_i B_o} \geq J_{\widetilde{\mathcal{N}}} - J_\mathcal{N}, \quad Y_{A_i B_o} \geq 0, \quad (10c)$$
$$J_{\widetilde{\mathcal{N}}} = \mathrm{Tr}_{A_o B_i}(J_m \otimes \mathbb{1}_{A_i B_o}) J_\Pi, \quad (10d)$$
$$J_\Pi \geq 0, \quad \mathrm{Tr}_{A_o B_o} J_\Pi = \mathbb{1}_{A_i B_i}, \quad (\text{CP, TP}) \quad (10e)$$
$$\mathrm{Tr}_{A_o} J_\Pi = \mathbb{1}_{A_i}/|A_i| \otimes \mathrm{Tr}_{A_o A_i} J_\Pi, \quad (A \nrightarrow B) \quad (10f)$$
$$\mathrm{Tr}_{B_o} J_\Pi = \mathbb{1}_{B_i}/|B_i| \otimes \mathrm{Tr}_{B_i B_o} J_\Pi, \quad (B \nrightarrow A) \quad (10g)$$

where the dimensions of A_o and B_i are equal to m.

The main idea to do the simplification is to utilize the symmetry of the noiseless quantum channel, i.e., the invariance of J_m under any local unitary $U_{A_o} \otimes \overline{U}_{B_i}$. Suppose J_Π is optimal for SDP (10), we can check that $(U_{A_o} \otimes \overline{U}_{B_i}) J_\Pi (U_{A_o} \otimes \overline{U}_{B_i})^\dagger$ is also optimal. Any convex combination of optimal solutions remains optimal. Thus, without loss of generality we can take [31], [34],

$$J_\Pi = \int dU (U_{A_o} \otimes \overline{U}_{B_i}) J_\Pi (U_{A_o} \otimes \overline{U}_{B_i})^\dagger \quad (11)$$
$$= \frac{J_m}{m} \otimes H_{A_i B_o} + (\mathbb{1}_{A_o B_i} - \frac{J_m}{m}) \otimes K_{A_i B_o}, \quad (12)$$

where the integral is taken over the Haar measure and H, K are operators on system $A_i B_o$.

Note that $J_m \cdot J_m = m J_m$. Taking Eq. (11) into the condition (10d), we obtain that

$$\text{Condition (10d)} \iff J_{\widetilde{\mathcal{N}}} = mH. \quad (13)$$

Taking Eq. (11) into the condition (10e), we have the equivalence

$$\text{Condition (10e)} \iff H \geq 0, \ K \geq 0,$$
$$\text{and } \text{Tr}_{B_o}(H + (m^2 - 1)K) = m\mathbb{1}_{A_i}. \quad (14)$$

Since $J_{\widetilde{\mathcal{N}}}$ is the Choi-Jamiołkowski matrix of the effective channel, we have $\text{Tr}_{B_o} J_{\widetilde{\mathcal{N}}} = \text{Tr}_{B_o} mH = \mathbb{1}_{A_i}$. Combining $\text{Tr}_{B_o}(H + (m^2 - 1)K) = m\mathbb{1}_{A_i}$, we have $\text{Tr}_{B_o} mK = \mathbb{1}_{A_i}$. This implies that the condition (10g) is trivial and

$$\text{Condition (10f)} \iff H + (m^2 - 1)K =$$
$$\mathbb{1}_{A_i}/|A_i| \otimes \text{Tr}_{A_i}(H + (m^2 - 1)K). \quad (15)$$

So far we have the simplified SDP as

$$\inf \quad \gamma \quad (16a)$$
$$\text{s.t.} \quad \text{Tr}_{B_o} Y_{A_i B_o} \leq \gamma \mathbb{1}_{A_i}, \quad (16b)$$
$$Y_{A_i B_o} \geq J_{\widetilde{\mathcal{N}}} - J_{\mathcal{N}}, \ Y_{A_i B_o} \geq 0, \quad (16c)$$
$$J_{\widetilde{\mathcal{N}}} = m H_{A_i B_o}, \quad (16d)$$
$$\text{Tr}_{B_o} J_{\widetilde{\mathcal{N}}} = \mathbb{1}_{A_i} \quad (16e)$$
$$H \geq 0, \ K \geq 0, \ \text{Tr}_{B_o} mK = \mathbb{1}_{A_i}, \quad (16f)$$
$$H + (m^2 - 1)K = \mathbb{1}_{A_i}/|A_i| \otimes$$
$$\text{Tr}_{A_i}(H + (m^2 - 1)K) \quad (16g)$$

Denote $V_{B_o} = m \text{Tr}_{A_i}(H + (m^2 - 1)K)/|A_i|$. By conditions (16g), we have

$$(m^2 - 1)mK = \mathbb{1}_{A_i} \otimes V_{B_o} - mH. \quad (17)$$

Together with the condition (16d), we can eliminate the variables K and H in the above SDP. Finally, replacing the subscript A_i to A and B_o to B, we have the desired SDP (9).

As for the PPT condition $J_{\Pi}^{T_{B_i B_o}} \geq 0$, we have $J_m^{T_{B_i}}/m \otimes H^{T_{B_o}} + (\mathbb{1} - J_m^{T_{B_i}}/m) \otimes K^{T_{B_o}} \geq 0$ from Eq. (11). Note that $J_m^{T_{B_i}}$ is the swap operator and we can decompose it into the sum of two orthogonal positive parts, i.e., $J_m^{T_{B_i}} = J_+ - J_-$ where $J_+ \geq 0, J_- \geq 0$ and $J_+ + J_- = \mathbb{1}$. Then the PPT condition is equivalent to $J_+ \otimes [H^{T_{B_o}} + (m - 1)K^{T_{B_o}}] + J_- \otimes [-H^{T_{B_o}} + (m + 1)K^{T_{B_o}}] \geq 0$. Thus $-(m - 1)K^{T_{B_o}} \leq H^{T_{B_o}} \leq (m + 1)K^{T_{B_o}}$. Combining with Eqs. (16d) and (17), we have $-\mathbb{1}_{A_i} \otimes V_{B_o}^T \leq m J_{\widetilde{\mathcal{N}}}^{T_{B_o}} \leq \mathbb{1}_{A_i} \otimes V_{B_o}^T$. ∎

From the definition of one-shot quantum simulation cost and Corollary 2, we have the following result.

Proposition 3 *For any quantum channel $\mathcal{N}$ and error tolerance $\varepsilon \geq 0$, the one-shot ε-error quantum simulation cost under NS-assisted codes is given by the following SDP,*

$$S_{NS,\varepsilon}^{(1)}(\mathcal{N}) = \log \inf \ \left\lceil \sqrt{\text{Tr} V_B} \right\rceil \quad (18a)$$
$$\text{s.t.} \quad \text{Tr}_B Y_{AB} \leq \varepsilon \mathbb{1}_A, \quad (18b)$$
$$Y_{AB} \geq J_{\widetilde{\mathcal{N}}} - J_{\mathcal{N}}, \ Y_{AB} \geq 0, \quad (18c)$$
$$J_{\widetilde{\mathcal{N}}} \geq 0, \ \text{Tr}_B J_{\widetilde{\mathcal{N}}} = \mathbb{1}_A, \quad (18d)$$
$$J_{\widetilde{\mathcal{N}}} \leq \mathbb{1}_A \otimes V_B. \quad (18e)$$

It is easy to check that $\delta = \log\lceil x \rceil - \log x \in [0, 1]$ for any $x \geq 1$. Thus we can use the least constant $\delta \in [0, 1]$ to adjust the r.h.s. of Eq. (18) to be the logarithm of an integer. That is,

$$S_{NS,\varepsilon}^{(1)}(\mathcal{N}) = \delta +$$
$$\frac{1}{2} \log \inf \quad \text{Tr} V_B \quad (19a)$$
$$\text{s.t.} \quad \text{Tr}_B Y_{AB} \leq \varepsilon \mathbb{1}_A, \quad (19b)$$
$$Y_{AB} \geq J_{\widetilde{\mathcal{N}}} - J_{\mathcal{N}}, \ Y_{AB} \geq 0, \quad (19c)$$
$$J_{\widetilde{\mathcal{N}}} \geq 0, \ \text{Tr}_B J_{\widetilde{\mathcal{N}}} = \mathbb{1}_A, \quad (19d)$$
$$J_{\widetilde{\mathcal{N}}} \leq \mathbb{1}_A \otimes V_B. \quad (19e)$$

Note that the one-shot quantum simulation cost under NS∩PPT-assisted codes is not an SDP, since the objective function m appears in the conditions $\text{Tr} V_B = m^2$ and $-\mathbb{1}_A \otimes V_B^T \leq m J_{\widetilde{\mathcal{N}}}^{T_B} \leq \mathbb{1}_A \otimes V_B^T$ with different powers. We do not see a way to obtain a linear objective function.

It is also worth mentioning that the zero-error quantum simulation cost was studied by Duan and Winter in [8]. The authors show that the zero-error NS-assisted simulation cost is given by the conditional min-entropy of the channel's Choi-Jamiołkowski matrix [8, Theorem 2]. The result we obtained in Proposition 3 is more general and can reduce to the zero-error case by letting $\varepsilon = 0$. More specifically, taking $\varepsilon = 0$ will lead to $Y_{AB} = 0$ and thus $J_{\widetilde{\mathcal{N}}} = J_{\mathcal{N}}$. Then we have

$$S_{NS,0}^{(1)}(\mathcal{N}) = \frac{1}{2} \log \inf \{ \text{Tr} V_B | \ J_{\mathcal{N}} \leq \mathbb{1}_A \otimes V_B \} + \delta, \quad (20)$$

where $\delta \in [0, 1]$ is the least constant such that the r.h.s. is the logarithm of an integer. Since the conditional min-entropy is additive (see [15]), we have

$$S_{NS,0}(\mathcal{N}) := \lim_{n \to \infty} \frac{1}{n} S_{NS,0}^{(1)}(\mathcal{N}^{\otimes n}) \quad (21)$$
$$= \frac{1}{2} \log \inf \{ \text{Tr} V_B | \ J_{\mathcal{N}} \leq \mathbb{1}_A \otimes V_B \}. \quad (22)$$

Remark 1 In the next section, we will see that the zero-error NS-assisted simulation cost could also be considered to be the max-information of the channel's Choi-Jamiołkowski state based on Eqs. (29) and (34).

