Entanglement cost and quantum channel simulation

Mark M. Wilde

Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

(Received 2 August 2018; published 31 October 2018)

This paper proposes a revised definition for the entanglement cost of a quantum channel $\mathcal{N}$. In particular, it is defined here to be the smallest rate at which entanglement is required, in addition to free classical communication, in order to simulate n calls to $\mathcal{N}$, such that the most general discriminator cannot distinguish the n calls to $\mathcal{N}$ from the simulation. The most general discriminator is one who tests the channels in a sequential manner, one after the other, and this discriminator is known as a quantum tester [Chiribella *et al.*, Phys. Rev. Lett. **101**, 060401 (2008)] or one who is implementing a quantum costrategy [Gutoski and Watrous, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing STOC '07* (ACM Press, New York, 2007), pp. 565–574]. As such, the proposed revised definition of entanglement cost of a quantum channel leads to a rate that cannot be smaller than the previous notion of a channel's entanglement cost [Berta *et al.*, IEEE Trans. Inf. Theory **59**, 6779 (2013)], in which the discriminator is limited to distinguishing parallel uses of the channel from the simulation. Under this revised notion, I prove that the entanglement cost of certain teleportation-simulable channels is equal to the entanglement cost of their underlying resource states. Then I find single-letter formulas for the entanglement cost of some fundamental channel models, including dephasing, erasure, three-dimensional Werner-Holevo channels, and epolarizing channels (complements of depolarizing channels), as well as single-mode pure-loss and pure-amplifier bosonic Gaussian channels. These examples demonstrate that the resource theory of entanglement for quantum channels is not reversible. Finally, I discuss how to generalize the basic notions to arbitrary resource theories.

DOI: 10.1103/PhysRevA.98.042338

I. INTRODUCTION

The resource theory of entanglement [1] has been one of the richest contributions to quantum information theory [2–5], and, these days, the seminal ideas coming from it are influencing diverse areas of physics [6]. A fundamental question in entanglement theory is to determine the smallest rate at which Bell states (or ebits) are needed, along with the assistance of free classical communication, in order to generate n copies of an arbitrary bipartite state ρ_{AB} reliably (in this introduction, n should be understood to be an arbitrarily large number) [1]. The optimal rate is known as the entanglement cost of ρ_{AB} [1], and a formal expression is known for this quantity in terms of a regularization of the entanglement of formation [7]. An upper bound in terms of entanglement of formation has been known for some time [1,7], while a lower bound in terms of a semidefinite programming quantity has been determined recently [8]. Conversely, a related fundamental question is to determine the largest rate at which one can distill ebits reliably from n copies of ρ_{AB}, again with the assistance of free classical communication [1]. This optimal rate is known as the distillable entanglement, and various lower bounds [9] and upper bounds [10–13] are known for it.

The above resource theory is quite rich and interesting, but soon after learning about it, one might immediately question its operational significance. How are the bipartite states ρ_{AB} established in the first place? Of course, a quantum communication channel, such as a fiber-optic or free-space link, is required. Consequently, in the same paper that introduced the resource theory of entanglement [1], the authors there appreciated the relevance of this point and proposed that the distillation question could be extended to quantum channels. The distillation question for channels is then as follows: given n uses of a quantum channel $\mathcal{N}_{A\to B}$ connecting a sender Alice to a receiver Bob, along with the assistance of free classical communication, what is the optimal rate at which these channels can produce ebits reliably [1]? By invoking the teleportation protocol [14] and the fact that free classical communication is allowed, this rate is also equal to the rate at which arbitrary qubits can be reliably communicated by using the channel n times [1]. The optimal rate is known as the distillable entanglement of the channel [1], and various lower bounds [9] and upper bounds [15–18] are now known for it, strongly related to the bounds for distillable entanglement of states, as given above.

Some years after the distillable entanglement of a channel was proposed in [1], the question converse to it was proposed and addressed in [19]. The authors of [19] defined the entanglement cost of a quantum channel $\mathcal{N}_{A\to B}$ as the smallest rate at which entanglement is required, in addition to the assistance of free classical communication, in order to simulate n uses of $\mathcal{N}_{A\to B}$. Key to their definition of entanglement cost is the particular notion of simulation considered. In particular, the goal of their simulation protocol is to simulate n parallel uses of the channel, written as $(\mathcal{N}_{A\to B})^{\otimes n}$. Furthermore, they considered a simulation protocol $\mathcal{P}_{A^n\to B^n}$ to have the following form:

$$\mathcal{P}_{A^n\to B^n}(\omega_{A^n}) \equiv \mathcal{L}_{A^n\overline{A}_0\overline{B}_0\to B^n}\big(\omega_{A^n}\otimes\Phi_{\overline{A}_0\overline{B}_0}\big), \tag{1}$$

where ω_{A^n} is an arbitrary input state, $\mathcal{L}_{A^n\overline{A}_0\overline{B}_0\to B^n}$ is a free channel, whose implementation is restricted to consist of local operations and classical communication (LOCC) [1,20],

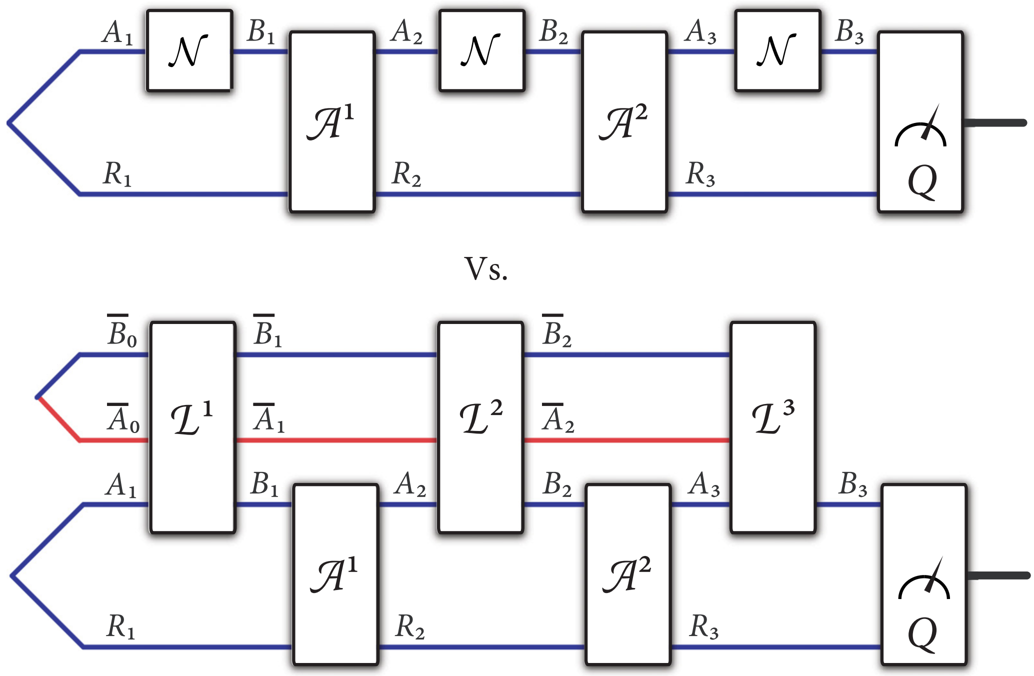

FIG. 1. Top part of the figure displays a three-round interaction between the discriminator and the simulator in the case that the actual channel $\mathcal{N}_{A\to B}$ is called three times. The bottom part of the figure displays the interaction between the discriminator and the simulator in the case that the simulation of three channel uses is called.

and $\Phi_{\overline{A}_0\overline{B}_0}$ is a maximally entangled resource state. For $\varepsilon \in [0, 1]$, the simulation is then considered ε distinguishable from $(\mathcal{N}_{A\to B})^{\otimes n}$ if the following condition holds:

$$\tfrac{1}{2}\|(\mathcal{N}_{A\to B})^{\otimes n} - \mathcal{P}_{A^n\to B^n}\|_\diamond \leqslant \varepsilon, \tag{2}$$

where $\|\cdot\|_\diamond$ denotes the diamond norm [21]. The physical meaning of the above inequality is that it places a limitation on how well any discriminator can distinguish the channel $(\mathcal{N}_{A\to B})^{\otimes n}$ from the simulation $\mathcal{P}_{A^n\to B^n}$ in a guessing game. Such a guessing game consists of the discriminator preparing a quantum state ρ_{RA^n}, the referee picking $(\mathcal{N}_{A\to B})^{\otimes n}$ or $\mathcal{P}_{A^n\to B^n}$ at random and then applying it to the A^n systems of ρ_{RA^n}, and the discriminator finally performing a quantum measurement on the systems RB^n. If the inequality in (2) holds, then the probability that the discriminator can correctly distinguish the channel from its simulation is bounded from above by $\frac{1}{2}(1+\varepsilon)$, regardless of the particular state ρ_{RA^n} and final measurement chosen for his distinguishing strategy [21–24]. Thus, if ε is close to zero, then this probability is not much better than random guessing, and in this case, the channels are considered nearly indistinguishable and the simulation thus reliable.

In parallel to the above developments in entanglement theory, there have indubitably been many advances in the theory of quantum channel discrimination [25–30] and related developments in the theory of quantum interactive proof systems [31–34]. Notably, the most general method for distinguishing a quantum memory channel from another one consists of a quantum-memory-assisted discrimination protocol [26,27]. In the language of quantum interactive proof systems, memory channels are called strategies and memory-assisted discrimination protocols are called costrategies [31–33]. For a visual illustration of the physical setup, please consult Fig. 2 of Ref. [26] or Fig. 2 of Ref. [31]. In subsequent work after [26,31], a number of theoretical results listed above have been derived related to memory channel discrimination or quantum strategies.

The aforementioned developments in the theory of quantum channel discrimination indicate that the notion of channel simulation proposed in [19] is not the most general notion that could be considered. In particular, if a simulator is claiming to have simulated n uses of the channel $\mathcal{N}_{A\to B}$, then the discriminator should be able to test this assertion in the most general way possible, as given in [26,27,31]. That is, we would like for the simulation to pass the strongest possible test that could be performed to distinguish it from the n uses of $\mathcal{N}_{A\to B}$. Such a test allows for the discriminator to prepare an arbitary state $\rho_{R_1A_1}$, call the first channel use $\mathcal{N}_{A_1\to B_1}$ or its simulation, apply an arbitrary channel $\mathcal{A}^{(1)}_{R_1B_1\to R_2A_2}$, call the second channel use or its simulation, etc. After the nth call is made, the discriminator then performs a joint measurement on the remaining quantum systems. See Fig. 1 for a visual depiction. If the simulation is good, then the probability for the discriminator to distinguish the n channels from the simulation should be no larger than $\frac{1}{2}(1+\varepsilon)$, for small ε.

In this paper, I propose an alternative definition for the entanglement cost of a channel $\mathcal{N}_{A\to B}$, such that it is the smallest rate at which ebits are needed, along with the assistance of free classical communication, in order to simulate n uses of $\mathcal{N}_{A\to B}$, in such a way that a discriminator performing the most stringest test, as described above, cannot distinguish the simulation from n actual calls of $\mathcal{N}_{A\to B}$ (Sec. II B). Here I denote the optimal rate by $E_C(\mathcal{N})$, and the prior quantity defined in [19] by $E_C^{(p)}(\mathcal{N})$, given that the simulation there was only required to pass a less stringent parallel discrimination test, as discussed above. Due to the fact that it is more difficult to pass the simulation test as specified by the alternative definition, it follows that $E_C(\mathcal{N}) \geqslant E_C^{(p)}(\mathcal{N})$ (discussed in more detail in what follows). After establishing definitions, I then prove a general upper bound on the entanglement cost of a quantum channel, using the notion of teleportation simulation (Sec. III A). I prove that the entanglement cost of certain "resource-seizable," teleportation-simulable channels takes on a particularly simple form (Sec. III B), which allows for concluding single-letter formulas for the entanglement cost of dephasing, erasure, three-dimensional Werner-Holevo channels, and epolarizing channels (complements of depolarizing channels), as detailed in Sec. IV. Note that the result about entanglement cost of dephasing channels solves an open question from [19]. I then extend the results to the case of bosonic Gaussian channels (Sec. V), proving single-letter formulas for the entanglement cost of fundamental channel models, including pure-loss and pure-amplifier channels (Theorem 2 in Sec. V G). These examples lead to the conclusion that the resource theory of entanglement for quantum channels is not reversible. I also prove that the entanglement cost of thermal, amplifier, and additive-noise bosonic Gaussian channels is bounded from below by the entanglement cost of their "Choi states." In Sec. VI, I discuss how to generalize the basic notions to other resource theories. Finally, Sec. VII concludes with a summary and some open questions.

II. NOTIONS OF QUANTUM CHANNEL SIMULATION

In this section, I review the definition of entanglement cost of a quantum channel, as detailed in [19], and I also review the main theorem from [19]. After that, I propose the revised definition of a channel's entanglement cost.