IV. THE CHANNEL'S MAX-INFORMATION AND ITS ASYMPTOTIC EQUIPARTITION PROPERTY

In this section, we introduce a novel entropy called the channel's smooth max-information and show that it has an operational interpretation regarding the quantum simulation cost of a channel. Furthermore, we prove the asymptotic equipartition property (AEP) of the channel's smooth max-information and explore its close relation to the well-known quantum reverse Shannon theorem (QRST).

Some basic notations will be used in this section. The set of sub-normalized quantum states is denoted as $S_{\leq}(A) := \{\rho \geq 0 \mid \text{Tr} \rho \leq 1\}$. The set of quantum states is denoted as $S_{=}(A) := \{\rho \geq 0 \mid \text{Tr} \rho = 1\}$. We denote ρ_A as the reduced state of ρ_{AB}, i.e. $\rho_A := \text{Tr}_B \rho_{AB}$. The purified distance based on the generalized fidelity is given by [15]

$$P(\rho, \sigma) := \sqrt{1 - F^2(\rho, \sigma)} \quad \text{with} \quad (23)$$
$$F(\rho, \sigma) := \|\rho^{1/2}\sigma^{1/2}\|_1 + \sqrt{(1 - \text{Tr}\rho)(1 - \text{Tr}\sigma)}. \quad (24)$$

We say ρ and σ are ε-close and write $\rho \approx^\varepsilon \sigma$ if and only if $P(\rho, \sigma) \leq \varepsilon$.

A. The Channel's Max-Information

The max-relative entropy of $\rho \in \mathcal{S}_\leq(A)$ with respect to $\sigma \geq 0$ is defined as [35], [36]

$$D_{\max}(\rho \| \sigma) := \inf \left\{ t \mid \rho \leq 2^t \cdot \sigma \right\}, \qquad (25)$$

which is a one-shot generalization of the equantum relative entropy

$$D(\rho \| \sigma) := \operatorname{Tr} \rho (\log \rho - \log \sigma) \qquad (26)$$

if $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$ and $+\infty$ otherwise. The max-information that B has about A for $\rho_{AB} \in \mathcal{S}_\leq(AB)$ is defined as [5]

$$I_{\max}(A : B)_\rho := \inf_{\sigma_B \in \mathcal{S}_=(B)} D_{\max}(\rho_{AB} \| \rho_A \otimes \sigma_B), \qquad (27)$$

which is a one-shot generalization of the quantum mutual information

$$I(A : B)_\rho := \inf_{\sigma_B \in \mathcal{S}_=(B)} D(\rho_{AB} \| \rho_A \otimes \sigma_B). \qquad (28)$$

Definition 4 *For any quantum channel $\mathcal{N}_{A' \to B}$ we define the channel's max-information of $\mathcal{N}$ as*

$$I_{\max}(A : B)_\mathcal{N} := I_{\max}(A : B)_{\mathcal{N}_{A' \to B}(\Phi_{AA'})}, \qquad (29)$$

where $\Phi_{AA'} = \frac{1}{|A|} \sum_{i,j=0}^{|A|-1} |i_A i_{A'}\rangle \langle j_A j_{A'}|$ is the maximally entangled state on AA'.

Remark 2 The following argument shows that this definition does not depend on the input state $\Phi_{AA'}$. That is, for any full rank state $\phi_{A'}$ with a purification $\phi_{AA'} = |A| \sqrt{\phi_A} \Phi_{AA'} \sqrt{\phi_A}$, we have

$$I_{\max}(A : B)_\mathcal{N} = I_{\max}(A : B)_{\mathcal{N}_{A' \to B}(\phi_{AA'})}. \qquad (30)$$

From the definitions (25), (27) and (29), we have

$$I_{\max}(A : B)_\mathcal{N} = \inf \ t \ \text{ s.t.}$$

$$\mathcal{N}_{A' \to B}(\Phi_{AA'}) \leq 2^t \cdot \frac{\mathbb{1}_A}{|A|} \otimes \sigma_B, \ \sigma_B \in \mathcal{S}_=(B). \qquad (31)$$

Since

$$\mathcal{N}_{A' \to B}(\phi_{AA'}) = |A| \cdot \mathcal{N}_{A' \to B}(\sqrt{\phi_A} \Phi_{AA'} \sqrt{\phi_A}) \qquad (32)$$

$$= |A| \cdot \sqrt{\phi_A} \mathcal{N}_{A' \to B}(\Phi_{AA'}) \sqrt{\phi_A}, \qquad (33)$$

the first condition in (31) is equivalent to $\mathcal{N}_{A' \to B}(\phi_{AA'}) \leq 2^t \cdot \phi_A \otimes \sigma_B$, which implies Eq. (30).

Remark 3 Note that the sandwiched Rényi version of the channel's mutual information was previously defined in [37], [38]. Our definition of the channel's max-information is compatible with the more general one and can be recovered in the limit of $\alpha \to \infty$ [29].

Comparing Eqs. (20) and (31), we can write the one-shot zero-error quantum simulation cost as the channel's max-information:

$$S_{\mathrm{NS},0}^{(1)}(\mathcal{N}) = \frac{1}{2} I_{\max}(A : B)_\mathcal{N} + \delta, \qquad (34)$$

where $\delta \in [0, 1]$ is the least constant such that the r.h.s. is the logarithm of an integer. In the following, we show this relation beyond the zero-error case. For this, we define the smoothed version of the channel's max-information.

Definition 5 *For any quantum channel $\mathcal{N}$, we define the channel's smooth max-information as*

$$I_{\max}^\varepsilon(A : B)_\mathcal{N} := \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}} - \mathcal{N}\|_\diamond \leq \varepsilon \\ \widetilde{\mathcal{N}} \in \mathrm{CPTP}(A':B)}} I_{\max}(A : B)_{\widetilde{\mathcal{N}}}, \qquad (35)$$

where $\mathrm{CPTP}(A' : B)$ denotes the set of all CPTP maps from A' to B.

We show that the one-shot ε-error quantum simulation cost is completely characterized by the channel's smooth max-information. This provides the operational meaning of this new entropy.

Theorem 6 *For any quantum channel $\mathcal{N}$ and given error tolerance $\varepsilon \geq 0$, it holds that*

$$S_{\mathrm{NS},\varepsilon}^{(1)}(\mathcal{N}) = \frac{1}{2} I_{\max}^\varepsilon(A : B)_\mathcal{N} + \delta, \qquad (36)$$

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

Proof: We first notice that the constraints in Eq. (19d) $J_{\widetilde{\mathcal{N}}} \geq 0$, $\operatorname{Tr}_B J_{\widetilde{\mathcal{N}}} = \mathbb{1}_A$ uniquely define a CPTP map $\widetilde{\mathcal{N}}$ due to the Choi-Jamiołkowski isomorphism. Applying the SDP of diamond norm in (7), we find

$$S_{\mathrm{NS},\varepsilon}^{(1)}(\mathcal{N}) = \delta + \frac{1}{2} \log \inf \quad \operatorname{Tr} V_B \qquad (37a)$$

$$\text{s.t.} \quad \frac{1}{2}\|\widetilde{\mathcal{N}} - \mathcal{N}\|_\diamond \leq \varepsilon, \qquad (37b)$$

$$\widetilde{\mathcal{N}} \in \mathrm{CPTP}(A' : B), \qquad (37c)$$

$$J_{\widetilde{\mathcal{N}}} \leq \mathbb{1}_A \otimes V_B. \qquad (37d)$$

From Eqs. (31), we know that

$$I_{\max}(A : B)_\mathcal{N} = \log \inf \left\{ \operatorname{Tr} V_B \mid J_\mathcal{N} \leq \mathbb{1}_A \otimes V_B \right\}. \qquad (38)$$

Combining Eqs. (37) and (38), we obtain the desired result. ∎

Remark 4 Note that an alternative way to write Eq. (36) is $S_{\mathrm{NS},\varepsilon}^{(1)}(\mathcal{N}) = \log \left\lceil \sqrt{2^{I_{\max}^\varepsilon(A:B)_\mathcal{N}}} \right\rceil$.

Remark 5 Since $I_{\max}^\varepsilon(A : B)_\mathcal{N}$ is introduced as an entropy of the channel $\mathcal{N}$, it is natural to consider its data-precessing inequality, i.e., $I_{\max}^\varepsilon(A : B)_{\Theta(\mathcal{N})} \leq I_{\max}^\varepsilon(A : B)_\mathcal{N}$ for any quantum superchannel Θ [28]. By the operational meaning above, this is equivalent to $S_{\mathrm{NS},\varepsilon}^{(1)}(\Theta(\mathcal{N})) \leq S_{\mathrm{NS},\varepsilon}^{(1)}(\mathcal{N})$ which can be understood as we need less resources

to simulate a quantum channel with higher noise. Specifically, this relation can be checked by the definition of quantum channel simulation cost. Suppose the simulation cost of $\mathcal{N}$ is given by $\log m$. This implies that there exists an NS-assisted code Π such that $\frac{1}{2}\|\Pi \circ \mathrm{id}_m - \mathcal{N}\|_\diamond \le \varepsilon$. For any quantum superchannel Θ, it can be written [28] as $\Theta(\mathcal{N}) = \mathcal{R}_{B_o E \to \bar{B}} \circ (\mathcal{N}_{A_i \to B_o} \otimes \mathrm{id}_E) \circ \mathcal{F}_{\bar{A} \to A_i E}$ with pre-processing channel $\mathcal{F}_{\bar{A} \to A_i E}$ and post-processing channel $\mathcal{R}_{B_o E \to \bar{B}}$. Note that the diamond norm is multiplicative under tensor product, sub-multiplicative under composition and equals to one for any quantum channels [39]. Then it holds $\frac{1}{2}\|\Theta(\Pi \circ \mathrm{id}_m) - \Theta(\mathcal{N})\|_\diamond \le \frac{1}{2}\|\Pi \circ \mathrm{id}_m - \mathcal{N}\|_\diamond \le \varepsilon$. This shows that we can use the noiseless channel id_m to simulate $\Theta(\mathcal{N})$ under the NS-assisted code $\Theta \circ \Pi$ within ε error. By definition, we have $S_{\mathrm{NS},\varepsilon}^{(1)}(\Theta(\mathcal{N})) \le \log m = S_{\mathrm{NS},\varepsilon}^{(1)}(\mathcal{N})$.