Before starting, let us define a maximally entangled state Φ_{AB} of Schmidt rank d as

$$\Phi_{AB} \equiv \frac{1}{d}\sum_{i,j=1}^{d} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_B, \tag{3}$$

where $\{|i\rangle_A\}_i$ and $\{|i\rangle_B\}_i$ are orthonormal bases. An LOCC channel $\mathcal{L}_{A'B'\to AB}$ is a bipartite channel that can be written in the following form:

$$\mathcal{L}_{A'B'\to AB} = \sum_y \mathcal{E}^y_{A'\to A} \otimes \mathcal{F}^y_{B'\to B}, \tag{4}$$

where $\{\mathcal{E}^y_{A'\to A}\}_y$ and $\{\mathcal{F}^y_{B'\to B}\}_y$ are sets of completely positive, trace-non-increasing maps, such that the sum map $\sum_y \mathcal{E}^y_{A'\to A} \otimes \mathcal{F}^y_{B'\to B}$ is a quantum channel (completely positive and trace preserving) [20]. However, not every channel of the form in (4) is an LOCC channel [there are separable channels of the form in (4) that are not implementable by LOCC [35]]. The diamond norm of the difference of two channels $\mathcal{R}_{A\to B}$ and $\mathcal{S}_{A\to B}$ is defined as [21]

$$\|\mathcal{R}-\mathcal{S}\|_\diamond \equiv \sup_{\psi_{RA}} \|\mathcal{R}_{A\to B}(\psi_{RA}) - \mathcal{S}_{A\to B}(\psi_{RA})\|_1, \tag{5}$$

where the optimization is with respect to all pure bipartite states ψ_{RA} with system R isomorphic to system A and the trace norm of an operator X is defined as $\|X\|_1 \equiv \mathrm{Tr}\{\sqrt{X^\dagger X}\}$. The operational interpretation of the diamond norm is that it is related to the maximum success probability $p_{\mathrm{succ}}(\mathcal{R},\mathcal{S})$ for any physical experiment, of the kind discussed after (2), to distinguish the channels $\mathcal{R}$ and $\mathcal{S}$:

$$p_{\mathrm{succ}}(\mathcal{R},\mathcal{S}) = \tfrac{1}{2}\big(1 + \tfrac{1}{2}\|\mathcal{R}-\mathcal{S}\|_\diamond\big). \tag{6}$$

A. Entanglement cost of a quantum channel from [19]

Let us now review the notion of entanglement cost from [19]. Fix $n, M \in \mathbb{N}$, $\varepsilon \in [0,1]$, and a quantum channel $\mathcal{N}_{A\to B}$. According to [19], an (n, M, ε) (parallel) LOCC-assisted channel simulation code consists of an LOCC channel $\mathcal{L}_{A^n\overline{A}_0\overline{B}_0\to B^n}$ and a maximally entangled resource state $\Phi_{\overline{A}_0\overline{B}_0}$ of Schmidt rank M, such that together they implement a simulation channel $\mathcal{P}_{A^n\to B^n}$, as defined in (1). In this model, to be clear, we assume that Alice has access to all systems labeled by A, Bob has access to all systems labeled by B, and they are in distant laboratories. The simulation $\mathcal{P}_{A^n\to B^n}$ is considered ε distinguishable from n parallel calls $(\mathcal{N}_{A\to B})^{\otimes n}$ of the actual channel $\mathcal{N}_{A\to B}$ if the condition in (2) holds. Note here again that the condition in (2) corresponds to a discriminator who is restricted to performing only a parallel test to distinguish the n calls of $\mathcal{N}_{A\to B}$ from its simulation. Let us also note here that the condition in (2) can be understood as the simulation $\mathcal{P}_{A^n\to B^n}$ providing an approximate teleportation simulation of $(\mathcal{N}_{A\to B})^{\otimes n}$, in the language of the later work of [36].

A rate R is said to be achievable for (parallel) channel simulation of $\mathcal{N}_{A\to B}$ if for all $\varepsilon \in (0,1]$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, \varepsilon)$ LOCC-assisted channel simulation code. The (parallel) entanglement cost $E_C^{(p)}(\mathcal{N})$ of the channel $\mathcal{N}$ is equal to the infimum of all achievable rates, with the superscript (p) indicating that the test of the simulation is restricted to being a parallel discrimination test.

The main result of [19] is that the channel's entanglement cost $E_C^{(p)}(\mathcal{N})$ is equal to the regularization of its entanglement of formation. To state this result precisely, recall that the entanglement of formation of a bipartite state ρ_{AB} is defined as [1]

$$E_F(A;B)_\rho \equiv \inf\left\{\sum_x p_X(x) H(A)_{\psi^x} : \rho_{AB} = \sum_x p_X(x)\psi^x_{AB}\right\}, \tag{7}$$

where the infimum is with respect to all convex decompositions of ρ_{AB} into pure states ψ^x_{AB} and

$$H(A)_{\psi^x} \equiv -\mathrm{Tr}\left\{\psi^x_A \log_2 \psi^x_A\right\} \tag{8}$$

is the quantum entropy of the marginal state $\psi^x_A = \mathrm{Tr}_B\{\psi^x_{AB}\}$. The entanglement of formation does not increase under the action of an LOCC channel [1]. A channel's entanglement of formation $E_F(\mathcal{N})$ is then defined as

$$E_F(\mathcal{N}) \equiv \sup_{\psi_{RA}} E_F(R;B)_\omega, \tag{9}$$

where $\omega_{RB} \equiv \mathcal{N}_{A\to B}(\psi_{RA})$, and it suffices to take the optimization with respect to a pure state input ψ_{RA}, with system R isomorphic to system A, due to purification, the Schmidt decomposition theorem, and the LOCC monotonicity of entanglement of formation [1]. We can now state the main result of [19] described above:

$$E_C^{(p)}(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} E_F(\mathcal{N}^{\otimes n}). \tag{10}$$

The regularized formula on the right-hand side may be difficult to evaluate in general, and thus can only be considered a formal expression, but if the additivity relation $\frac{1}{n}E_F(\mathcal{N}^{\otimes n}) = E_F(\mathcal{N})$ holds for a given channel $\mathcal{N}$ for all $n \geqslant 1$, then it simplifies significantly as $E_C^{(p)}(\mathcal{N}) = E_F(\mathcal{N})$.

B. Proposal for a revised notion of entanglement cost of a channel

Now I propose the revised definition for entanglement cost of a channel. As motivated in the Introduction, a parallel test of channel simulation is not the most general kind of test that can be considered. Thus the revised definition proposes that the entanglement cost of a channel should incorporate the most stringent test possible.

To begin with, let us fix $n, M \in \mathbb{N}$, $\varepsilon \in [0, 1]$, and a quantum channel $\mathcal{N}_{A\to B}$. We define an (n, M, ε) (sequential) LOCC-assisted channel simulation code to consist of a maximally entangled resource state $\Phi_{\overline{A}_0\overline{B}_0}$ of Schmidt rank M and a set

$$\left\{\mathcal{L}^{(i)}_{A_i\overline{A}_{i-1}\overline{B}_{i-1}\to B_i\overline{A}_i\overline{B}_i}\right\}_{i=1}^{n} \tag{11}$$

of LOCC channels. Note that the systems $\overline{A}_n\overline{B}_n$ of the final LOCC channel $\mathcal{L}^{(n)}_{A_n\overline{A}_{n-1}\overline{B}_{n-1}\to B_n\overline{A}_n\overline{B}_n}$ can be taken trivial without loss of generality. As before, Alice has access to all systems labeled by A, Bob has access to all systems labeled by B, and they are in distant laboratories. The structure of this simulation protocol is intended to be compatible with a discrimination strategy that can test the actual n channels versus the above simulation in a sequential way, along the lines discussed in [26,27,33]. I later show how this encompasses the parallel tests discussed in the previous section.

A sequential discrimination strategy consists of an initial state $\rho_{R_1A_1}$, a set $\{\mathcal{A}^{(i)}_{R_iB_i\to R_{i+1}A_{i+1}}\}_{i=1}^{n-1}$ of adaptive channels, and a quantum measurement $\{Q_{R_nB_n}, I_{R_nB_n} - Q_{R_nB_n}\}$. Let us employ the shorthand $\{\rho, \mathcal{A}, Q\}$ to abbreviate such a discrimination strategy. Note that, in performing a discrimination strategy, the discriminator has a full description of the channel $\mathcal{N}_{A\to B}$ and the simulation protocol, which consists of $\Phi_{\overline{A}_0\overline{B}_0}$ and the set in (11). If this discrimination strategy is performed on the n uses of the actual channel $\mathcal{N}_{A\to B}$, the relevant states involved are

$$\rho_{R_{i+1}A_{i+1}} \equiv \mathcal{A}^{(i)}_{R_iB_i\to R_{i+1}A_{i+1}}\big(\rho_{R_iB_i}\big), \tag{12}$$

for $i \in \{1, \dots, n-1\}$, and

$$\rho_{R_iB_i} \equiv \mathcal{N}_{A_i\to B_i}\big(\rho_{R_iA_i}\big), \tag{13}$$

for $i \in \{1, \dots, n\}$. If this discrimination strategy is performed on the simulation protocol discussed above, then the relevant states involved are

$$\begin{aligned}\tau_{R_1B_1\overline{A}_1\overline{B}_1} &\equiv \mathcal{L}^{(1)}_{A_1\overline{A}_0\overline{B}_0\to B_1\overline{A}_1\overline{B}_1}\big(\tau_{R_1A_1}\otimes\Phi_{\overline{A}_0\overline{B}_0}\big),\\ \tau_{R_{i+1}A_{i+1}\overline{A}_i\overline{B}_i} &\equiv \mathcal{A}^{(i)}_{R_iB_i\to R_{i+1}A_{i+1}}\big(\tau_{R_iB_i\overline{A}_i\overline{B}_i}\big),\end{aligned} \tag{14}$$

for $i \in \{1, \dots, n-1\}$, where $\tau_{R_1A_1} = \rho_{R_1A_1}$, and

$$\tau_{R_iB_i\overline{A}_i\overline{B}_i} \equiv \mathcal{L}^{(i)}_{A_i\overline{A}_{i-1}\overline{B}_{i-1}\to B_i\overline{A}_i\overline{B}_i}\big(\tau_{R_iA_i\overline{A}_{i-1}\overline{B}_{i-1}}\big), \tag{15}$$

for $i \in \{2, \dots, n\}$. The discriminator then performs the measurement $\{Q_{R_nB_n}, I_{R_nB_n} - Q_{R_nB_n}\}$ and guesses "actual channel" if the outcome is $Q_{R_nB_n}$ and "simulation" if the outcome is $I_{R_nB_n} - Q_{R_nB_n}$. Figure 1 depicts the discrimination strategy in the case that the actual channel is called $n = 3$ times and in the case that the simulation is performed.

If the *a priori* probabilities for the actual channel or simulation are equal, then the success probability of the discriminator in distinguishing the channels is given by

$$\begin{aligned}&\tfrac{1}{2}\big[\operatorname{Tr}\{Q_{R_nB_n}\rho_{R_nB_n}\} + \operatorname{Tr}\{(I_{R_nB_n} - Q_{R_nB_n})\tau_{R_nB_n}\}\big]\\ &\quad\leqslant \tfrac{1}{2}\big(1 + \tfrac{1}{2}\|\rho_{R_nB_n} - \tau_{R_nB_n}\|_1\big),\end{aligned} \tag{16}$$

where the latter inequality is well known from the theory of quantum state discrimination [22–24]. For this reason, we say that the n calls to the actual channel $\mathcal{N}_{A\to B}$ are ε distinguishable from the simulation if the following condition holds for the respective final states:

$$\tfrac{1}{2}\|\rho_{R_nB_n} - \tau_{R_nB_n}\|_1 \leqslant \varepsilon. \tag{17}$$

If this condition holds for all possible discrimination strategies $\{\rho, \mathcal{A}, Q\}$, i.e., if

$$\frac{1}{2}\sup_{\{\rho,\mathcal{A}\}}\|\rho_{R_nB_n} - \tau_{R_nB_n}\|_1 \leqslant \varepsilon, \tag{18}$$

then the simulation protocol constitutes an (n, M, ε) channel simulation code. It is worthwhile to remark the following: if we ascribe the shorthand $(\mathcal{N})^n$ for the n uses of the channel and the shorthand $(\mathcal{L})^n$ for the simulation, then the condition in (18) can be understood in terms of the n-round strategy norm of [26,27,33]

$$\tfrac{1}{2}\|(\mathcal{N})^n - (\mathcal{L})^n\|_{\diamond,n} \leqslant \varepsilon. \tag{19}$$

As before, a rate R is achievable for (sequential) channel simulation of $\mathcal{N}$ if for all $\varepsilon \in (0, 1]$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, \varepsilon)$ (sequential) channel simulation code for $\mathcal{N}$. We define the (sequential) entanglement cost $E_C(\mathcal{N})$ of the channel $\mathcal{N}$ to be the infimum of all achievable rates. Due to the fact that this notion is more general, we sometimes simply refer to $E_C(\mathcal{N})$ as the entanglement cost of the channel $\mathcal{N}$ in what follows.

C. LOCC monotonicity of the entanglement cost

Let us note here that if a channel $\mathcal{N}_{A\to B}$ can be realized from another channel $\mathcal{M}_{A'\to B'}$ via a preprocessing LOCC channel $\mathcal{L}^{\text{pre}}_{A\to A'A_MB_M}$ and a postprocessing LOCC channel $\mathcal{L}^{\text{post}}_{B'A_MB_M\to B}$ as

$$\mathcal{N}_{A\to B} = \mathcal{L}^{\text{post}}_{B'A_MB_M\to B}\circ\mathcal{M}_{A'\to B'}\circ\mathcal{L}^{\text{pre}}_{A\to A'A_MB_M}, \tag{20}$$

then it follows that any (n, M, ε) protocol for sequential channel simulation of $\mathcal{M}_{A'\to B'}$ realizes an (n, M, ε) protocol for sequential channel simulation of $\mathcal{N}_{A\to B}$. This is an immediate consequence of the fact that the best strategy for discriminating $\mathcal{N}_{A\to B}$ from its simulation can be understood as a particular strategy for discriminating $\mathcal{M}_{A'\to B'}$ from a simulation of $\mathcal{M}_{A'\to B'}$, due to the structural decomposition in

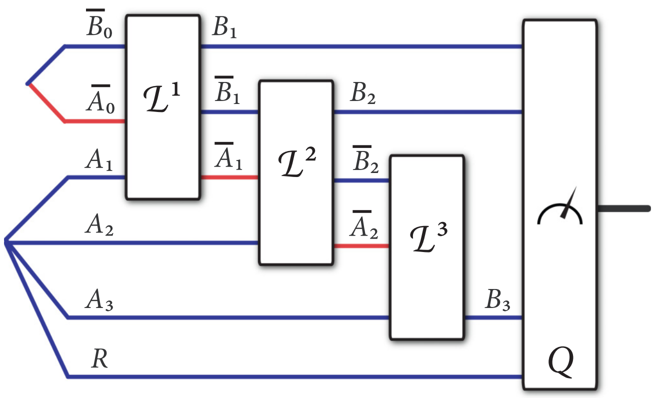

FIG. 2. Simulation protocol from the bottom part of Fig. 1 rewritten to clarify that it can participate in a parallel channel simulation test.

(20). Following definitions, a simple consequence is the following LOCC monotonicity inequality for the entanglement cost of these channels:

$$E_C(\mathcal{N}) \leqslant E_C(\mathcal{M}). \tag{21}$$

Thus it takes more or the same entanglement to simulate the channel $\mathcal{M}$ than it does to simulate $\mathcal{N}$. Furthermore, the decomposition in (20) and the bound in (21) can be used to bound the entanglement cost of a channel $\mathcal{M}$ from below. Note that the structure in (20) was discussed recently in the context of general resource theories in Ref. [6], Sec. III-D-5.

D. Parallel tests as a special case of sequential tests

A parallel test of the form described in Sec. II A is a special case of the sequential test outlined above. One can see this in two seemingly different ways. First, we can think of the sequential strategy taking a particular form. The state $\xi_{RA_1A_2\cdots A_n}$ is prepared, and here we identify systems $RA_2\cdots A_n$ with system R_1 of $\rho_{R_1A_1}$ in an adaptive protocol and system A_1 of $\xi_{RA_1A_2\cdots A_n}$ with system A_1 of $\rho_{R_1A_1}$. Then the channel $\mathcal{N}_{A_1\to B_1}$ or its simulation is called. After that, the action of the first adaptive channel is simply to swap in system A_2 of $\xi_{RA_1A_2\cdots A_n}$ to the second call of the channel $\mathcal{N}_{A_2\to B_2}$ or its simulation, while keeping systems $RB_1A_3\cdots A_n$ as part of the reference R_2 of the state $\rho_{R_2A_2}$. Then this iterates and the final measurement is performed on all of the remaining systems.

The other way to see how a parallel test is a special kind of sequential test is to rearrange the simulation protocol as has been done in Fig. 2. Here, we see that the simulation protocol has a memory structure, and it is clear that the simulation protocol can accept as input a state $\xi_{RA_1A_2\cdots A_n}$ and outputs a state on systems $RB_1\cdots B_n$, which can subsequently be measured.

As a consequence of this reduction, any (n, M, ε) sequential channel simulation protocol can serve as an (n, M, ε) parallel channel simulation protocol. Furthermore, if R is an achievable rate for sequential channel simulation, then it is also an achievable rate for parallel channel simulation. Finally, these reductions imply the following inequality:

$$E_C(\mathcal{N}) \geqslant E_C^{(p)}(\mathcal{N}). \tag{22}$$

Intuitively, one might sometimes require more entanglement in order to pass the more stringent test that occurs in sequential channel simulation. As a consequence of (10) and (22), we have that

$$E_C(\mathcal{N}) \geqslant \lim_{n\to\infty} \frac{1}{n} E_F(\mathcal{N}^{\otimes n}). \tag{23}$$

It is an interesting question (not addressed here) to determine if there exists a channel such that the inequality in (22) is strict.

If desired, it is certainly possible to obtain a nonasymptotic, weak-converse bound that implies the above bound after taking limits. Let us state this bound as follows.

Proposition 1. Let $\mathcal{N}_{A\to B}$ be a quantum channel, and let $n, M \in \mathbb{N}$ and $\varepsilon \in [0,1]$. Set $d = \min\{|A|, |B|\}$, i.e., the minimum of the input and output dimensions of the channel $\mathcal{N}_{A\to B}$. Then the following bound holds for any (n, M, ε) sequential channel simulation code:

$$\frac{1}{n}\log_2 M \geqslant \frac{1}{n} E_F(\mathcal{N}^{\otimes n}) - \sqrt{\varepsilon}\log_2 d - \frac{1}{n} g_2(\sqrt{\varepsilon}), \tag{24}$$

where $\frac{1}{n}\log_2 M$ is understood as the nonasymptotic entanglement cost of the protocol and the bosonic entropy function $g_2(x)$ is defined for $x \geqslant 0$ as

$$g_2(x) \equiv (x+1)\log_2(x+1) - x\,\log_2 x. \tag{25}$$

Proof. To see this, suppose that there exists an (n, M, ε) protocol for sequential channel simulation. Then by the above reasoning (also see Fig. 2), it can be thought of as a parallel channel simulation protocol, such that the criterion in (2) holds. Suppose that $\psi_{RA_1\cdots A_n}$ is a test input state, with $|R| = |A|^n$, leading to $\omega_{RB_1\cdots B_n} = (\mathcal{N}_{A\to B})^{\otimes n}(\psi_{RA_1\cdots A_n})$ when the actual channels are applied and $\sigma_{RB_1\cdots B_n}$ when the simulation is applied. Then we have that

$$\begin{aligned} &E_F(R; B_1\cdots B_n)_\omega \\ &\quad\leqslant E_F(R; B_1\cdots B_n)_\sigma + n\sqrt{\varepsilon}\log_2 d + g_2(\sqrt{\varepsilon}) \\ &\quad\leqslant E_F(RA_1\cdots A_n\overline{A}_0; \overline{B}_0)_{\psi\otimes\Phi} + n\sqrt{\varepsilon}\log_2 d + g_2(\sqrt{\varepsilon}) \\ &\quad= E_F(\overline{A}_0; \overline{B}_0)_\Phi + n\sqrt{\varepsilon}\log_2 d + g_2(\sqrt{\varepsilon}) \\ &\quad= \log_2 M + n\sqrt{\varepsilon}\log_2 d + g_2(\sqrt{\varepsilon}). \end{aligned} \tag{26}$$

The first inequality follows from the condition in (18), as well as from the continuity bound for entanglement of formation from Ref. [37], Corollary 4. The second inequality follows from the LOCC monotonicity of the entanglement of formation [1], here thinking of the person who possesses systems $RA_1\cdots A_n$ to be in the same laboratory as the one possessing the systems $\overline{A}_i$, while the person who possesses the $\overline{B}_i$ systems is in a different laboratory. The first equality follows from the fact that $\psi_{RA_1\cdots A_n}$ is in tensor product with $\Phi_{\overline{A}_0\overline{B}_0}$, so that, by a local channel, one may remove $\psi_{RA_1\cdots A_n}$ or append it for free. The final equality follows because the entanglement of formation of the maximally entangled state is equal to the logarithm of its Schmidt rank. Since the bound holds uniformly regardless of the input state $\psi_{RA_1\cdots A_n}$, after an optimization and a rearrangement we conclude the stated lower bound on the nonasymptotic entanglement cost $\frac{1}{n}\log_2 M$ of the protocol. ∎

Remark 1. Let us note here that the entanglement cost of a quantum channel is equal to zero if and only if the channel is entanglement breaking [38,39]. The "if part" follows as a straightforward consequence of definitions and the fact that these channels can be implemented as a measurement followed by a preparation [38,39], given that this measure-prepare procedure is a particular kind of LOCC and thus allowed for free (without any cost) in the above model. The "only-if" part follows from (22) and Ref. [19], Corollary 18, the latter of which depends on the result from [40].

III. BOUNDS FOR THE ENTANGLEMENT COST OF TELEPORTATION-SIMULABLE CHANNELS

A. Upper bound on the entanglement cost of teleportation-simulable channels

The most trivial method for simulating a channel is to employ the teleportation protocol [14] directly. In this method, Alice and Bob could use the teleportation protocol so that Alice could transmit the input of the channel to Bob, who could then apply the channel. Repeating this n times, this trivial method would implement an $(n, |A|^n, 0)$ simulation protocol in either the parallel or sequential model. Alternatively, Alice could apply the channel first and then teleport the output to Bob, and repeating this n times would implement an $(n, |B|^n, 0)$ simulation protocol in either the parallel or sequential model. Thus they could always achieve a rate of $\log_2(\min\{|A|, |B|\})$ using this approach, and this reasoning establishes a simple dimension upper bound on the entanglement cost of a channel:

$$E_C(\mathcal{N}_{A\to B}) \leqslant \log_2(\min\{|A|, |B|\}). \tag{27}$$

In this context, also see Ref. [36], Proposition 9.

A less trivial approach is to exploit the fact that some channels of interest could be teleportation simulable with associated resource state $\omega_{A'B'}$, in which the resource state need not be a maximally entangled state [see Ref. [1], Sec. V, and Ref. [41], Eq. (11)]. Recall from these references that a channel $\mathcal{N}_{A\to B}$ is teleportation simulable with associated resource state $\omega_{A'B'}$ if there exists an LOCC channel $\mathcal{L}_{AA'B'\to B}$ such that the following equality holds for all input states ρ_A:

$$\mathcal{N}_{A\to B}(\rho_A) = \mathcal{L}_{AA'B'\to B}(\rho_A \otimes \omega_{A'B'}). \tag{28}$$

If a channel possesses this structure, then we arrive at the following upper bound on the entanglement cost.

Proposition 2. Let $\mathcal{N}_{A\to B}$ be a quantum channel that is teleportation simulable with associated resource state $\omega_{A'B'}$, as defined in (28). Let $n, M \in \mathbb{N}$ and $\varepsilon \in (0,1)$. Then there exists an $(n, M, \sqrt{\varepsilon})$ sequential channel simulation code satisfying the following bound:

$$\frac{1}{n}\log_2 M \leqslant \frac{1}{n} E_{F,0}^{\varepsilon/2}(A'^n; B'^n)_{\omega^{\otimes n}}, \tag{29}$$

where $\frac{1}{n}\log_2 M$ is understood as the nonasymptotic entanglement cost of the protocol and $E_{F,0}^{\varepsilon/2}(A'^n; B'^n)_{\omega^{\otimes n}}$ is the $\varepsilon/2$-smooth entanglement of formation (EOF) [42] recalled in Definition 1 below.