Remark 6 Note that the Choi-Jamiołkowski matrix of a constant channel $\mathcal{M}(\rho) = \sigma, \forall \rho$ is given by $J_\mathcal{M} = \mathbb{1}_A \otimes \sigma_B$. Thus from the perspective of quantum resource theory, the channel's smooth max-information can be written as the "distance" between the given channel $\mathcal{N}$ and the set of constant channels

$$\mathcal{G} := \{ \mathcal{M} \in \mathrm{CPTP}(A : B) \mid \exists \sigma \text{ s.t. } \mathcal{M}(\rho) = \sigma, \forall \rho \}. \quad (39)$$

More specifically, for any quantum channel $\mathcal{N}$, we have

$$I_{\max}^\varepsilon(A : B)_\mathcal{N} = \min_{\mathcal{M} \in \mathcal{G}} D_{\max}^\varepsilon(\mathcal{N}\|\mathcal{M}), \quad (40)$$

where

$$D_{\max}^\varepsilon(\mathcal{N}\|\mathcal{M}) := \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}} - \mathcal{N}\|_\diamond \le \varepsilon \\ \widetilde{\mathcal{N}} \in \mathrm{CPTP}(A':B)}} D_{\max}(\widetilde{\mathcal{N}}\|\mathcal{M}) \quad (41)$$

and $D_{\max}(\mathcal{N}\|\mathcal{M}) := D_{\max}(J_\mathcal{N}\|J_\mathcal{M})$ with Choi-Jamiołkowski matrices $J_\mathcal{N}$, $J_\mathcal{M}$. Since the max-relative entropy is closely related with the robustness [40]—the minimal mixing required to make the given resource useless, we can also define the channel's analogy of robustness as (see also [41])

$$\mathcal{R}_g(\mathcal{N}) := \inf \left\{ t \ge 0 \mid \exists \mathcal{M} \in \mathrm{CPTP}(A : B) \right.$$
$$\left. \text{s.t. } \frac{\mathcal{N} + t\mathcal{M}}{1 + t} \in \mathcal{G} \right\}, \quad (42)$$

and its smoothed version

$$R_g^\varepsilon(\mathcal{N}) := \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}} - \mathcal{N}\|_\diamond \le \varepsilon \\ \widetilde{\mathcal{N}} \in \mathrm{CPTP}(A':B)}} R_g(\widetilde{\mathcal{N}}). \quad (43)$$

Then the channel's smooth max-information can be written as

$$I_{\max}^\varepsilon(A : B)_\mathcal{N} = \log(1 + \mathcal{R}_g^\varepsilon(\mathcal{N})). \quad (44)$$

B. Asymptotic Equipartition Property and the Quantum Reverse Shannon Theorem

In the framework of quantum channel simulation, the quantum capacity is given by the optimal rate of using $\mathcal{N}$ to simulate the qubit noiseless channel id_2, while the channel

simulation cost is given by the optimal rate of using id_2 to simulate the channel $\mathcal{N}$. Thus, it operationally holds that

$$Q_\mathrm{E}(\mathcal{N}) \le Q_\mathrm{NS}(\mathcal{N}) \le S_\mathrm{NS}(\mathcal{N}) \le S_\mathrm{E}(\mathcal{N}), \quad (45)$$

where the above four notations represent entanglement-assisted quantum capacity, NS-assisted quantum capacity, NS-assisted quantum simulation cost and entanglement-assisted quantum simulation cost, respectively. The quantum reverse Shannon theorem [4], [5] states that the quantum simulation cost is equal to its quantum capacity under entanglement-assistance, i.e., $Q_\mathrm{E}(\mathcal{N}) = S_\mathrm{E}(\mathcal{N})$. In the following, the quantum reverse Shannon theorem under NS-assistance means that $Q_\mathrm{NS}(\mathcal{N}) = S_\mathrm{NS}(\mathcal{N})$.

The AEP of the channel's smooth max-information is the claim that

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} I_{\max}^\varepsilon(A : B)_{\mathcal{N}^{\otimes n}} = I(A : B)_\mathcal{N}, \quad (46)$$

where

$$I(A : B)_\mathcal{N} := \max_{\rho_A \in S_=(A)} I(A : B)_{\mathcal{N}_{A' \to B}(\phi_{AA'})} \quad (47)$$

is the mutual information of the quantum channel and $\phi_{AA'}$ is a purification of the state ρ_A. Based on the operational interpretation in Theorem 6, the definition of asymptotic simulation cost in Eq. (5) and the known result that $Q_\mathrm{E}(\mathcal{N}) = \frac{1}{2} I(A : B)_\mathcal{N}$ [2], we have

$$\mathrm{AEP\ (46)} \iff Q_\mathrm{E}(\mathcal{N}) = S_\mathrm{NS}(\mathcal{N}). \quad (48)$$

Hence the QRST implies the AEP for the channel's smooth max-information. On the other hand, we can directly prove the AEP, as is done in Theorem 8. This proof then implies the QRST in the presence of NS correlations.

In the following, we utilize the smooth max-information of a quantum state and its variation:

$$I_{\max}^\varepsilon(A : B)_\rho := \inf_{\widetilde{\rho} \approx^\varepsilon \rho} I_{\max}(A : B)_{\widetilde{\rho}}, \quad (49)$$

$$\widehat{I}_{\max}^\varepsilon(A : B)_\rho := \inf_{\substack{\widetilde{\rho} \approx^\varepsilon \rho \\ \widetilde{\rho}_A = \rho_A}} I_{\max}(A : B)_{\widetilde{\rho}}. \quad (50)$$

The first smooth max-information is most often used in the literature [5], [24]. The second variation naturally appears in our discussion of the channel simulation problem. The restricted smoothing such that the marginal state is fixed comes from the definition of diamond norm where the reference system of the input state is untouched. Using ideas from [42], the following lemma shows that these two quantities are equivalent up to some correction terms. The proof can be found in Appendix B.

Lemma 7 For any quantum state ρ_{AB} and $\varepsilon \in (0, 1)$, it holds

$$\widehat{I}_{\max}^\varepsilon(A : B)_\rho \le I_{\max}^{\varepsilon/6}(A : B)_\rho + g(\varepsilon), \quad (51)$$

with $g(\varepsilon) = \log(1 + 72/\varepsilon^2)$.

Theorem 8 For any quantum channel $\mathcal{N}$ we have the AEP for the channel's smooth max-information:

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} I_{\max}^\varepsilon(A : B)_{\mathcal{N}^{\otimes n}} = I(A : B)_\mathcal{N}. \quad (52)$$

Proof: The proof strategy is as follows. We first use the post-selection technique to show that the channel's smooth max-information is upper bounded by the quantity in Eq. (50). By Eq. (51) we can then use the basic properties of the smooth max-information developed in [5] to show one direction of the proof. The other direction can be proved via the continuity of the mutual information of quantum states.

Consider n uses of the channel $\mathcal{N}$ and let $\omega_{RAA'}^n$ be the purification of the de Finetti state $\omega_{AA'}^n = \int \sigma_{AA'}^{\otimes n} d(\sigma_{AA'})$ with pure states $\sigma_{AA'} = |\sigma\rangle\langle\sigma|_{AA'}$ and $d(\cdot)$ the measure on the normalized pure states induced by the Haar measure. Furthermore we can assume without loss of generality that $|R| \leq (n+1)^{|A|^2-1}$ [43]. Note that $\omega_{A'}^n$ is a full rank state. We have the following chain of inequalities

$$I_{\max}^\varepsilon(A:B)_{\mathcal{N}^{\otimes n}}$$
$$= \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n}\|_\diamond \leq \varepsilon \\ \widetilde{\mathcal{N}}^n \in \text{CPTP}(A'^n:B^n)}} I_{\max}(RA:B)_{\widetilde{\mathcal{N}}^n(\omega_{RAA'}^n)} \tag{53a}$$

$$\leq \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n}\|_\diamond \leq \varepsilon \\ \widetilde{\mathcal{N}}^n \in \text{Perm}(A'^n:B^n)}} I_{\max}(RA:B)_{\widetilde{\mathcal{N}}^n(\omega_{RAA'}^n)} \tag{53b}$$

$$\leq \inf_{\substack{\frac{1}{2}\|(\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n})(\omega_{RAA'}^n)\|_1 \leq \varepsilon_1 \\ \widetilde{\mathcal{N}}^n \in \text{Perm}(A'^n:B^n)}} I_{\max}(RA:B)_{\widetilde{\mathcal{N}}^n(\omega_{RAA'}^n)} \tag{53c}$$

$$\leq \inf_{\substack{\widetilde{\mathcal{N}}^n(\omega_{RAA'}^n) \approx_{\varepsilon_2} \mathcal{N}^{\otimes n}(\omega_{RAA'}^n) \\ \widetilde{\mathcal{N}}^n \in \text{Perm}(A'^n:B^n)}} I_{\max}(RA:B)_{\widetilde{\mathcal{N}}^n(\omega_{RAA'}^n)}, \tag{53d}$$

where $\varepsilon_1 = \varepsilon(n+1)^{-(|A'|^2-1)}$ and we take $\varepsilon_2 = \varepsilon_1$. Note that the discussion works for any $0 < \varepsilon_2 \leq \varepsilon_1$. In (53a), we choose $\omega_{RAA'}^n$ as the input state of the channel's max-information (30). In (53b), we restrict the channel $\widetilde{\mathcal{N}}^n$ to be permutation invariant, where $\text{Perm}(A'^n : B^n) := \{\mathcal{N}^n \in \text{CPTP}(A'^n : B^n) \mid \pi_B^n \circ \mathcal{N}^n \circ \pi_{A'}^n = \mathcal{N}, \text{ for all permutation } \pi^n\}$ denotes the set of all permutation invariant channels. In (53c), we use post-selection technique (see [5, Prop. D.4]), which relaxes the diamond norm to the trace norm. In (53d), we replace the trace norm with the purified distance due to the inequality $\frac{1}{2}\|\rho - \sigma\|_1 \leq P(\rho, \sigma)$.