Definition 1 (Smooth EOF [42]). Let $\delta \in (0,1)$ and τ_{CD} be a bipartite state. Let $\mathcal{E} = \{p_X(x), \phi^x_{CD}\}$ denote a pure-state ensemble decomposition of τ_{CD}, meaning that $\tau_{CD} = \sum_x p_X(x)\phi^x_{CD}$, where ϕ^x_{CD} is a pure state and p_X is a probability distribution. Define the conditional entropy of order zero $H_0(K|L)_\omega$ of a bipartite state ω_{KL} as

$$H_0(K|L)_\omega \equiv \max_{\sigma_L} \log_2 \operatorname{Tr}\left\{\Pi^\omega_{KL}(I_K \otimes \sigma_L)\right\}, \tag{30}$$

where Π^ω_{KL} denotes the projection onto the support of ω_{KL} and σ_L is a density operator. Then the δ-smooth entanglement of formation of τ_{CD} is given by

$$E^\delta_{F,0}(C; D)_\tau \equiv \min_{\mathcal{E}, \widetilde{\tau}_{XC} \in B^\delta_{cq}(\tau_{XC})} H_0(C|X)_{\widetilde{\tau}}, \tag{31}$$

where the minimization is with respect to all pure-state ensemble decompositions $\mathcal{E}$ of τ_{CD}, $\tau_{XCD} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \phi^x_{CD}$ is a labeled pure-state extension of τ_{CD}, and the δ-ball $B^\delta_{cq}(\tau_{XC})$ of cq states for a cq state τ_{XC} is defined as

$$B^\delta_{cq}(\tau_{XC}) \equiv \left\{\omega_{XC} : \omega_{XC} \geqslant 0,\ \omega_{XC} = \sum_x |x\rangle\langle x| \otimes \omega^x_C,\right.$$
$$\left.\|\omega_{XC} - \tau_{XC}\|_1 \leqslant \delta\right\}. \tag{32}$$

The δ-smooth entanglement of formation has the property that, for a tensor-power state $\tau^{\otimes n}_{CD}$, the following limit holds (Ref. [42], Theorem 2):

$$\lim_{\delta\to 0}\lim_{n\to\infty} \frac{1}{n} E^\delta_{F,0}(C^n; D^n)_{\tau^{\otimes n}} = \lim_{n\to\infty} \frac{1}{n} E_F(C; D)_\tau, \tag{33}$$
$$= E_C(\tau_{CD}), \tag{34}$$

where the latter quantity denotes the entanglement cost of the state τ_{CD} [7].

Proof of Proposition 2. The approach for an (n, M, ε) sequential channel simulation consists of the following steps.

First, employ the one-shot entanglement cost protocol from Ref. [42], Theorem 1, which consumes a maximally entangled state $\Phi_{\overline{A}_0\overline{B}_0}$ of Schmidt rank M along with an LOCC channel $\mathcal{P}_{\overline{A}_0\overline{B}_0\to A'^nB'^n}$ to generate n approximate copies of the resource state $\omega_{A'B'}$. In particular, using the maximally entangled state $\Phi_{\overline{A}_0\overline{B}_0}$ with

$$\log_2 M = E^{\varepsilon/2}_{F,0}(A'^n; B'^n)_{\omega^{\otimes n}}, \tag{35}$$

one can achieve the following approximation (Ref. [42], Theorem 1):

$$\tfrac{1}{2}\left\|\omega^{\otimes n}_{A'B'} - \widetilde{\omega}_{A'^nB'^n}\right\|_1 \leqslant \sqrt{\varepsilon}, \tag{36}$$

where

$$\widetilde{\omega}_{A'^nB'^n} \equiv \mathcal{P}_{\overline{A}_0\overline{B}_0\to A'^nB'^n}\left(\Phi_{\overline{A}_0\overline{B}_0}\right). \tag{37}$$

Next, at the first instance in which the channel should be simulated, Alice and Bob apply the LOCC channel $\mathcal{L}_{AA'B'\to B}$ from (28) to the A'_1 and B'_1 systems of $\widetilde{\omega}_{A'^nB'^n}$. For the second instance, they apply the LOCC channel $\mathcal{L}_{AA'B'\to B}$ from (28) to the A'_2 and B'_2 systems of $\widetilde{\omega}_{A'^nB'^n}$. This continues for the next $n-2$ rounds of the sequential channel simulation.

By the data processing inequality for trace distance, it is guaranteed that the following bound holds on the performance of this protocol for sequential channel simulation:

$$\tfrac{1}{2}\|(\mathcal{N})^n - (\mathcal{L})^n\|_{\diamond,n} \leqslant \tfrac{1}{2}\left\|\omega_{A'B'}^{\otimes n} - \widetilde{\omega}_{A'^n B'^n}\right\|_1 \leqslant \sqrt{\varepsilon}. \quad (38)$$

This follows because the distinguishability of the simulation from the actual channel uses is limited by the distinguishability of the states $\omega_{A'B'}^{\otimes n}$ and $\widetilde{\omega}_{A'^n B'^n}$, due to the assumed structure of the channel in (28), as well as the structure of the sequential channel simulation. ∎

By applying definitions, the bound in Proposition 2, taking the limits $n \to \infty$ and then $\varepsilon \to 0$ (with $M = 2^{n[R+\delta]}$ for a fixed rate R and arbitrary $\delta > 0$), and applying (33), we conclude the following statement.

Corollary 1. Let $\mathcal{N}_{A\to B}$ be a quantum channel that is teleportation simulable with associated resource state $\omega_{A'B'}$, as defined in (28). Then the entanglement cost of the channel $\mathcal{N}$ is never larger than the entanglement cost of the resource state $\omega_{A'B'}$:

$$E_C(\mathcal{N}) \leqslant E_C(\omega_{A'B'}). \quad (39)$$

The above corollary captures the intuitive idea that if a single instance of the channel $\mathcal{N}$ can be simulated via LOCC starting from a resource state $\omega_{A'B'}$, then the entanglement cost of the channel should not exceed the entanglement cost of the resource state. The idea of the above proof is simply to prepare a large number n of copies of $\omega_{A'B'}$ approximately and then use these to simulate n uses of the channel $\mathcal{N}$, such that the simulation could not be distinguished from n uses of the channel $\mathcal{N}$ in any sequential test.

B. Entanglement cost of resource-seizable, teleportation-simulable channels

In this section, I define teleportation-simulable channels that are resource seizable, meaning that one can seize the channel's underlying resource state by the following procedure: (1) prepare a free, separable state, (2) input one of its systems to the channel, and then (3) postprocess with a free, LOCC channel.

This procedure is indeed related to the channel processing described earlier in (20). After that, I prove that the entanglement cost of a resource-seizable channel is equal to the entanglement cost of its underlying resource state.

Definition 2 (Resource-seizable channel). Let $\mathcal{N}_{A\to B}$ be a teleportation-simulable channel with associated resource state $\omega_{A'B'}$, as defined in (28). Suppose that there exists a separable input state $\rho_{A_M A B_M}$ to the channel and a postprocessing LOCC channel $\mathcal{D}_{A_M B B_M \to A'B'}$ such that the resource state $\omega_{A'B'}$ can be seized from the channel $\mathcal{N}_{A\to B}$ as follows:

$$\mathcal{D}_{A_M B B_M\to A'B'}\left[\mathcal{N}_{A\to B}\left(\rho_{A_M A B_M}\right)\right] = \omega_{A'B'}. \quad (40)$$

Then we say that the channel is a resource-seizable, teleportation-simulable channel.

In Appendix A, I discuss how resource-seizable channels are related to those that are "implementable from their image," as defined in Ref. [43], Appendix A. In Sec. VI, I also discuss how to generalize the notion of a resource-seizable channel to an arbitrary resource theory.

The main result of this section is the following simplifying form for the entanglement cost of a resource-seizable channel (as defined above), establishing that its entanglement cost in the asymptotic regime is the same as the entanglement cost of the underlying resource state. Furthermore, for these channels, the entanglement cost is not increased by the need to pass a more stringest test for channel simulation as required in a sequential test.

Theorem 1. Let $\mathcal{N}_{A\to B}$ be a resource-seizable, teleportation-simulable channel with associated resource state $\omega_{A'B'}$, as given in Definition 2. Then the entanglement cost of the channel $\mathcal{N}_{A\to B}$ is equal to its parallel entanglement cost, which in turn is equal to the entanglement cost of the resource state $\omega_{A'B'}$:

$$E_C(\mathcal{N}) = E_C^{(p)}(\mathcal{N}) = E_C(\omega_{A'B'}). \quad (41)$$

Proof. Consider from (22) that

$$E_C(\mathcal{N}) \geqslant E_C^{(p)}(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} E_F(\mathcal{N}^{\otimes n}). \quad (42)$$

Let $\psi_{RA^n} \equiv \psi_{RA_1\cdots A_n}$ be an arbitrary pure input state to consider at the input of the tensor-power channel $(\mathcal{N}_{A\to B})^{\otimes n}$, leading to the state

$$\sigma_{RB^n} \equiv (\mathcal{N}_{A\to B})^{\otimes n}(\psi_{RA_1\cdots A_n}). \quad (43)$$

From the assumption that the channel is teleportation-simulable with associated resource state $\omega_{A'B'}$, we have from (28) that

$$\sigma_{RB^n} = (\mathcal{L}_{AA'B'\to B})^{\otimes n}\left(\psi_{RA^n} \otimes \omega_{A'B'}^{\otimes n}\right). \quad (44)$$

Then

$$E_F(R;B^n)_\sigma \leqslant E_F(RA^n A'^n; B'^n)_{\psi\otimes\omega^{\otimes n}} \quad (45)$$

$$= E_F(A'^n; B'^n)_{\omega^{\otimes n}}, \quad (46)$$

where the inequality follows from LOCC monotonicity of the entanglement of formation. Since the bound holds for an arbitrary input state, we conclude that the following inequality holds for all $n \in \mathbb{N}$:

$$\frac{1}{n} E_F(\mathcal{N}^{\otimes n}) \leqslant \frac{1}{n} E_F(A'^n; B'^n)_{\omega^{\otimes n}}. \quad (47)$$

Now taking the limit $n \to \infty$, we conclude that

$$E_C^{(p)}(\mathcal{N}) \leqslant E_C(\omega_{A'B'}). \quad (48)$$

To see the other inequality, let a decomposition of the separable input state $\rho_{A_M A B_M}$ be given by

$$\rho_{A_M A B_M} = \sum_x p_X(x) \psi^x_{A_M A} \otimes \phi^x_{B_M}. \quad (49)$$

Considering that $[\psi^x_{A_M A}]^{\otimes n}$ is a particular input to the tensor-power channel $(\mathcal{N}_{A\to B})^{\otimes n}$, we conclude that

$$E_F(\mathcal{N}^{\otimes n}) \geqslant E_F\left(A_M^n; B^n\right)_{[\mathcal{N}(\psi^x)]^{\otimes n}}. \quad (50)$$

Since this holds for all x, we have that

$$E_F(\mathcal{N}^{\otimes n}) \geqslant \sum_x p_X(x) E_F\left(A_M^n; B^n\right)_{[\mathcal{N}(\psi^x)]^{\otimes n}}$$

$$= \sum_x p_X(x) E_F\left(A_M^n; B^n B_M^n\right)_{[\mathcal{N}(\psi^x)\otimes\phi^x]^{\otimes n}}$$

$$\geqslant E_F\big(A_M^n; B^n B_M^n\big)_{[\mathcal{N}(\rho)]^{\otimes n}}$$
$$\geqslant E_F(A'^n; B'^n)_{\omega^{\otimes n}}, \quad (51)$$

where the equality follows because introducing a product state locally does not change the entanglement, the second inequality follows from convexity of entanglement of formation [1], and the last inequality follows from the assumption in (40) and the LOCC monotonicity of the entanglement of formation. Since the inequality holds for all $n \in \mathbb{N}$, we can divide by n and take the limit $n \to \infty$ to conclude that

$$E_C^{(p)}(\mathcal{N}) \geqslant E_C(\omega_{A'B'}), \quad (52)$$

and in turn, from (48), that

$$E_C^{(p)}(\mathcal{N}) = E_C(\omega_{A'B'}). \quad (53)$$

Combining this equality with the inequalities in (39) and (42) leads to the statement of the theorem. ■

IV. EXAMPLES

The equality in Theorem 1 provides a formal expression for the entanglement cost of any resource-seizable, teleportation-simulable channel, given in terms of the entanglement cost of the underlying resource state $\omega_{A'B'}$. Due to the fact that the entanglement cost of a state is generally not equal to its entanglement of formation [44], it could still be a significant challenge to compute the entanglement cost of these special channels. However, for some special states, the equality $E_C(\omega_{A'B'}) = E_F(A'; B')_\omega$ does hold, and I discuss several of these examples and related channels here.

Let us begin by recalling the notion of a covariant channel $\mathcal{N}_{A\to B}$ [45]. For a group G with unitary channel representations $\{\mathcal{U}_A^g\}_g$ and $\{\mathcal{V}_B^g\}_g$ acting on the input system A and output system B of the channel $\mathcal{N}_{A\to B}$, the channel $\mathcal{N}_{A\to B}$ is covariant with respect to the group G if the following equality holds:

$$\mathcal{N}_{A\to B} \circ \mathcal{U}_A^g = \mathcal{V}_B^g \circ \mathcal{N}_{A\to B}. \quad (54)$$

If the averaging channel is such that $\frac{1}{|G|}\sum_g \mathcal{U}_A^g(X) = \mathrm{Tr}[X]I/|A|$ (implementing a unitary one-design), then we simply say that the channel $\mathcal{N}_{A\to B}$ is covariant.