Exploiting the permutation invariance of $\mathcal{N}^{\otimes n}$, we know that the optimal solution of (53d) can be still taken at a permutation invariant channel even if we relax the set $\text{Perm}(A'^n : B^n)$ to all CPTP maps, which is then equivalent to optimize over all quantum states with marginal ω_{RA}^n and ε_2-close to $\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)$. Specifically, from Lemma 11 in Appendix B, we know that the optimization in (53d) is equivalent to

$$\inf_{\sigma_{RAB}^n \in K} I_{\max}(RA:B)_{\sigma_{RAB}^n} \quad \text{with} \tag{54}$$

$$K := \left\{ \sigma_{RAB}^n \,\middle|\, \sigma_{RAB}^n \approx_{\varepsilon_2} \mathcal{N}^{\otimes n}(\omega_{RAA'}^n), \sigma_{RA}^n = \omega_{RA}^n \right\},$$

which is exactly the definition of $\widehat{I}_{\max}^{\varepsilon_2}(RA:B)_{\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)}$. Thus we have

$$I_{\max}^\varepsilon(A:B)_{\mathcal{N}^{\otimes n}} \leq \widehat{I}_{\max}^{\varepsilon_2}(RA:B)_{\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)}. \tag{55}$$

From Eq. (51), denote $\varepsilon_3 = \varepsilon_2/6$, we have

$$I_{\max}^\varepsilon(A:B)_{\mathcal{N}^{\otimes n}} \leq I_{\max}^{\varepsilon_3}(RA:B)_{\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)} + g(\varepsilon_2). \tag{56}$$

Then we can use some known properties of the smooth max-information from [5], [24], which leads to

$$I_{\max}^{\varepsilon_3}(RA:B)_{\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)}$$
$$\leq I_{\max}^{\varepsilon_4}(B:RA)_{\mathcal{N}^{\otimes n}(\omega_{RAA'}^n)} + f(\varepsilon_4) \tag{57a}$$

$$\leq I_{\max}^{\varepsilon_4}(B:A)_{\mathcal{N}^{\otimes n}(\omega_{AA'}^n)} + 2\log|R| + f(\varepsilon_4) \tag{57b}$$

$$= I_{\max}^{\varepsilon_4}(B:A)_{\mathcal{N}^{\otimes n}(\sum_{i\in I} p_i(\sigma_{AA'}^i)^{\otimes n})} + 2\log|R| + f(\varepsilon_4) \tag{57c}$$

$$\leq \max_{\sigma_{AA'}^i} I_{\max}^{\varepsilon_4}(B:A)_{\mathcal{N}^{\otimes n}((\sigma_{AA'}^i)^{\otimes n})} + \log|I| + 2\log|R| + f(\varepsilon_4) \tag{57d}$$

$$\leq \max_{\sigma_{AA'}} I_{\max}^{\varepsilon_4}(B:A)_{\mathcal{N}(\sigma_{AA'})^{\otimes n}} + \log|I| + 2\log|R| + f(\varepsilon_4), \tag{57e}$$

where $\varepsilon_4 = \varepsilon_3/2$, $f(\varepsilon) = \log(\frac{1}{1-\sqrt{1-\varepsilon^2}} + \frac{1}{1-\varepsilon})$ and $|I| = (n+1)^{2|A||A'|-2}$. In (57a), we swap the system order according to [24, Corollary 5]. In (57b), we get rid of purification system R according to [5, Lemma B.12]. In (57c), we express the integral $\omega_{AA'}^n = \int \sigma_{AA'}^{\otimes n} d(\sigma_{AA'})$ into convex combination of finite number of operators according to [5, Corollary D.6]. In (57d), we use the quasi-convexity of the smooth max-information [5, Lemma B.21]. In (57e), we relax the maximization to all pure states $\sigma_{AA'}$.

Combining Eqs. (56), (57e) and the AEP for the smooth max-information from [5, Lemma B.24], we get

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} I_{\max}^\varepsilon(A:B)_{\mathcal{N}^{\otimes n}}$$
$$\leq \max_{\sigma_{AA'}} I(A:B)_{\mathcal{N}(\sigma_{AA'})} = I(A:B)_{\mathcal{N}}. \tag{58}$$

On the other hand, suppose the optimal solution of $I(A:B)_{\mathcal{N}}$ is taken at $\rho_{A'}$ with a purification $\phi_{AA'}$. Since we can always find a full rank state that is arbitary close to $\rho_{A'}$, thus it gives the mutual information arbitary close to $I(A:B)_{\mathcal{N}}$ due to the continuty. In the following, we can assume that $\rho_{A'}$ is of full rank without loss of generality and have the chain of inequalities in (60), as shown at the bottom of the next page, where $h_2(\cdot)$ is the binary entropy. In (60b), we use the fact that max-relative entropy is never smaller than the relative entropy [35]. The line (60c) follows from the definition of the mutual information of a quantum state. The line (60d) follows from the continuity of quantum mutual information in Lemma 12 (Appendix B). The line (60e) follows from the additivity of quantum mutual information. The line (60f) follows from the assumption that $\phi_{AA'}$ is the optimizer of $I(A:B)_{\mathcal{N}}$. Finally, we have

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} I_{\max}^\varepsilon(A:B)_{\mathcal{N}^{\otimes n}} \geq I(A:B)_{\mathcal{N}}. \tag{59}$$

Combining Eqs. (58) and (59), we conclude the claim. ∎

After this work there is an alternative proof of Eq. (58) given by Gour and Wilde in [29]. Their proof uses the sandwiched Rényi mutual information and its relation with the max-information, different from the post-selection technique we use in this work.
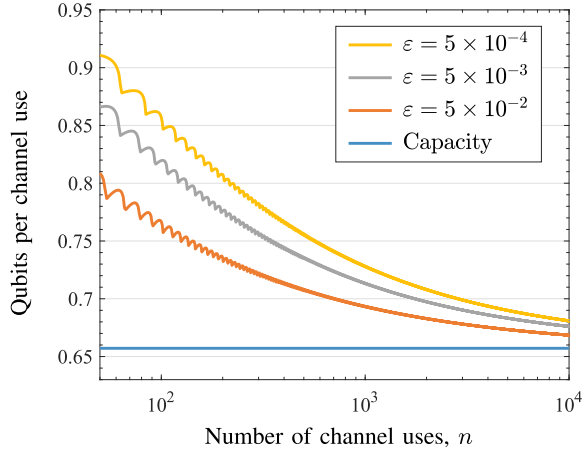
Fig. 2. Exact value by the linear program (80) of the average simulation cost for three different error tolerances $\varepsilon \in \{5 \times 10^{-4}, 5 \times 10^{-3}, 5 \times 10^{-2}\}$ and the qubit depolarizing channel with failure probability $p = 0.15$. The lowest line marks the entanglement-assisted quantum capacity of the channel (roughly 0.657 qubits per channel use).

V. EXAMPLES

In this section, we apply our results to some basic channels. For classical channels, the one-shot ε-error quantum simulation cost can be given by a linear program as shown in Eq. (76) (Appendix A). Using the symmetry of the quantum depolarizing channel, we can also simplify its n-shot simulation cost as a linear program. Moreover, the zero-error simulation cost of various channels can be solved analytically.

Example 1. The quantum depolarizing channel is given by $\mathcal{D}_p(\rho) = (1 - p)\rho + p \cdot \frac{1}{d}$ with dimension d. Its Choi-Jamiołkowski matrix $J_{\mathcal{D}_p}$ commutes with any local unitary $U \otimes \overline{U}$ and $J_{\mathcal{D}_p}^{\otimes n}$ is invariant under any permutation of the tensor factors. Exploiting these symmetries, we can simplify the SDP (18) for $\mathcal{D}_p^{\otimes n}$ to a linear program (80) in Appendix A. Numerical implementation is shown in Fig. 2. We can see that as the number of channel uses n increases, the average quantum simulation cost will approach to its entanglement-assisted quantum capacity [44], i.e., half of the quantum mutual information of the channel.

Recall that the primal and dual SDPs of the zero-error simulation cost are given by [8]

Primal: $S_{\text{NS},0}(\mathcal{N}) = \frac{1}{2} \log \inf \left\{ \text{Tr } V_B \mid J_{\mathcal{N}} \leq \mathbb{1}_A \otimes V_B \right\},$

Dual: $S_{\text{NS},0}(\mathcal{N}) = \frac{1}{2} \log \sup \left\{ \text{Tr } J_{\mathcal{N}} X_{AB} \mid \right.$
$$\left. \text{Tr}_A X_{AB} \leq \mathbb{1}_B, X_{AB} \geq 0 \right\}.$$

We study some fundamental channels and show their analytical solutions by explicitly constructing feasible solutions in both primal and dual problems, respectively. Using the weak duality, we can argue that the feasible solutions we construct are optimal.

Example 2. The quantum depolarizing channel is $\mathcal{D}_p(\rho) = (1 - p)\rho + p \cdot \frac{1}{d}$ with dimension d. Taking

$$V_B = (d(1-p) + \frac{p}{d})\mathbb{1}_B, \text{ and } X_{AB} = \sum_{i,j=0}^{d-1} |ii\rangle\langle jj|, \quad (61)$$

in the primal and dual problems respectively, we can verify that they are feasible solutions. Thus, we have

$$\frac{1}{2} \log(d^2(1 - p) + p) = \frac{1}{2} \log \text{Tr } J_{\mathcal{D}_p} X_{AB} \tag{62}$$

$$\leq S_{\text{NS},0}(\mathcal{D}_p) \tag{63}$$

$$\leq \frac{1}{2} \log \text{Tr } V_B \tag{64}$$

$$= \frac{1}{2} \log(d^2(1 - p) + p). \tag{65}$$

We find that

$$S_{\text{NS},0}(\mathcal{D}_p) = \frac{1}{2} \log(d^2(1 - p) + p). \tag{66}$$

Example 3. The amplitude damping channel is $\mathcal{N}_r(\rho) = \sum_{i=0}^{1} E_i \rho E_i^{\dagger}$ with $E_0 = |0\rangle\langle 0| + \sqrt{1 - r}|1\rangle\langle 1|$, $E_1 = \sqrt{r}|0\rangle\langle 1|$ and $0 \leq r \leq 1$. The optimal solutions are given by

$$V_B = (1 + \sqrt{1 - r})|0\rangle\langle 0| + (\sqrt{1 - r} + 1 - r)|1\rangle\langle 1| \tag{67}$$

$$\text{and} \quad X_{AB} = (|00\rangle + |11\rangle)(\langle 00| + \langle 11|). \tag{68}$$

We find that

$$S_{\text{NS},0}(\mathcal{N}_r) = \frac{1}{2} \log(2(1 + \sqrt{1 - r}) - r). \tag{69}$$

Example 4. The dephasing channel is $\mathcal{Z}_p(\rho) = (1 - p)\rho + pZ\rho Z$ with $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$. The optimal solutions are given by

$$V_B = (|2p - 1| + 1)\mathbb{1}_B \quad \text{and} \tag{70}$$

$$X_{AB} = (|00\rangle + |11\rangle)(\langle 00| + \langle 11|). \tag{71}$$

$$I_{\max}^{\varepsilon}(A : B)_{\mathcal{N}^{\otimes n}} = \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n}\|_{\diamond} \leq \varepsilon \\ \widetilde{\mathcal{N}}^n \in \text{CPTP}(A'^n : B^n)}} \inf_{\sigma_B^n \in \mathcal{S}_=(B^{\otimes n})} D_{\max}(\widetilde{\mathcal{N}}_{A' \to B}^n(\phi_{AA'}^{\otimes n})\|\phi_A^{\otimes n} \otimes \sigma_B^n) \tag{60a}$$

$$\geq \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n}\|_{\diamond} \leq \varepsilon \\ \widetilde{\mathcal{N}}^n \in \text{CPTP}(A'^n : B^n)}} \inf_{\sigma_B^n \in \mathcal{S}_=(B^{\otimes n})} D(\widetilde{\mathcal{N}}_{A' \to B}^n(\phi_{AA'}^{\otimes n})\|\phi_A^{\otimes n} \otimes \sigma_B^n) \tag{60b}$$

$$= \inf_{\substack{\frac{1}{2}\|\widetilde{\mathcal{N}}^n - \mathcal{N}^{\otimes n}\|_{\diamond} \leq \varepsilon \\ \widetilde{\mathcal{N}}^n \in \text{CPTP}(A'^n : B^n)}} I(A : B)_{\widetilde{\mathcal{N}}_{A' \to B}^n(\phi_{AA'}^{\otimes n})} \tag{60c}$$

$$\geq I(A : B)_{\mathcal{N}_{A' \to B}^{\otimes n}(\phi_{AA'}^{\otimes n})} - (2n\varepsilon \log |A| + (1 + \varepsilon)h_2(\varepsilon/(1 + \varepsilon))) \tag{60d}$$

$$= nI(A : B)_{\mathcal{N}_{A' \to B}(\phi_{AA'})} - (2n\varepsilon \log |A| + (1 + \varepsilon)h_2(\varepsilon/(1 + \varepsilon))) \tag{60e}$$

$$= nI(A : B)_{\mathcal{N}} - (2n\varepsilon \log |A| + (1 + \varepsilon)h_2(\varepsilon/(1 + \varepsilon))) \tag{60f}$$
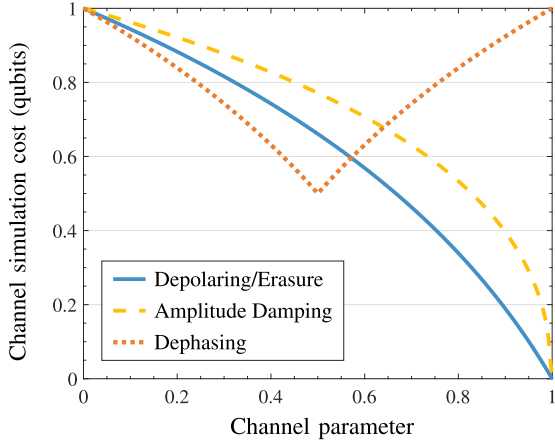
Fig. 3. The zero-error NS-assisted channel simulation cost of the qubit depolarizing channel/qubit erasure channel, qubit amplitude damping channel and qubit dephasing channel as a function of each channel's noise parameter.

We find that

$$S_{\text{NS},0}(\mathcal{Z}_p) = \frac{1}{2} \log(|4p - 2| + 2). \tag{72}$$

Example 5. The quantum erasure channel is $\mathcal{E}_p(\rho) = (1 - p)\rho + p|e\rangle\langle e|$ with $|e\rangle$ orthogonal to the input Hilbert space. The optimal solutions are given by

$$V_B = d(1 - p) \sum_{i,j=0}^{d-1} |i\rangle\langle i| + p|d\rangle\langle d| \quad \text{and} \tag{73}$$

$$X_{AB} = \sum_{i,j=0}^{d-1} |ii\rangle\langle jj| + \frac{1}{d} \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes |d\rangle\langle d|. \tag{74}$$

We find that

$$S_{\text{NS},0}(\mathcal{E}_p) = \frac{1}{2} \log(d^2(1 - p) + p). \tag{75}$$

Finally, we plot the zero-error NS-assisted channel simulation cost of these four channels as a function of their noise parameters in the following Figure 3. The figure is plotted for the qubit case, i.e, $d = 2$. Note that the quantum depolarizing channel and the quantum erasure channel have exactly the same rate given by Eqs. (66) and (75).

VI. DISCUSSION

Since the entanglement-assisted capacity allows a single-letter characterization, it is natural to consider a second-order refinement thereof. A second-order expansion of an achievable rate was established in [45] but no matching second-order converse bound is known. Our one-shot NS-assisted quantum simulation cost and the channel's smooth max-information may provide some insights in this direction.

Suppose a quantum channel $\mathcal{N}$ can be used to simulate a noiseless channel id_{m_1} with dimension m_1 and on the other hand it requires a noiseless channel id_{m_2} with dimension m_2 to simulate itself, then we necessarily have $m_2 \geq m_1$ by the definition of simulation. This means that the simulation cost of a channel operationally provides a converse bound for

its channel capacity. However this approach does not provide a tighter bound than the NS-assisted capacity in the one-shot and asymptotic setting (see, e.g., [19], [27], [31]).

From Eq. (40), the AEP of the channel's smooth max-information can be equivalently written as

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \inf_{\mathcal{M}^n \in \mathcal{G}_n} D_{\max}^\varepsilon(\mathcal{N}^{\otimes n} \| \mathcal{M}^n) = \inf_{\mathcal{M} \in \mathcal{G}} D(\mathcal{N} \| \mathcal{M}),$$

where $\mathcal{G}_n$ represents the set of constant channels from A'^n to B^n and the channel divergence is defined as [38], [46] $D(\mathcal{N} \| \mathcal{M}) := \max_{\rho_A} D(\mathcal{N}_{A' \to B}(\phi_{AA'}) \| \mathcal{M}_{A' \to B}(\phi_{AA'}))$, with $\phi_{AA'}$ a purification of ρ_A. An interesting question is to consider the channel's AEP beyond the set of constant channels $\mathcal{G}_n$, such as the singleton $\{\mathcal{M}^{\otimes n}\}$ [47]. Can we obtain a channel's generalization of quantum Stein's lemma [48]? A partial progress to this problem regarding classical-quantum channels has been recently given in [49].

APPENDIX A
LINEAR PROGRAMS

For any classical channel $\mathcal{N}(y|x)$, the SDP (18) will naturally reduce to a linear program. Specifically, its one-shot simulation cost is given by a linear program,

$$S_{\text{NS},\varepsilon}^{(1)}(\mathcal{N}) = \log \inf \left\lceil \sqrt{\sum_y V_y} \right\rceil \quad \text{s.t.} \tag{76a}$$

$$Y_{xy} \geq \widetilde{\mathcal{N}}(y|x) - \mathcal{N}(y|x), \, Y_{xy} \geq 0, \forall x, y, \tag{76b}$$

$$\widetilde{\mathcal{N}}(y|x) \geq 0, \forall x, y, \quad \sum_y \widetilde{\mathcal{N}}(y|x) = 1, \forall x, \tag{76c}$$

$$\widetilde{\mathcal{N}}(y|x) \leq V_y, \forall x, y, \quad \sum_y Y_{xy} \leq \varepsilon, \forall x. \tag{76d}$$