Then from Sec. 7 of Ref. [46] (see also Ref. [47], Appendix A), we conclude that any covariant channel is teleportation simulable with associated resource state given by the Choi state of the channel, i.e., $\omega_{A'B'} = \mathcal{N}_{A\to B}(\Phi_{A'A})$. As such, covariant channels are resource seizable, so that the equality in Theorem 1 applies to all covariant channels. Thus the entanglement cost of a covariant channel is equal to the entanglement cost of its Choi state. In spite of this reduction, it could still be a great challenge to compute formulas for the entanglement cost of these channels, due to the fact that the entanglement of formation is not necessarily equal to the entanglement cost for the Choi states of these channels. For example, the entanglement cost of an isotropic state [48,49], which is the Choi state of a depolarizing channel, is not known. In the next few subsections, I detail some example channels for which it is possible to characterize their entanglement cost.

A. Erasure channels

A simple example of a channel that is covariant is the quantum erasure channel, defined as [50]

$$\mathcal{E}^q(\rho) \equiv (1-q)\rho + q|e\rangle\langle e|, \quad (55)$$

where ρ is a d-dimensional input state, $q \in [0,1]$ is the erasure probability, and $|e\rangle\langle e|$ is a pure erasure state orthogonal to any input state, so that the output state has $d+1$ dimensions. By the remark above, we conclude that $E_C(\mathcal{E}^q) = E_C(\mathcal{E}^q_{A\to B}(\Phi_{RA}))$, and so determining the entanglement cost boils down to determining the entanglement cost of the Choi state

$$\mathcal{E}^q_{A\to B}(\Phi_{RA}) = (1-q)\Phi_{RA} + \frac{I_R}{d} \otimes |e\rangle\langle e|. \quad (56)$$

An obvious pure-state decomposition for $\mathcal{E}^q_{A\to B}(\Phi_{RA})$ [see Ref. [19], Eqs. (93)–(95)] leads to

$$E_C(\mathcal{E}^q_{A\to B}(\Phi_{RA})) \leqslant E_F\big(\mathcal{E}^q_{A\to B}(\Phi_{RA})\big) \quad (57)$$
$$\leqslant (1-q)\log_2 d. \quad (58)$$

As it turns out, these inequalities are tight, due to an operational argument. In particular, the distillable entanglement of $\mathcal{E}^q_{A\to B}(\Phi_{RA})$ is exactly equal to $(1-q)\log_2 d$ [51], and due to the operational fact that the distillable entanglement of a state cannot exceed its entanglement cost [1], we conclude that $E_C(\mathcal{E}^q_{A\to B}(\Phi_{RA})) = (1-q)\log_2 d$, and in turn that

$$E_C(\mathcal{E}^q) = E_C^{(p)}(\mathcal{E}^q) = (1-q)\log_2 d. \quad (59)$$

This result generalizes the finding from [19], which is that $E_C^{(p)}(\mathcal{E}^q) = (1-q)\log_2 d$, and so we conclude that, for erasure channels, the entanglement cost of these channels is not increased by the need to pass a more stringent test for channel simulation, as posed by a sequential test. Note also that the distillable entanglement of the erasure channel is given by $E_D(\mathcal{E}^q) = (1-q)\log_2 d$, due to [51].

The fact that the distillable entanglement of an erasure channel is equal to its entanglement cost implies that, if we restrict the resource theory of entanglement for quantum channels to consist solely of erasure channels, then it is reversible. By this, we mean that, in the limit of many channel uses, if one begins with an erasure channel of parameter q and distills ebits from it at a rate $(1-q)\log_2 d$, then one can subsequently use these distilled ebits to simulate the same erasure channel again. As we see below, this reversibility breaks down when considering other channels.

B. Dephasing channels

A d-dimensional dephasing channel has the following action:

$$\mathcal{D}^{\mathbf{q}}(\rho) = \sum_{i=0}^{d-1} q_i Z^i \rho Z^{i\dagger}, \quad (60)$$

where $\mathbf{q}$ is a vector containing the probabilities q_i and Z has the following action on the computational basis $Z|x\rangle = e^{2\pi i x/d}|x\rangle$. This channel is covariant with respect to the Heisenberg-Weyl group of unitaries, which are well known to be a unitary one design. Furthermore, as remarked previously (e.g., in [52]), the Choi state $\mathcal{D}^{\mathbf{q}}_{A\to B}(\Phi_{RA})$ of this channel is a maximally correlated state [10,11], which has the form

$$\sum_{i,j} \alpha_{i,j} |i\rangle\langle j|_R \otimes |i\rangle\langle j|_B. \tag{61}$$

As such, Theorem 1 applies to these channels, implying that

$$E_C(\mathcal{D}^{\mathbf{q}}) = E_C^{(p)}(\mathcal{D}^{\mathbf{q}}) = E_C\big(\mathcal{D}^{\mathbf{q}}_{A\to B}(\Phi_{RA})\big) \tag{62}$$

$$= E_F\big(\mathcal{D}^{\mathbf{q}}_{A\to B}(\Phi_{RA})\big), \tag{63}$$

with the final equality resulting from the fact that the entanglement cost is equal to the entanglement of formation for maximally correlated states [53,54]. In Ref. [54], Sec. VI-A, an optimization procedure is given for calculating the entanglement of formation of maximally correlated states, which is simpler than that needed from the definition of entanglement of formation.

A qubit dephasing channel with a single dephasing parameter $q \in [0, 1]$ is defined as

$$\mathcal{D}^q(\rho) = (1-q)\rho + qZ\rho Z. \tag{64}$$

For the Choi state of this channel, there is an explicit formula for its entanglement of formation [55], from which we can conclude that

$$E_C(\mathcal{D}^q) = E_C^{(p)}(\mathcal{D}^q) = h_2[1/2 + \sqrt{q(1-q)}], \tag{65}$$

where

$$h_2(x) \equiv -x\,\log_2 x - (1-x)\log_2(1-x) \tag{66}$$

is the binary entropy. The equality in (65) solves an open question from [19], where it had only been shown that $E_C^{(p)}(\mathcal{D}^q) \leqslant h_2[1/2 + \sqrt{q(1-q)}]$.

The results of Ref. [1], Eq. (57) and Ref. [3], Eq. (8.114) gave a simple formula for the distillable entanglement of the qubit dephasing channel:

$$E_D(\mathcal{D}^q) = 1 - h_2(q). \tag{67}$$

Thus this formula and the formula in (65) demonstrate that the resource theory of entanglement for these channels is irreversible. That is, if one started from a qubit dephasing channel with parameter $q \in (0, 1)$ and distilled ebits from it at the ideal rate of $1 - h_2(q)$, and then subsequently wanted to use these ebits to simulate a qubit dephasing channel with the same parameter, this is not possible, because the rate at which ebits are distilled is not sufficient to simulate the channel again. Figure 3 compares the formulas for entanglement cost and distillable entanglement of the qubit dephasing channel, demonstrating that there is a noticeable gap between them. At $q = 1/2$, the qubit dephasing channel is a completely dephasing, classical channel, so that $E_C(\mathcal{D}^{1/2}) = E_D(\mathcal{D}^{1/2}) = 0$. Thus a reasonable approximation to the difference is given by a Taylor expansion about $q = 1/2$:

$$E_C(\mathcal{D}^q) - E_D(\mathcal{D}^q)$$
$$= \frac{1}{\ln 2}\left[2\,\ln\left(\frac{1}{|q-\frac{1}{2}|}\right) - 1\right]\left(q-\frac{1}{2}\right)^2 + O\left(\left(q-\frac{1}{2}\right)^4\right). \tag{68}$$

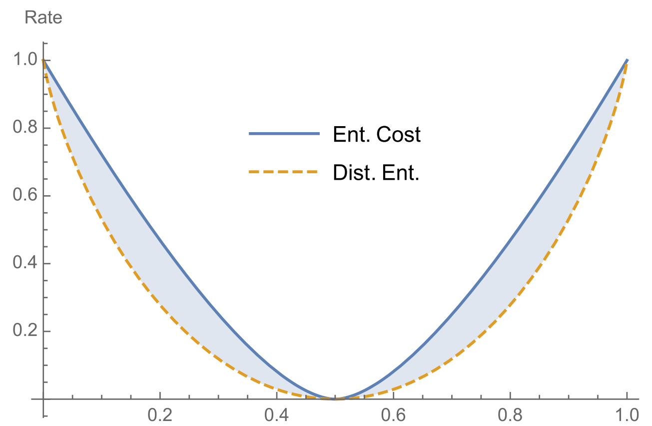

FIG. 3. Entanglement cost $E_C(\mathcal{D}^q) = h_2[1/2 + \sqrt{q(1-q)}]$ and distillable entanglement $E_D(\mathcal{D}^q) = 1 - h_2(q)$ of the qubit dephasing channel $\mathcal{D}^q$ as a function of the dephasing parameter $q \in [0, 1]$, with the shaded area demonstrating the gap between them. The units for rate on the vertical axis are ebits per channel use and q on the horizontal axis is dimensionless.

C. Werner-Holevo channels

A particular kind of Werner-Holevo channel performs the following transformation on a d-dimensional input state ρ [56]:

$$\mathcal{W}^{(d)}(\rho) \equiv \frac{1}{d-1}[\mathrm{Tr}\{\rho\}I - T(\rho)], \tag{69}$$

where T denotes the transpose map $T(\cdot) = \sum_{i,j} |i\rangle\langle j|(\cdot)|i\rangle\langle j|$. As observed in Ref. [56], Section II and Ref. [57], Section VII, this channel is covariant, and so an immediate consequence of Ref. [46], Section 7 is that these channels are teleportation simulable with associated resource state given by their Choi state. The latter fact was explicitly observed in Ref. [57], Sections VI and VII, as well as Ref. [43], Appendix A. Furthermore, its Choi state is given by

$$\mathcal{W}^{(d)}_{A\to B}(\Phi_{RA}) = \alpha_d \equiv \frac{1}{d(d-1)}(I_{RB} - F_{RB}), \tag{70}$$

where α_d is the antisymmetric state, i.e., the maximally mixed state on the antisymmetric subspace of a $d \times d$ quantum system and $F_{RB} \equiv \sum_{i,j} |i\rangle\langle j|_R \otimes |j\rangle\langle i|_B$ denotes the unitary swap operator. Theorem 1 thus applies to these channels, and we find that

$$E_C(\mathcal{W}^{(d)}) = E_C^{(p)}(\mathcal{W}^{(d)}) \tag{71}$$

$$= E_C(\alpha_d) \tag{72}$$

$$\geqslant \log_2(4/3) \approx 0.415, \tag{73}$$

with the inequality following from Ref. [58], Theorem 2. We also have that

$$E_C(\mathcal{W}^{(d)}) = E_C^{(p)}(\mathcal{W}^{(d)}) = E_C(\alpha_d) \leqslant E_F(\alpha_d) = 1, \tag{74}$$

with the last equality following from the result stated in Ref. [59], Section IV-C. For $d = 3$, the entanglement cost $E_C(\alpha_3)$ is known to be equal to exactly one ebit [60]:

$$E_C(\mathcal{W}^{(3)}) = E_C^{(p)}(\mathcal{W}^{(3)}) = 1. \tag{75}$$

It was observed in Ref. [43], Appendix A (as well as [61]) that the distillable entanglement of the Werner-Holevo channel $\mathcal{W}^{(d)}$ is equal to the distillable entanglement of its Choi state:

$$E_D(\mathcal{W}^{(d)}) = E_D(\alpha_d). \tag{76}$$

Thus, an immediate consequence of Ref. [58], Theorem 1 and Eq. (5), is that

$$E_D(\mathcal{W}^{(d)}) \leqslant \begin{Bmatrix} \log_2 \frac{d+2}{d} & \text{if } d \text{ is even} \\ \frac{1}{2}\log_2 \frac{d+3}{d-1} & \text{if } d \text{ is odd} \end{Bmatrix} \tag{77}$$

$$= \frac{2}{d \ln 2}\left(1 - \frac{1}{d}\right) + O\left(\frac{1}{d^3}\right). \tag{78}$$

We can now observe that the resource theory of entanglement is generally not reversible when restricted to Werner-Holevo channels. The case $d = 2$ is somewhat trivial: in this case, one can verify that the channel $\mathcal{W}^{(2)}$ is a unitary channel, equivalent to acting on the input state with the Pauli Y unitary. Thus, for $d = 2$, the channel is a noiseless qubit channel, and we trivially have that

$$E_D(\mathcal{W}^{(2)}) = E_C(\mathcal{W}^{(2)}) = 1, \tag{79}$$

so that the resource theory of entanglement is clearly reversible in this case. For $d = 3$, the upper bound on distillable entanglement in (77) evaluates to $\frac{1}{2}\log_2(3) \approx 0.793$, while the entanglement cost is equal to one, as stated in (75), so that

$$E_D(\mathcal{W}^{(3)}) \leqslant 0.793 < 1 = E_C(\mathcal{W}^{(3)}). \tag{80}$$

Thus the resource theory of entanglement is not reversible for $\mathcal{W}^{(3)}$. For $d \in \{4, 5, 6\}$, the upper bound in (77) and the lower bound in (73) are not strong enough to make a definitive statement [interestingly, the bounds in (77) and (73) are actually equal for $d = 6$]. Then, for $d \geqslant 7$, the upper bound in (77) and the lower bound in (73) are strong enough to conclude that

$$E_D(\mathcal{W}^{(d)}) < E_C(\mathcal{W}^{(d)}), \tag{81}$$

so that the resource theory is not reversible for $\mathcal{W}^{(d)}$. Figure 4 summarizes these observations.

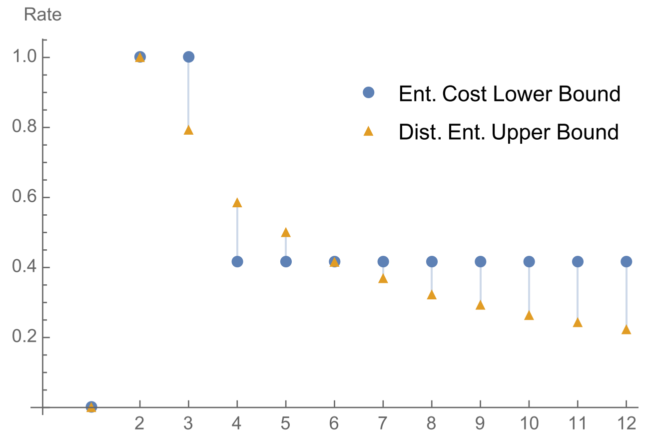

FIG. 4. Lower bound on the entanglement cost $E_C(\mathcal{W}^{(d)})$ from (73) and upper bound on distillable entanglement $E_D(\mathcal{W}^{(d)})$ from (77) for the Werner-Holevo channel $\mathcal{W}^{(d)}$ as a function of the parameter $d \geqslant 4$, with the lines connecting the dots demonstrating the gap between them. For $d = 2$, the points are exact due to (79), and reversibility holds. For $d = 3$, the entanglement cost $E_C(\mathcal{W}^{(3)})$ is exactly equal to one, as recalled in (75), while (77) applies to $E_D(\mathcal{W}^{(3)})$, and the resource theory is irreversible. For $d \in \{4, 5, 6\}$, the bounds are not strong enough to reach a conclusion about reversibility. For $d \geqslant 7$, the resource theory is irreversible, and the gap $E_C(\mathcal{W}^{(d)}) - E_D(\mathcal{W}^{(d)})$ grows at least as large as the difference of (73) and (78). The units for rate on the vertical axis are ebits per channel use and d on the horizontal axis is dimensionless.

D. Epolarizing channels (complements of depolarizing channels)

The d-dimensional depolarizing channel is a common model of noise in quantum information, transmitting the input state with probability $1 - q \in [0, 1]$ and replacing it with the maximally mixed state $\pi \equiv \frac{I}{d}$ with probability q:

$$\Delta^q(\rho) = (1 - q)\rho + q\pi. \tag{82}$$

According to Stinespring's theorem [62], every quantum channel $\mathcal{N}_{A\to B}$ can be realized by the action of some isometric channel $\mathcal{U}_{A\to BE}$ followed by a partial trace:

$$\mathcal{N}_{A\to B}(\rho_A) = \mathrm{Tr}_E\{\mathcal{U}_{A\to BE}(\rho_A)\}. \tag{83}$$

Due to the partial trace and its invariance with respect to isometric channels acting exclusively on the E system, the extending channel $\mathcal{U}_{A\to BE}$ is not unique in general, but it is unique up to this freedom. Then given an isometric channel $\mathcal{U}_{A\to BE}$ extending $\mathcal{N}_{A\to B}$ as in (83), the complementary channel $\mathcal{N}^c_{A\to E}$ is defined by a partial trace over the system B and is interpreted physically as the channel from the input to the environment:

$$\mathcal{N}^c_{A\to E}(\rho_A) = \mathrm{Tr}_B\{\mathcal{U}_{A\to BE}(\rho_A)\}. \tag{84}$$

Due to the fact that properties of the original channel are related to properties of its complementary channel [63,64], there has been significant interest in understanding complementary channels. In this spirit, and due to the prominent role of the depolarizing channel, researchers have studied its complementary channels [65,66]. In Ref. [65], Eq. (3.6), the following form was given for a complementary channel of Δ^q:

$$\rho \to S^q_{AF}(\rho_A \otimes I_F)S^{q\dagger}_{AF}, \tag{85}$$

where I_F is a d-dimensional identity operator and

$$S^q_{AF} \equiv \sqrt{\frac{q}{d}} I_{AF} + \sqrt{d}\left(-\frac{\sqrt{q}}{d} + \sqrt{1 - q\left(\frac{d^2-1}{d^2}\right)}\right)\Phi_{AF}. \tag{86}$$

A channel complementary to Δ^q has been called an "epolarizing channel" in [66].

An alternative complementary channel, related to the above one by an isometry acting on the output systems AF, but perhaps more intuitive, is realized in the following way [66, Eq. (28)]. Consider the isometry $U_{A\to SG_1G_2A}$ defined as

$$U_{A\to SG_1G_2A}|\psi\rangle_A \equiv \text{C-SWAP}_{SG_1A}(|\phi^q\rangle_S \otimes |\Phi\rangle_{G_1G_2} \otimes |\psi\rangle_A), \tag{87}$$

where the control qubit $|\phi^q\rangle_S \equiv \sqrt{1-q}|0\rangle_S + \sqrt{q}|1\rangle_S$, $|\Phi\rangle_{G_1G_2}$ is a maximally entangled state of Schmidt rank d, and the controlled-SWAP unitary is given by

$$\text{C-SWAP}_{SG_1A} \equiv |0\rangle\langle 0|_S \otimes I_{G_1A} + |1\rangle\langle 1|_S \otimes \text{SWAP}_{G_1A}, \tag{88}$$

with SWAP_{G_1A} denoting a unitary swap operation. By tracing over the systems SG_1G_2, we recover the original depolarizing channel

$$\Delta^q(\rho_A) = \text{Tr}_{SG_1G_2}\{U\rho_A U^\dagger\}. \tag{89}$$

Thus, by definition, a channel complementary to Δ^q is realized by

$$\Lambda^q_{A\to SG_1G_2}(\rho_A) \equiv \text{Tr}_A\{U\rho_A U^\dagger\}, \tag{90}$$

and, in what follows, let us refer to $\Lambda^q_{A\to SG_1G_2}$ as *the* epolarizing channel.

The isometry $U_{A\to SG_1G_2A}$ in (87) is unitarily covariant, in the sense that for an arbitrary unitary V_A acting on the input, we have that

$$U_{A\to SG_1G_2A}V_A = \left(V_{G_1}\otimes \overline{V}_{G_2}\otimes V_A\right)U_{A\to SG_1G_2A}, \tag{91}$$

where $\overline{V}$ denotes the complex conjugate of V. The identity in (91) follows because

$$\begin{aligned} &U_{A\to SG_1G_2A}V_A|\psi\rangle_A \\ &= \text{C-SWAP}_{SG_1A}\left(|\phi^q\rangle_S|\Phi\rangle_{G_1G_2}V_A|\psi\rangle_A\right) \\ &= \text{C-SWAP}_{SG_1A}\left[|\phi^q\rangle_S\left(V_{G_1}\overline{V}_{G_2}\right)|\Phi\rangle_{G_1G_2}V_A|\psi\rangle_A\right] \\ &= \left(V_{G_1}\otimes\overline{V}_{G_2}\otimes V_A\right)\text{C-SWAP}_{SG_1A}\left(|\phi^q\rangle_S|\Phi\rangle_{G_1G_2}|\psi\rangle_A\right) \\ &= \left(V_{G_1}\otimes\overline{V}_{G_2}\otimes V_A\right)U_{A\to SG_1G_2A}|\psi\rangle_A. \end{aligned} \tag{92}$$

The above analysis omits some tensor-product symbols for brevity. The third equality uses the well known fact that $|\Phi\rangle_{G_1G_2} = (V_{G_1}\otimes\overline{V}_{G_2})|\Phi\rangle_{G_1G_2}$. In the fourth equality, we have exploited the facts that $\overline{V}_{G_2}$ commutes with C-SWAP_{SG_1A} and that

$$\text{SWAP}_{G_1A}\left(V_{G_1}\otimes V_A\right) = \left(V_{G_1}\otimes V_A\right)\text{SWAP}_{G_1A}. \tag{93}$$

The covariance in (91) then implies that the epolarizing channel is covariant in the following sense:

$$\left(\Lambda^q_{A\to SG_1G_2}\circ\mathcal{V}_A\right)(\rho_A) = \left[\left(\mathcal{V}_{G_1}\otimes\overline{\mathcal{V}}_{G_2}\right)\circ\Lambda^q_{A\to SG_1G_2}\right](\rho_A), \tag{94}$$

where $\mathcal{V}$ denotes the unitary channel realized by the unitary operator V.

As such, by the discussion after (54), the epolarizing channel is a resource-seizable, teleportation-simulable channel with associated resource state given by $\Lambda^q_{A\to SG_1G_2}(\Phi_{RA})$. Thus Theorem 1 applies to these channels, implying that the first two of the following equalities hold:

$$\begin{aligned} E_C(\Lambda^q) = E_C^{(p)}(\Lambda^q) &= E_C(\Lambda^q(\Phi_{RA})) &(95)\\ &= E_F(\Lambda^q(\Phi_{RA})) &(96)\\ &= -\left(1-q+\frac{q}{d}\right)\log_2\left(1-q+\frac{q}{d}\right) \\ &\quad -(d-1)\frac{q}{d}\log_2\left(\frac{q}{d}\right). &(97)\end{aligned}$$

Let us now justify the final two equalities, which give a simple formula for the entanglement cost of epolarizing channels. First, consider that the Choi state $\Lambda^q_{A'\to SG_1G_2}(\Phi_{A'A})$ of the epolarizing channel is equal to the state resulting from sending in the maximally mixed state to the isometric channel $\mathcal{U}_{A\to SG_1G_2A}$, defined from (87):

$$\Lambda^q_{A'\to SG_1G_2}(\Phi_{A'A}) = \mathcal{U}_{A\to SG_1G_2A}(\pi_A), \tag{98}$$

where system A' is isomorphic to A. This equality is shown in Appendix B. As such, then Ref. [67], Theorem 3 applies, as discussed in Example 6 therein, and as a consequence we can conclude the second and third equalities in the following, with the bipartite cut of systems taken as $SG_1G_2|A$:

$$\begin{aligned} E_C\left(\Lambda^q_{A'\to SG_1G_2}(\Phi_{A'A})\right) &= E_C(\mathcal{U}_{A\to SG_1G_2A}(\pi_A)) &(99)\\ &= E_F\left(\mathcal{U}_{A\to SG_1G_2A}(\pi_A)\right) &(100)\\ &= H_{\min}(\Delta^q). &(101)\end{aligned}$$

The last line features the minimum output entropy of the depolarizing channel, which was identified in [68] and shown to be equal to (97).

As discussed in previous examples, it is worthwhile to consider the reversibility of the resource theory of entanglement for epolarizing channels. In this spirit, by invoking the covariance of Λ^q, the discussion after (54), Ref. [1], Eq. (55), and Ref. [11], Theorem 4.13, we find the following bound on the distillable entanglement of the epolarizing channel Λ^q:

$$E_D(\Lambda^q) \leqslant R(A;SG_1G_2)_{\Lambda^q(\Phi)}, \tag{102}$$

where $R(A;SG_1G_2)_{\Lambda^q(\Phi)}$ denotes the Rains relative entropy of the state $\Lambda^q_{A'\to SG_1G_2}(\Phi_{A'A})$. Recall that the Rains relative entropy for an arbitrary state ρ_{AB} is defined as [11]

$$R(A;B)_\rho \equiv \min_{\tau_{AB}\in\text{PPT}'(A;B)} D(\rho_{AB}\|\tau_{AB}), \tag{103}$$

where the quantum relative entropy is defined as [69]

$$D(\rho\|\tau) \equiv \text{Tr}\{\rho[\log_2\rho - \log_2\tau]\} \tag{104}$$

and the Rains set $\text{PPT}'(A;B)$ [70] is given by

$$\text{PPT}'(A;B) \equiv \{\tau_{AB} : \tau_{AB} \geqslant 0 \wedge \|T_B(\tau_{AB})\|_1 \leqslant 1\}, \tag{105}$$

with T_B denoting the partial transpose. Appendix C details a Matlab program taking advantage of recent advances in [71,72], in order to compute the Rains relative entropy of any bipartite state.