For the quantum depolarizing channel $\mathcal{D}_p(\rho) = (1 - p)\rho + p \cdot \frac{\mathbb{1}}{d}$, its Choi-Jamiołkowski matrix is given by $J_{\mathcal{D}_p} = q_1\Phi_d + q_2\Phi_d^\perp$ where $q_1 = d(1 - p) + \frac{p}{d}$, $q_2 = \frac{p}{d}$ and Φ_d is the maximally entangled state with dimension d, $\Phi_d^\perp = \mathbb{1} - \Phi_d$. Then we have

$$J_{\mathcal{D}_p}^{\otimes n} = \sum_{k=0}^n p_k P_k^n(\Phi_d, \Phi_d^\perp) \quad \text{with} \quad p_k = q_1^k q_2^{n-k}, \tag{77}$$

and $P_k^n(\Phi_d, \Phi_d^\perp)$ denotes the summation of n-fold tensor products of Φ_d and $\Phi_d^\perp$ with exactly k factors of Φ_d. For example, $P_1^3(\Phi_d, \Phi_d^\perp) = \Phi_d^\perp \otimes \Phi_d^\perp \otimes \Phi_d + \Phi_d^\perp \otimes \Phi_d \otimes \Phi_d^\perp + \Phi_d \otimes \Phi_d^\perp \otimes \Phi_d^\perp$. Due to the symmetries of $J_{\mathcal{D}_p}^{\otimes n}$, we can take the optimal solution in SDP (18) in form of

$$J_{\widetilde{\mathcal{N}}^n} = \sum_{k=0}^n r_k P_k^n(\Phi_d, \Phi_d^\perp), \quad Y = \sum_{k=0}^n y_k P_k^n(\Phi_d, \Phi_d^\perp), \tag{78}$$

$$\text{and} \quad V = s\mathbb{1}_d^{\otimes n}. \tag{79}$$

Then we have the LP as follows,

$$S_{\text{NS},\varepsilon}^{(1)}(\mathcal{D}_p^{\otimes n}) = \log \inf \left\lceil \sqrt{d^n \cdot s} \right\rceil \quad \text{s.t.} \tag{80a}$$

$$y_k - r_k + p_k \geq 0, \, y_k \geq 0, \, 0 \leq r_k \leq s, \forall k \tag{80b}$$

$$\sum_{k=0}^n \binom{n}{k} (\frac{1}{d})^k (d - \frac{1}{d})^{n-k} r_k = 1, \tag{80c}$$

$$\sum_{k=0}^n \binom{n}{k} (\frac{1}{d})^k (d - \frac{1}{d})^{n-k} y_k \leq \varepsilon. \tag{80d}$$

APPENDIX B
TECHNICAL LEMMAS

Lemma 9 *For any quantum state ρ_{AB} and $\varepsilon \in (0, 1)$, it holds*

$$\widehat{I}_{\max}^{\varepsilon}(A : B)_\rho \leq I_{\max}^{\varepsilon/6}(A : B)_\rho + g(\varepsilon), \qquad (81)$$

with $g(\varepsilon) = \log(1 + 72/\varepsilon^2)$.

Proof: Recall the definitions of the smooth max-information for quantum states:

$$\widehat{I}_{\max}^{\varepsilon}(A : B)_\rho = \inf_{\substack{\widetilde{\rho} \approx^\varepsilon \rho \\ \widetilde{\rho}_A = \rho_A}} \inf_{\sigma_B} D_{\max}(\widetilde{\rho}_{AB} \| \rho_A \otimes \sigma_B), \qquad (82)$$

$$I_{\max}^{\varepsilon}(A : B)_\rho = \inf_{\widetilde{\rho} \approx^\varepsilon \rho} \inf_{\sigma_B} D_{\max}(\widetilde{\rho}_{AB} \| \widetilde{\rho}_A \otimes \sigma_B). \qquad (83)$$

In [42] the authors also discuss the variation

$$\widetilde{I}_{\max}^{\varepsilon}(A : B)_\rho = \inf_{\widetilde{\rho} \approx^\varepsilon \rho} \inf_{\sigma_B} D_{\max}(\widetilde{\rho}_{AB} \| \rho_A \otimes \sigma_B), \qquad (84)$$

where the marginal state in the second term is fixed to be ρ_A. It was shown in [42] that for any quantum state ρ_{AB} and $\varepsilon \in (0, 1)$ we have

$$\widehat{I}_{\max}^{\varepsilon}(A : B)_\rho \leq \widetilde{I}_{\max}^{\varepsilon/3}(A : B)_\rho + \log(1 + 72/\varepsilon^2). \qquad (85)$$

To show the result as Eq. (81), we only need to prove

$$\widetilde{I}_{\max}^{\varepsilon}(A : B)_\rho \leq I_{\max}^{\varepsilon/2}(A : B)_\rho. \qquad (86)$$

Denote $I_{\max}^{\varepsilon/2}(A : B)_\rho = \lambda$ and suppose the optimal solution is taken at $\overline{\rho}_{AB}$ and σ_B. Let

$$\widetilde{\rho}_{AB} = \rho_A^{\frac{1}{2}} V_A \overline{\rho}_A^{-\frac{1}{2}} \overline{\rho}_{AB} \overline{\rho}_A^{-\frac{1}{2}} V_A^\dagger \rho_A^{\frac{1}{2}}, \qquad (87)$$

where V_A is the unitary such that $F(\rho_A, \overline{\rho}_A) = \operatorname{Tr} \rho_A^{\frac{1}{2}} \overline{\rho}_A^{\frac{1}{2}} V_A$. By direct calculation, we have

$$\widetilde{\rho}_{AB} = \rho_A^{\frac{1}{2}} V_A \overline{\rho}_A^{-\frac{1}{2}} \overline{\rho}_{AB} \overline{\rho}_A^{-\frac{1}{2}} V_A^\dagger \rho_A^{\frac{1}{2}} \qquad (88)$$

$$\leq 2^\lambda \cdot \rho_A^{\frac{1}{2}} V_A P_{\overline{\rho}_A} V_A^\dagger \rho_A^{\frac{1}{2}} \otimes \sigma_B \qquad (89)$$

$$\leq 2^\lambda \rho_A \otimes \sigma_B, \qquad (90)$$

where $P_{\overline{\rho}_A}$ is the projector on the support of $\overline{\rho}_A$. Then Eq. (86) follows as soon as we show $\widetilde{\rho}_{AB} \approx^\varepsilon \rho_{AB}$. This is actually shown in the final steps of the proof for Theorem 2 in [42]. We repeat these steps here for completeness. Deonte the purification of $\overline{\rho}_{AB}$ as $|\overline{\rho}_{ABC}\rangle = \overline{\rho}_A^{\frac{1}{2}} |\Phi\rangle_{A:BC}$ where $|\Phi\rangle_{A:BC}$ denotes the non-normalized maximally entangled pure state in the cut $A : BC$. Then $|\widetilde{\rho}_{ABC}\rangle = \rho_A^{\frac{1}{2}} V_A \overline{\rho}_A^{-\frac{1}{2}} |\overline{\rho}_{ABC}\rangle$ is a purification of $\widetilde{\rho}_{AB}$. We have

$$F^2(\widetilde{\rho}_{AB}, \overline{\rho}_{AB}) \geq F^2(\widetilde{\rho}_{ABC}, \overline{\rho}_{ABC}) \qquad (91)$$

$$= |\langle \overline{\rho}_{ABC} | \widetilde{\rho}_{ABC} \rangle|^2 \qquad (92)$$

$$= |\langle \Phi_{A:BC} | \overline{\rho}_A^{\frac{1}{2}} \rho_A^{\frac{1}{2}} V_A \overline{\rho}_A^{-\frac{1}{2}} \overline{\rho}_A^{\frac{1}{2}} |\Phi\rangle_{A:BC}\rangle|^2 \qquad (93)$$

$$= \left| \operatorname{Tr} \overline{\rho}_A^{\frac{1}{2}} \rho_A^{\frac{1}{2}} V_A \right|^2 = F^2(\rho_A, \overline{\rho}_A), \qquad (94)$$

which implies $P(\widetilde{\rho}_{AB}, \overline{\rho}_{AB}) \leq P(\rho_A, \overline{\rho}_A)$. Thus it holds

$$P(\widetilde{\rho}_{AB}, \rho_{AB}) \leq P(\widetilde{\rho}_{AB}, \overline{\rho}_{AB}) + P(\overline{\rho}_{AB}, \rho_{AB}) \qquad (95)$$

$$\leq P(\rho_A, \overline{\rho}_A) + P(\overline{\rho}_{AB}, \rho_{AB}) \qquad (96)$$

$$\leq 2P(\overline{\rho}_{AB}, \rho_{AB}) \qquad (97)$$

$$\leq \varepsilon, \qquad (98)$$

which completes the proof. ∎

Lemma 10 *For any pure state $\phi_{AA'}$ and quantum state ρ_{AB} such that $\phi_A = \rho_A$, the following two sets are the same,*

$$\{\mathcal{N}_{A' \to B}(\phi_{AA'}) \approx^\varepsilon \rho_{AB} \mid \mathcal{N} \in \operatorname{CPTP}(A' : B)\}$$
$$= \{\sigma_{AB} \approx^\varepsilon \rho_{AB} \mid \sigma_A = \rho_A\}. \quad (99)$$

Proof: Denote the L.H.S and R.H.S as $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. It is clear that $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and we now show the other direction. For any quantum state $\sigma_{AB} \in \mathcal{S}_2$, denote $\overline{\sigma}_{AB} = \sigma_A^{-1/2} \sigma_{AB} \sigma_A^{-1/2}$. Then, we have $\overline{\sigma}_{AB} \geq 0$ and $\overline{\sigma}_A = \mathbb{1}_A$. From the Choi-Jamiołkowski isomorphism, we know that there exists a CPTP map $\mathcal{N}_{A' \to B}$ such that $\overline{\sigma}_{AB} = \mathcal{N}_{A' \to B}(\Phi_{AA'})$, where $\Phi_{AA'}$ denotes the un-normalized maximally entangled state. Thus, we get $\sigma_{AB} = \mathcal{N}_{A' \to B}(\sigma_A^{1/2} \Phi_{AA'} \sigma_A^{1/2})$. Denoting $\psi_{AA'} = \sigma_A^{1/2} \Phi_{AA'} \sigma_A^{1/2}$, we have that $\psi_{AA'}$ is a purification of σ_A and since $\sigma_A = \rho_A = \phi_A$ we get that $\phi_{AA'}$ is also a purification of σ_A. Due to Uhlmann's theorem [50], there exists a unitary U on the system A' such that $\psi_{AA'} = \mathcal{U}(\phi_{AA'})$ with $\mathcal{U}(\cdot) = U \cdot U^\dagger$. Hence, we find $\sigma_{AB} = \mathcal{N} \circ \mathcal{U}(\phi_{AA'}) \in \mathcal{S}_1$. This completes the proof. ∎