Figure 5 plots the entanglement cost of the epolarizing channel for $d=2$ (qubit input), and it also plots the Rains bound on distillable entanglement in (102). There is a gap for every value of $q\in(0,1)$, demonstrating that the resource theory of entanglement is irreversible for epolarizing channels. The figure also plots the coherent information of the state $\Lambda^q_{A'\to SG_1G_2}(\Psi^s_{A'A})$, optimized with respect to $|\Psi^s\rangle_{A'A} \equiv \sqrt{s}|00\rangle_{A'A} + \sqrt{1-s}|11\rangle_{A'A}$ for $s\in[0,1]$, which is known to be a lower bound on the distillable entanglement of Λ^q [9]. Note that the coherent information plot is not in contradiction with the recent result of [66], which states that the coherent information is strictly greater than zero for all $q\in(0,1]$. It is simply that the coherent information is so small for $q \lesssim 0.18$, that it is difficult to witness its strict positivity numerically.

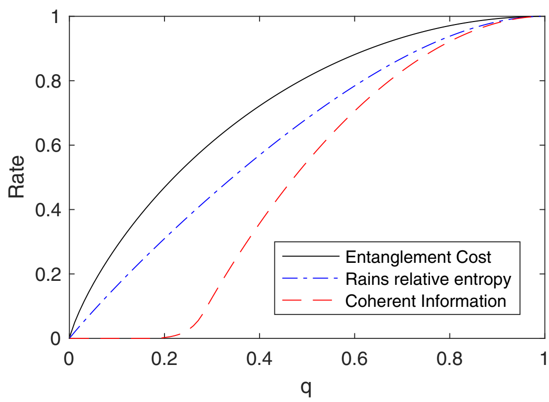

FIG. 5. Figure depicts the entanglement cost, the Rains bound, and the coherent information of the epolarizing channel Λ^q, for $d = 2$ and $q \in [0, 1]$. The gap between the entanglement cost and the Rains bound for all $q \in (0, 1)$ demonstrates that the resource theory of entanglement is irreversible for epolarizing channels. The units for rate on the vertical axis are ebits per channel use and q on the horizontal axis is dimensionless.

Matlab files to generate Fig. 5 are available with the arXiv posting of this paper.

V. BOSONIC GAUSSIAN CHANNELS

In this section, I extend the main ideas of the paper in order to characterize the entanglement cost of all single-mode bosonic Gaussian channels [73]. From a practical perspective, we should be most interested in the single-mode thermal, amplifier, and additive-noise channels, as these are of the greatest interest in applications, as stressed in Ref. [2], Sec. 12.6.3, and Ref. [74], Sec. 3.5. However, it also turns out that these are the only nontrivial cases to consider among all single-mode bosonic Gaussian channels, as discussed below.

A. On the definition of entanglement cost for infinite-dimensional channels

Before beginning, let us note that there are some subtleties involved when dealing with quantum information theory in infinite-dimensional Hilbert spaces [2]. For example, as advised in [75], the direct use of the diamond norm in infinite-dimensional Hilbert spaces could be too strong for applications, and this observation has motivated some recent work [76,77] on modifications of the diamond norm that take into account physical constraints such as energy limitations. On the other hand, the recent findings in [78] suggest that the direct use of the diamond norm is reasonable when considering single-mode thermal, amplifier, and additive-noise channels, as well as some multimode bosonic Gaussian channels. As it turns out, we can indeed directly employ the diamond norm when analyzing the entanglement cost of these channels. In fact, one of the main contributions of [78] was to consider uniform convergence issues in the teleportation simulation of bosonic Gaussian channels, and, due to the fact that the operational framework of entanglement cost is directly related to the approximate teleportation simulation of a channel, one should expect that the findings of [78] would be related to the issues involved in the entanglement cost of bosonic Gaussian channels.

With this in mind, let us define the entanglement cost for an infinite-dimensional channel almost exactly as it has been defined in Sec. II B, with the exception that we allow for LOCC channels that have a continuous classical index (e.g., as considered in Ref. [79], Section 4), thus going beyond the LOCC channels considered in (4). Specifically, let us define an (n, M, ε) sequential channel simulation code as it has been defined in Sec. II B, noting that the ε-error criterion is given by (18), representing the direct generalization of the strategy norm of [26,27,33] to infinite-dimensional systems. Achievable rates and the entanglement cost are then defined in the same way.

B. Preliminary observations about the entanglement cost of single-mode bosonic Gaussian channels

The starting point for our analysis of single-mode bosonic Gaussian channels is the Holevo classification from [80], in which canonical forms for all single-mode bosonic Gaussian channels have been given, classifying them up to local Gaussian unitaries acting on the input and output of the channel. It then suffices for us to focus our attention on the canonical forms, as it is self-evident from definitions that local unitaries do not alter the entanglement cost of a quantum channel. The thermal and amplifier channels form the class C discussed in [80], and the additive-noise channels form the class B_2 discussed in the same work. The classes that remain are labeled A, B_1, and D in [80]. The channels in A and D are entanglement breaking [39], and as a consequence of the "if part" of Remark 1, they have zero entanglement cost. Channels in the class B_1 are perhaps not interesting for practical applications, and as it turns out, they have infinite quantum capacity [80]. Thus their entanglement cost is also infinite, because a channel's quantum capacity is a lower bound on its distillable entanglement, which is in turn a lower bound on its entanglement cost—these relationships are a direct consequence of the definitions of the underlying quantities. For the same reason, the entanglement cost of the bosonic identity channel is also infinite.

C. Thermal, amplifier, and additive-noise channels

In light of the previous discussion, for the remainder of the paper, let us focus our attention on the thermal, amplifier, and additive-noise channels. Each of these are defined respectively by the following Heisenberg input-output relations:

$$\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}, \tag{106}$$

$$\hat{b} = \sqrt{G}\hat{a} + \sqrt{G-1}\hat{e}^{\dagger}, \tag{107}$$

$$\hat{b} = \hat{a} + (x + ip)/\sqrt{2}, \tag{108}$$

where $\hat{a}$, $\hat{b}$, and $\hat{e}$ are the field-mode annihilation operators for the sender's input, the receiver's output, and the environment's input of these channels, respectively.

The channel in (106) is a thermalizing channel, in which the environmental mode is prepared in a thermal state $\theta(N_B)$

of mean photon number $N_B \geqslant 0$, defined as

$$\theta(N_B) \equiv \frac{1}{N_B+1}\sum_{n=0}^{\infty}\left(\frac{N_B}{N_B+1}\right)^n |n\rangle\langle n|, \tag{109}$$

where $\{|n\rangle\}_{n=0}^{\infty}$ is the orthonormal, photonic number-state basis. When $N_B = 0$, $\theta(N_B)$ reduces to the vacuum state, in which case the resulting channel in (106) is called the pure-loss channel—it is said to be quantum-limited in this case because the environment is injecting the minimum amount of noise allowed by quantum mechanics. The parameter $\eta \in (0,1)$ is the transmissivity of the channel, representing the average fraction of photons making it from the input to the output of the channel. Let $\mathcal{L}_{\eta,N_B}$ denote this channel, and we make the further abbreviation $\mathcal{L}_\eta \equiv \mathcal{L}_{\eta,N_B=0}$ when it is the pure-loss channel. The channel in (106) is entanglement-breaking when $(1-\eta)N_B \geqslant \eta$ [39] and, by Remark 1, the entanglement cost is equal to zero for these values.

The channel in (107) is an amplifier channel and the parameter $G > 1$ is its gain. For this channel, the environment is prepared in the thermal state $\theta(N_B)$. If $N_B = 0$, the amplifier channel is called the pure-amplifier channel—it is said to be quantum limited for a similar reason as stated above. Let $\mathcal{A}_{G,N_B}$ denote this channel, and we make the further abbreviation $\mathcal{A}_G \equiv \mathcal{A}_{G,N_B=0}$ when it is the quantum-limited amplifier channel. The channel in (107) is entanglement-breaking when $(G-1)N_B \geqslant 1$ [39] and, by Remark 1, the entanglement cost is equal to zero for these values.

Finally, the channel in (108) is an additive-noise channel, representing a quantum generalization of the classical additive white Gaussian noise channel. In (108), x and p are zero-mean, independent Gaussian random variables each having variance $\xi \geqslant 0$. Let $\mathcal{T}_\xi$ denote this channel. The channel in (108) is entanglement breaking when $\xi \geqslant 1$ [39] and, by Remark 1, the entanglement cost is equal to zero for these values.

Kraus representations for the channels in (106)–(108) are available in [81], which can be helpful for further understanding their action on input quantum states.

Due to the entanglement-breaking regions discussed above, we are left with a limited range of single-mode bosonic Gaussian channels to consider, which is delineated by the white strip in Fig. 1 of [82].

D. Upper bound on the entanglement cost of teleportation-simulable channels with bosonic Gaussian resource states

In this section, I determine an upper bound on the entanglement cost of any channel $\mathcal{N}_{A\to B}$ that is teleportation simulable with associated resource state given by a bosonic Gaussian state. Related bosonic teleportation channels have been considered previously [83–88], in the case that the LOCC channel associated to $\mathcal{N}_{A\to B}$ is a Gaussian LOCC channel. Proposition 3 below states that the entanglement cost of these channels is bounded from above by the Gaussian entanglement of formation [89] of the underlying bosonic Gaussian resource state and, as such, this proposition represents a counterpart to Proposition 2. Before stating it, let us note that the Gaussian entanglement of formation $E_F^g(A;B)_\rho$ of a bipartite state ρ_{AB} [89] is given by the same formula as in (7), with the exception that the pure states ψ^x_{AB} in the ensemble decomposition are required to be Gaussian. Note that continuous probability measures are allowed for the decomposition (for an explicit definition, see Ref. [89], Sec. III). Let us note here that the first part of the proof outlines a procedure for the formation of n approximate copies of a bipartite state and, even though this kind of protocol has been implicit in prior literature, I have included explicit steps for clarity. After proving Proposition 2, I discuss its application to thermal, amplifier, and additive-noise bosonic Gaussian channels.

Proposition 3. Let $\mathcal{N}_{A\to B}$ be a channel that is teleportation simulable as defined in (28), where the resource state $\omega_{A'B'}$ is a bosonic Gaussian state composed of k modes for system A' and ℓ modes for system B', with $k,\ell \geqslant 1$. Then the entanglement cost of $\mathcal{N}_{A\to B}$ is never larger than the Gaussian entanglement of formation of the bosonic Gaussian resource state $\omega_{A'B'}$:

$$E_C(\mathcal{N}) \leqslant E_F^g(A';B')_\omega. \tag{110}$$

Proof. The main idea of the proof is to first form n approximate copies of the bosonic Gaussian resource state $\omega_{A'B'}$, by using entanglement and LOCC as related to the approach from [90] and then, after that, simulate n uses of the channel $\mathcal{N}_{A\to B}$ by employing the structure of the channel $\mathcal{N}_{A\to B}$ from (28). Indispensable to the proof is the analysis in Ref. [89], Secs. II and III, where it is shown that every bosonic Gaussian state can be decomposed as a Gaussian mixture of local displacements acting on a fixed Gaussian pure state and that such a decomposition is optimal for the Gaussian entanglement of formation (Ref. [89], Proposition 1). The Gaussian mixture of local displacements can be understood as an LOCC channel $\mathcal{G}_{A'B'}$, and let $\psi^\omega_{A'B'}$ denote the aforementioned fixed Gaussian pure state such that $\mathcal{G}_{A'B'}(\psi^\omega_{A'B'}) = \omega_{A'B'}$.