Lemma 11 *For any quantum channel $\mathcal{N}_{A' \to B}$ and $\varepsilon \in (0, 1)$, denote $k_i := \inf_{\sigma_{RAB}^n \in K_i} I_{\max}(RA : B)_{\sigma_{RAB}^n}$ as the optimal values, where $i = 1, 2$ and the feasible sets*

$$K_1 := \left\{ \sigma_{RAB}^n \middle| \sigma_{RAB}^n \approx^\varepsilon \mathcal{N}_{A' \to B}^{\otimes n}(\omega_{RAA'}^n), \sigma_{RA}^n = \omega_{RA}^n \right\},$$

$$K_2 := \Big\{ \sigma_{RAB}^n \,\Big|\, \sigma_{RAB}^n = \widetilde{\mathcal{N}}_{A' \to B}^n(\omega_{RAA'}^n) \approx^\varepsilon$$
$$\mathcal{N}_{A' \to B}^{\otimes n}(\omega_{RAA'}^n), \widetilde{\mathcal{N}}^n \in \operatorname{Perm}(A'^n : B^n) \Big\}.$$

where $\omega_{RAA'}^n$ is the purification of the de Finetti state $\omega_{AA'}^n = \int \sigma_{AA'}^{\otimes n} d(\sigma_{AA'})$ with pure states $\sigma_{AA'} = |\sigma\rangle\langle\sigma|_{AA'}$ and $d(\cdot)$ the measure on the normalized pure states induced by the Haar measure. Then $k_1 = k_2$.

Proof: It is clear that $K_2 \subseteq K_1$ thus $k_1 \leq k_2$. We need to show the opposite direction. In the following, let us consider $R^n A^n$ as the reference system. For any optimal quantum state $\sigma_{RAB}^n \in K_1$, according to Lemma 10, there exists a quantum channel $\widetilde{\mathcal{N}}_{A' \to B}^n$ such that

$$\sigma_{RAB}^n = \widetilde{\mathcal{N}}_{A' \to B}^n(\omega_{RAA'}^n) \approx^\varepsilon \mathcal{N}_{A' \to B}^{\otimes n}(\omega_{RAA'}^n). \qquad (100)$$

Then for any permutation operation π^n, we have $\pi_{A'}^n(\omega_{A'}^n) = \omega_{A'}^n$. Then both $\pi_{A'}^n(\omega_{RAA'}^n)$ and $\omega_{RAA'}^n$ are purifications of $\omega_{A'}^n$. By Uhlmann's theorem [50], there exists a unitary $U_{\pi, RA}^n$ acting on the reference system $R^n A^n$ such that $\pi_{A'}^n(\omega_{RAA'}^n) = \mathcal{U}_{\pi, RA}^n(\omega_{RAA'}^n)$ with $\mathcal{U}_{\pi, RA}^n(\cdot) = U_{\pi, RA}^n(\cdot)(U_{\pi, RA}^n)^\dagger$. Then we have

$$\pi_B^n \circ \widetilde{\mathcal{N}}_{A' \to B}^n \circ \pi_{A'}^n(\omega_{RAA'}^n)$$
$$= \pi_B^n \circ \widetilde{\mathcal{N}}_{A' \to B}^n \circ \mathcal{U}_{\pi, RA}^n(\omega_{RAA'}^n) \qquad (101)$$
$$= \mathcal{U}_{\pi, RA}^n \circ \pi_B^n \circ \widetilde{\mathcal{N}}_{A' \to B}^n(\omega_{RAA'}^n), \qquad (102)$$

and

$$\pi_B^n \circ \mathcal{N}_{A' \to B}^{\otimes n} \circ \pi_{A'}^n(\omega_{RAA'}^n)$$
$$= \pi_B^n \circ \mathcal{N}_{A' \to B}^{\otimes n} \circ \mathcal{U}_{\pi, RA}^n(\omega_{RAA'}^n) \qquad (103)$$
$$= \mathcal{U}_{\pi, RA}^n \circ \pi_B^n \circ \mathcal{N}_{A' \to B}^{\otimes n}(\omega_{RAA'}^n). \qquad (104)$$

Since the purified distance is invariant under unitary operations and $\widetilde{\mathcal{N}}^n_{A'\to B}(\omega^n_{RAA'}) \approx^\varepsilon \mathcal{N}^{\otimes n}_{A'\to B}(\omega^n_{RAA'})$, we have

$$\pi^n_B \circ \widetilde{\mathcal{N}}^n_{A'\to B} \circ \pi^n_{A'}(\omega^n_{RAA'})$$

$$\approx^\varepsilon \pi^n_B \circ \mathcal{N}^{\otimes n}_{A'\to B} \circ \pi^n_{A'}(\omega^n_{RAA'}) \qquad (105)$$

$$= \mathcal{N}^{\otimes n}_{A'\to B}(\omega^n_{RAA'}), \qquad (106)$$

where the equality follows from the permutation invariance of $\mathcal{N}^{\otimes n}_{A'\to B}$. Due to the convexity of the purified distance, we have

$$\widetilde{\sigma}^n_{RAB} := \frac{1}{n!} \sum_{\pi_n} \pi^n_B \circ \widetilde{\mathcal{N}}^n_{A'\to B} \circ \pi^n_{A'}(\omega^n_{RAA'}) \qquad (107)$$

$$\approx^\varepsilon \mathcal{N}^{\otimes n}_{A'\to B}(\omega^n_{RAA'}). \qquad (108)$$

Note that $\frac{1}{n!} \sum_{\pi_n} \pi^n_B \circ \widetilde{\mathcal{N}}^n_{A'\to B} \circ \pi^n_{A'}$ is permutation invariant, which implies $\widetilde{\sigma}^n_{RAB} \in K_2$. Then we have

$$k_2 \leq I_{\max}(RA : B)_{\widetilde{\sigma}^n_{RAB}} \qquad (109)$$

$$= I_{\max}(RA : B)_{\Theta(\widetilde{\mathcal{N}}^n_{A'\to B})} \qquad (110)$$

$$\leq I_{\max}(RA : B)_{\widetilde{\mathcal{N}}^n_{A'\to B}} \qquad (111)$$

$$= I_{\max}(RA : B)_{\sigma^n_{RAB}} \qquad (112)$$

$$= k_1, \qquad (113)$$

where the first and second equalities follows from the fact that channel's max-information is independent on the input state (see Remark 2), the second inequality follows by the monotonicity of the channel's max-information (see Remark 5) under the superchannel $\Theta(\cdot) = \frac{1}{n!} \sum_{\pi_n} \pi^n_B(\cdot)\pi^n_{A'}$, the last equality follows from the optimality assumption of σ^n_{RAB}. This completes the proof. ∎

Lemma 12 *For any quantum states ρ_{AB} and σ_{AB} such that $\rho_A = \sigma_A$ and $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$, it holds that*

$$|I(A : B)_\rho - I(A : B)_\sigma| \leq 2\varepsilon \log |A| + (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1 + \varepsilon}\right),$$

where $h_2(\cdot)$ is the binary entropy.

Proof: Since $I(A : B)_\rho = H(A)_\rho - H(A|B)_\rho$, we have

$$|I(A : B)_\rho - I(A : B)_\sigma|$$

$$= |H(A|B)_\rho - H(A|B)_\sigma| \qquad (114)$$

$$\leq 2\varepsilon \log |A| + (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1 + \varepsilon}\right), \qquad (115)$$

where $H(A)$ and $H(A|B)$ are von Neumann entropy and conditional entropy respectively. The second inequality follows from [51, Lemma 2.]. ∎

Lemma 13 *Suppose the effective channel $\widetilde{\mathcal{M}}_{A_i\to B_o} = \Pi_{A_iB_i\to A_oB_o} \circ \mathcal{N}_{A_o\to B_i}$ with quantum channel $\mathcal{N}$ and bipartite no-signalling quantum supermap Π, then we have the relation in terms of the their Choi-Jamiołkowski matrices $J^{\widetilde{\mathcal{M}}}_{A_iB_o} = \text{Tr}_{A_oB_i}[(J^{\mathcal{N}}_{A_oB_i})^T \otimes \mathbb{1}_{A_iB_o}]J^\Pi_{A_iB_iA_oB_o}$.*

Proof: This result is widely used when considering no-signalling codes, e.g., [8], [31]. We include its proof here only for completeness. Since $\Pi_{A_iB_i\to A_oB_o}$ is required to be B to A no-signalling, we can write it as

$\Pi_{A_iB_i\to A_oB_o} = \mathcal{D}_{E_oB_i\to B_o}\mathcal{F}_{E_i\to E_o}\mathcal{E}_{A_i\to A_oE_i}$ with quantum operations $\mathcal{E}, \mathcal{F}, \mathcal{D}$. Note that the inverse Choi-Jamiołkowski isomorphism is given by $\mathcal{N}_{A\to B}(X_A) = \text{Tr}_A J^{\mathcal{N}}_{AB}(X_A^T \otimes \mathbb{1}_B)$. In the following we will not explicitly write out the identity $\mathbb{1}_B$ for simplicity. For any operator $X_{A_iB_i}$, we have

$$\Pi_{A_iB_i\to A_oB_o}(X_{A_iB_i})$$

$$= \mathcal{D}_{E_oB_i\to B_o}\mathcal{F}_{E_i\to E_o}\mathcal{E}_{A_i\to A_oE_i}(X_{A_iB_i})$$

$$= \mathcal{D}_{E_oB_i\to B_o}\mathcal{F}_{E_i\to E_o} \text{Tr}_{A_i} J^{\mathcal{E}}_{A_iA_oE_i} X^{T_{A_i}}_{A_iB_i}$$

$$= \mathcal{D}_{E_oB_i\to B_o} \text{Tr}_{E_i} J^{\mathcal{F}}_{E_iE_o}\left[\text{Tr}_{A_i}(J^{\mathcal{E}}_{A_iA_oE_i})^{T_{E_i}} X^{T_{A_i}}_{A_iB_i}\right]$$

$$= \text{Tr}_{E_oB_i} J^{\mathcal{D}}_{E_oB_iB_o}\left[\text{Tr}_{E_i}(J^{\mathcal{F}}_{E_iE_o})^{T_{E_o}}\left[\text{Tr}_{A_i}(J^{\mathcal{E}}_{A_iA_oE_i})^{T_{E_i}} X^{T_{A_iB_i}}_{A_iB_i}\right]\right]$$

$$= \text{Tr}_{A_iB_i}\left[\text{Tr}_{E_iE_o} J^{\mathcal{D}}_{E_oB_iB_o}(J^{\mathcal{F}}_{E_iE_o})^{T_{E_o}}(J^{\mathcal{E}}_{A_iA_oE_i})^{T_{E_i}}\right]X^{T_{A_iB_i}}_{A_iB_i}.$$

Thus we have

$$J^\Pi_{A_iB_iA_oB_o} = \text{Tr}_{E_iE_o} J^{\mathcal{D}}_{E_oB_iB_o}(J^{\mathcal{F}}_{E_iE_o})^{T_{E_o}}(J^{\mathcal{E}}_{A_iA_oE_i})^{T_{E_i}}. \qquad (116)$$

Repeating the above steps again for $\widetilde{\mathcal{M}}_{A_i\to B_o}(X_{A_i}) = \mathcal{D}_{E_oB_i\to B_o}\mathcal{F}_{E_i\to E_o}\mathcal{E}_{A_i\to A_oE_i}\mathcal{N}_{A_o\to B_i}(X_{A_i})$, we have

$$J^{\widetilde{M}}_{A_iB_o} = \text{Tr}_{A_oB_iE_iE_o} J^{\mathcal{D}}_{E_oB_iB_o}(J^{\mathcal{F}}_{E_iE_o})^{T_{E_o}}(J^{\mathcal{E}}_{A_iA_oE_i})^{T_{A_oE_i}}(J^{\mathcal{N}}_{A_oB_i})^{T_{B_i}}$$

By Eq. (116), we have

$$J^{\widetilde{M}}_{A_iB_o} = \text{Tr}_{A_oB_i}(J^\Pi_{A_iB_iA_oB_o})^{T_{A_o}}(J^{\mathcal{N}}_{A_oB_i})^{T_{B_i}} \qquad (117)$$

$$= \text{Tr}_{A_oB_i}(J^{\mathcal{N}}_{A_oB_i})^{T_{A_oB_i}} J^\Pi_{A_iB_iA_oB_o}, \qquad (118)$$

which completes the proof. ∎

REFERENCES

[1] D. Kretschmann and R. F. Werner, "Tema con variazioni: Quantum channel capacity," *New J. Phys.*, vol. 6, no. 1, p. 26, 2004.

[2] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, Oct. 2002.

[3] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[4] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 2926–2959, May 2014.

[5] M. Berta, M. Christandl, and R. Renner, "The quantum reverse Shannon theorem based on one-shot information theory," *Commun. Math. Phys.*, vol. 306, no. 3, pp. 579–615, Sep. 2011.

[6] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

[7] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, "Zero-error channel capacity and simulation assisted by non-local correlations," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5509–5523, Aug. 2011.

[8] R. Duan and A. Winter, "No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the Lovász number," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 891–914, Feb. 2016.

[9] X. Wang and R. Duan, "On the quantum no-signalling assisted zero-error classical simulation cost of non-commutative bipartite graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2244–2248.

[10] X. Wang and M. M. Wilde, "Exact entanglement cost of quantum states and channels under PPT-preserving operations," 2018, *arXiv:1809.09592*. [Online]. Available: https://arxiv.org/abs/1809.09592

[11] M. Berta, F. G. S. L. Brandão, M. Christandl, and S. Wehner, "Entanglement cost of quantum channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6779–6795, Oct. 2013.

[12] K. B. Dana, M. G. Díaz, M. Mejatty, and A. Winter, "Resource theory of coherence: Beyond states," *Phys. Rev. A, Gen. Phys.*, vol. 95, no. 6, 2017, Art. no. 062327.

[13] 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.

[14] 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.

[15] M. Tomamichel, "A framework for non-asymptotic quantum information theory," 2012, *arXiv:1203.2142*. [Online]. Available: https://arxiv.org/abs/1203.2142

[16] 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.

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

[18] 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.

[19] 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.

[20] 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.

[21] N. Datta, J. M. Renes, R. Renner, and M. M. Wilde, "One-shot lossy quantum data compression," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8057–8076, Dec. 2013.

[22] M. Berta, M. Christandl, and D. Touchette, "Smooth entropy bounds on one-shot quantum state redistribution," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1425–1439, Mar. 2016.

[23] A. Anshu, R. Jain, P. Mukhopadhyay, A. Shayeghi, and P. Yao, "New one shot quantum protocols with application to communication complexity," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7566–7577, Dec. 2016.

[24] N. Ciganović, N. J. Beaudry, and R. Renner, "Smooth max-information as one-shot generalization for mutual information," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1573–1581, Mar. 2014.

[25] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.

[26] X. Wang, K. Fang, and M. Tomamichel, "On converse bounds for classical communication over quantum channels," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4609–4619, Jul. 2019.

[27] W. Matthews, "A linear program for the finite block length converse of polyanskiy–poor–Verdú via nonsignaling codes," *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7036–7044, Dec. 2012.

[28] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Transforming quantum operations: Quantum supermaps," *EPL Europhys. Lett.*, vol. 83, no. 3, p. 30004, 2008.

[29] G. Gour and M. M. Wilde, "Entropy of a quantum channel," 2018, *arXiv:1808.06980*. [Online]. Available: https://arxiv.org/abs/1808.06980

[30] G. Gour, "Comparison of quantum channels by superchannels," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5880–5904, Sep. 2019.

[31] 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.

[32] X. Wang, K. Fang, and R. Duan, "Semidefinite programming converse bounds for quantum communication," *IEEE Trans. Inf. Theory*, vol. 65, no. 4, pp. 2583–2592, Apr. 2019.

[33] J. Watrous, "Semidefinite programs for completely bounded norms," *Theory Comput.*, vol. 5, pp. 217–238, 2009.

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

[35] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2816–2826, Jun. 2009.

[36] R. Renner, "Security of quantum key distribution," *Int. J. Quantum Inf.* vol. 6, no. 1, pp. 1–127, Feb. 2008.

[37] M. K. Gupta and M. M. Wilde, "Multiplicativity of completely bounded *p*-norms implies a strong converse for entanglement-assisted capacity," *Commun. Math. Phys.*, vol. 334, no. 2, pp. 867–887, Mar. 2015.

[38] T. Cooney, M. Mosonyi, and M. M. Wilde, "Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication," *Commun. Math. Phys.*, vol. 344, no. 3, pp. 797–829, Aug. 2016.

[39] A. Y. Kitaev, "Quantum computations: Algorithms and error correction," *Russian Math. Surv.*, vol. 52, no. 6, pp. 1191–1249, 1997.

[40] N. Datta, "Max-relative entropy of entanglement, alias log robustness," *Int. J. Quantum Inf.*, vol. 7, no. 2, pp. 475–491, 2009.

[41] M. G. Díaz *et al.*, "Using and reusing coherence to realize quantum processes," *Quantum*, vol. 2, p. 100, 2018.

[42] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "Partially smoothed information measures," 2018, *arXiv:1807.05630*. [Online]. Available: https://arxiv.org/abs/1807.05630

[43] M. Christandl, R. König, and R. Renner, "Postselection technique for quantum channels with applications to quantum cryptography," *Phys. Rev. Lett.*, vol. 102, no. 2, 2009, Art. no. 020504.

[44] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, no. 15, p. 3081, 1999.

[45] N. Datta, M. Tomamichel, and M. M. Wilde, "On the second-order asymptotics for entanglement-assisted communication," *Quantum Inf. Process.*, vol. 15, no. 6, pp. 2569–2591, 2016.

[46] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, "Approaches for approximate additivity of the Holevo information of quantum channels," *Phys. Rev. A, Gen. Phys.*, vol. 97, no. 1, 2018, Art. no. 012332.

[47] *This Problem Has Been Independently Found and Presented by Andreas Winter in The Open Problems Session at The Rocky Mountain Summit on Quantum Information. See.* Accessed: Jun. 2018. [Online]. Available: http://jila.colorado.edu/rmsqi/open_problems/open-problem-winter.jpg

[48] 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.

[49] M. Berta, C. Hirche, E. Kaur, and M. M. Wilde, "Amortized channel divergence for asymptotic quantum channel discrimination," 2018, *arXiv:1808.01498*. [Online]. Available: https://arxiv.org/abs/1808.01498

[50] A. Uhlmann, "The 'transition probability' in the state space of a -algebra," *Rep. Math. Phys.*, vol. 9, no. 2, pp. 273–279, 1976.

[51] A. Winter, "Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints," *Commun. Math. Phys.*, vol. 347, no. 1, pp. 291–313, 2016.

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 Research Scientist at the Institute for Quantum Computing at Baidu Research. He was a Hartree Postdoctoral Fellow at the Joint Center for Quantum Information and Computer Science 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 theory, quantum computation, and optimization theory.

Marco Tomamichel (M'13–SM'16) is an Associate Professor 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.

Mario Berta is a Lecturer (Assistant Professor) at the Department of Computing Imperial College London. He was a Postdoctoral Researcher at the Institute for Quantum Information and Matter of the California Institute of Technology. He received a Ph.D. in theoretical physics from ETH Zurich in 2013. His research interests include quantum Shannon theory and quantum cryptography.