Since $\psi^\omega_{A'B'}$ is Gaussian, the marginal state $\psi^\omega_{B'}$ is Gaussian, and thus it has finite entropy $H(B')_{\psi^\omega}$, as well as finite entropy variance, i.e.,

$$V(B')_{\psi^\omega} \equiv \mathrm{Tr}\{\psi^\omega_{B'}[-\log_2\psi^\omega_{B'} - H(B')_{\psi^\omega}]^2\} < \infty, \tag{111}$$

the latter statement following from the Williamson decomposition [91] for Gaussian states as well as the formula for the entropy variance of a bosonic thermal state [92]. For $\delta > 0$, recall that the entropy-typical projector $\Pi^\delta_{B'^n}$ [93,94] of the state $\psi^\omega_{B'}$ is defined as the projection onto

$$\mathrm{span}\{|\xi_{z^n}\rangle : |-n^{-1}\log_2(p_{Z^n}(z^n)) - H(B')_{\psi^\omega}| \leqslant \delta\}, \tag{112}$$

where a countable spectral decomposition of $\psi^\omega_{B'}$ is given by

$$\psi^\omega_{B'} = \sum_z p_Z(z)|\xi_z\rangle\langle\xi_z| \tag{113}$$

and

$$|\xi_{z^n}\rangle \equiv |\xi_{z_1}\rangle \otimes \cdots \otimes |\xi_{z_n}\rangle, \tag{114}$$

$$p_{Z^n}(z^n) \equiv p_Z(z_1)\cdots p_Z(z_n). \tag{115}$$

The entropy-typical projector $\Pi^\delta_{B'^n}$ projects onto a finite-dimensional subspace of $[\psi^\omega_{B'}]^{\otimes n}$, and satisfies the conditions

$[\Pi^{\delta}_{B'^n}, [\psi^{\omega}_{B'}]^{\otimes n}] = 0$ and

$$2^{-n[H(B')_{\psi^\omega}+\delta]}\Pi^{\delta}_{B'^n} \leqslant \Pi^{\delta}_{B'^n}[\psi^{\omega}_{B'}]^{\otimes n}\Pi^{\delta}_{B'^n} \leqslant 2^{-n[H(B')_{\psi^\omega}-\delta]}\Pi^{\delta}_{B'^n}. \tag{116}$$

It then follows that $\mathrm{Tr}\{\Pi^{\delta}_{B'^n}\} \leqslant 2^{n[H(B')_{\psi^\omega}+\delta]}$. Furthermore, consider that the entropy-typical projector $\Pi^{\delta}_{B'^n}$ for the state $[\psi^{\omega}_{B'}]^{\otimes n}$ satisfies

$$\mathrm{Tr}\left\{\left(I_{A'^n}\otimes\Pi^{\delta}_{B'^n}\right)\left[\psi^{\omega}_{A'B'}\right]^{\otimes n}\right\} = \mathrm{Tr}\left\{\Pi^{\delta}_{B'^n}\left[\psi^{\omega}_{B'}\right]^{\otimes n}\right\} \geqslant 1-\frac{V(B')_{\psi^\omega}}{\delta^2 n}, \tag{117}$$

with the inequality following from the definition of the entropy-typical projector and an application of the Chebshev inequality. By the gentle measurement lemma [95,96] (see Ref. [4], Lemma 9.4.1 for the version employed here), we conclude that

$$\frac{1}{2}\left\|\left[\psi^{\omega}_{A'B'}\right]^{\otimes n}-\widetilde{\psi}^{\omega}_{A'^n B'^n}\right\|_1 \leqslant \sqrt{\frac{V(B')_{\psi^\omega}}{\delta^2 n}}, \tag{118}$$

where

$$\widetilde{\psi}^{\omega}_{A'^n B'^n} \equiv \frac{\left(I_{A'^n}\otimes\Pi^{\delta}_{B'^n}\right)\left[\psi^{\omega}_{A'B'}\right]^{\otimes n}\left(I_{A'^n}\otimes\Pi^{\delta}_{B'^n}\right)}{\mathrm{Tr}\left\{\left(I_{A'^n}\otimes\Pi^{\delta}_{B'^n}\right)\left[\psi^{\omega}_{A'B'}\right]^{\otimes n}\right\}}. \tag{119}$$

Observe that the system B'^n of $\widetilde{\psi}^{\omega}_{A'^n B'^n}$ is supported on a finite-dimensional subspace of B'^n.

Now, the idea of forming n approximate copies $\psi^{\omega}_{A'B'}$ is then the same as it is in [90]: Alice prepares the state $\widetilde{\psi}^{\omega}_{A'^n B'^n}$ locally, Alice and Bob require beforehand a maximally entangled state of Schmidt rank no larger than $2^{n[H(B')_{\psi^\omega}+\delta]}$, and then they perform quantum teleportation [14] to teleport the B'^n system to Bob. At this point, they share exactly the state $\widetilde{\psi}^{\omega}_{A'^n B'^n}$, which becomes less and less distinguishable from $[\psi^{\omega}_{A'B'}]^{\otimes n}$ as n grows large, due to (118). Now applying the Gaussian LOCC channel $(\mathcal{G}_{A'B'})^{\otimes n}$, the data processing inequality to (118), and the fact that $\mathcal{G}_{A'B'}(\psi^{\omega}_{A'B'}) = \omega_{A'B'}$, we conclude that

$$\frac{1}{2}\left\|\omega^{\otimes n}_{A'B'}-(\mathcal{G}_{A'B'})^{\otimes n}(\widetilde{\psi}^{\omega}_{A'^n B'^n})\right\|_1 \leqslant \sqrt{\frac{V(B')_{\psi^\omega}}{\delta^2 n}}. \tag{120}$$

Thus, to see that $H(B')_{\psi^\omega}$ is an achievable rate for forming $\omega^{\otimes n}_{A'B'}$, fix $\varepsilon\in(0,1]$ and $\delta>0$. Then choose n large enough so that $\sqrt{\frac{V(B')_{\psi^\omega}}{\delta^2 n}} \leqslant \varepsilon$. Apply the above procedure, using LOCC and a maximally entangled state of Schmidt rank no larger than $2^{n[H(B')_{\psi^\omega}+\delta]}$. Then the rate of entanglement consumption to produce n approximate copies of $\omega_{A'B'}$ satisfying (120) is $H(B')_{\psi^\omega}+\delta$. Since this is possible for $\varepsilon\in(0,1]$, $\delta>0$, and sufficiently large n, we conclude that $H(B')_{\psi^\omega}$ is an achievable rate for the formation of $\omega_{A'B'}$. Now, since achieving this rate is possible for any pure state $\psi^{\omega}_{A'B'}$ such that $\omega_{A'B'} = \mathcal{G}_{A'B'}(\psi^{\omega}_{A'B'})$, we conclude that the infimum of $H(B')_{\psi^\omega}$ with respect to all such pure states is an achievable rate. But this latter quantity is exactly the Gaussian entanglement of formation according to Ref. [89], Proposition 1.

The idea for simulating n uses of the channel $\mathcal{N}_{A\to B}$ is then the same as the idea used in the proof of Proposition 2. First form n approximate copies of $\omega_{A'B'}$ according to the procedure described above. Then, when the ith call to the channel $\mathcal{N}_{A\to B}$ is made, use the LOCC channel $\mathcal{L}_{AA'B'\to B}$ from the definition in (28) along with the ith A' and B' systems of the state approximating $\omega^{\otimes n}_{A'B'}$ to simulate it. By the same reasoning that led to (38), the distinguishability of the final states of any sequential test is limited by the distinguishability of the state $\omega^{\otimes n}_{A'B'}$ from its approximation, which I argued in (120) can be made arbitrarily small with increasing n. Thus the Gaussian entanglement of formation $\omega_{A'B'}$ is an achievable rate for sequential channel simulation of $\mathcal{N}_{A\to B}$. ■

1. *Upper bound for the entanglement cost of thermal, amplifier, and additive-noise bosonic Gaussian channels*

I now discuss how to apply Proposition 2 to single-mode thermal, amplifier, and additive-noise channels. Some recent papers [97–99] have shown how to simulate each of these channels by using a bosonic Gaussian resource state along with variations of the continuous-variable quantum teleportation protocol [83]. Of these works, the one most relevant for us is the latest one [99], because these authors proved that the entanglement of formation of the underlying resource state is equal to the entanglement of formation that results from transmitting through the channel one share of a two-mode squeezed vacuum state with arbitrarily large squeezing strength. That is, let $\mathcal{N}_{A\to B}$ denote a single-mode thermal, amplifier, or additive-noise channel. Then one of the main results of [99] is that, associated to this channel, there is a bosonic Gaussian resource state $\omega_{A'B'}$ and a Gaussian LOCC channel $\mathcal{G}_{AA'B'\to B}$ such that

$$E_F(A';B')_\omega = \sup_{N_S\geqslant 0} E_F(R;B)_{\sigma(N_S)} \tag{121}$$

$$= \lim_{N_S\to\infty} E_F(R;B)_{\sigma(N_S)}, \tag{122}$$

where

$$\sigma(N_S) \equiv \mathcal{N}_{A\to B}\left(\phi^{N_S}_{RA}\right), \tag{123}$$

$$\phi^{N_S}_{RA} \equiv |\phi^{N_S}\rangle\langle\phi^{N_S}|_{RA}, \tag{124}$$

$$|\phi^{N_S}\rangle_{RA} \equiv \frac{1}{\sqrt{N_S+1}}\sum_{n=0}^{\infty}\sqrt{\left(\frac{N_S}{N_S+1}\right)^n}|n\rangle_R|n\rangle_A, \tag{125}$$

and for all input states ρ_A,

$$\mathcal{N}_{A\to B}(\rho_A) = \mathcal{G}_{AA'B'\to B}(\rho_A\otimes\omega_{A'B'}). \tag{126}$$

In the above, $\phi^{N_S}_{RA}$ is the two-mode squeezed vacuum state [73]. Note that the equality in (122) holds because one can always produce $\phi^{N_S}_{RA}$ from $\phi^{N_S'}_{RA}$ such that $N_S'\geqslant N_S$, by using Gaussian LOCC and the local displacements involved in the Gaussian LOCC commute with the channel $\mathcal{N}_{A\to B}$ [100] (whether it be thermal, amplifier, or additive noise). Furthermore, the entanglement of formation does not increase under the action of an LOCC channel.

Thus, applying the above observations and Proposition 3, it follows that there exist bosonic Gaussian resource states $\omega^{\eta,N_B}_{A'B'}$, $\omega^{G,N_B}_{A'B'}$, and $\omega^{\xi}_{A'B'}$ associated to the respective thermal, amplifier, and additive-noise channels in (106)–(108), such that the following inequalities hold:

$$E_C(\mathcal{L}_{\eta,N_B}) \leqslant E_F(A';B')_{\omega^{\eta,N_B}}, \tag{127}$$

$$E_C(\mathcal{A}_{G,N_B}) \leqslant E_F(A';B')_{\omega^{G,N_B}}, \tag{128}$$

$$E_C(\mathcal{T}_\xi) \leqslant E_F(A';B')_{\omega^{\xi}}. \tag{129}$$

Analytical formulas for the upper bounds on the right can be found in Ref. [99], Eqs. (4)–(6).

E. Lower bound on the entanglement cost of bosonic Gaussian channels

In this section, I establish a lower bound on the nonasymptotic entanglement cost of thermal, amplifier, or additive-noise bosonic Gaussian channels. After that, I show how this bound implies a lower bound on the entanglement cost. Finally, by proving that the state resulting from sending one share of a two-mode squeezed vacuum through a pure-loss or pure-amplifier channel has entanglement cost equal to entanglement of formation, I establish the exact entanglement cost of these channels by combining with the results from the previous section.

Proposition 4. Let $\mathcal{N}_{A\to B}$ be a thermal, amplifier, or additive-noise channel, as defined in (106)–(108). Let $n, M \in \mathbb{N}$, $\varepsilon \in [0,1/2)$, $\varepsilon' \in (\sqrt{2\varepsilon}, 1]$, $\delta = [\varepsilon' - \sqrt{2\varepsilon}]/[1+\varepsilon']$, and $N_S \in [0,\infty)$. Then the following bound holds for any (n, M, ε) sequential or parallel channel simulation code for $\mathcal{N}_{A\to B}$:

$$\frac{1}{n}\log_2 M \geqslant \frac{1}{n}E_F(R^n;B^n)_{\omega^{\otimes n}} - (\varepsilon' + 2\delta)H\big(\phi_R^{N_S/\delta}\big) - \frac{1}{n}[2(1+\varepsilon')g_2(\varepsilon') + 2h_2(\delta)], \tag{130}$$

where $\omega_{RB} \equiv \mathcal{N}_{A\to B}(\phi_{RA}^{N_S})$ and $\frac{1}{n}\log_2 M$ is understood as the nonasymptotic entanglement cost of the protocol.

Proof. The reasoning here is very similar to that given in the proof of Proposition 1, but we can instead make use of the continuity bound for the entanglement of formation of energy-constrained states (Ref. [101], Proposition 5). To begin, suppose that there exists an (n, M, ε) protocol for sequential channel simulation. Then by previous reasoning (also see Fig. 2), it can be thought of as a parallel channel simulation protocol, such that the criterion in (2) holds. Let us take $(\phi_{RA}^{N_S})^{\otimes n}$ to be a test input state, leading to $\omega_{RB}^{\otimes n} = [\mathcal{N}_{A\to B}(\phi_{RA}^{N_S})]^{\otimes n}$ when the actual channels are applied and $\sigma_{R_1\cdots R_n B_1\cdots B_n}$ when the simulation is applied. Set

$$f(n,\varepsilon,\varepsilon',N_S) \equiv n(\varepsilon'+2\delta)H\big(\phi_R^{N_S/\delta}\big) + 2(1+\varepsilon')g_2(\varepsilon') + 2h_2(\delta). \tag{131}$$

Then we have that

$$\begin{aligned} &E_F(R^n;B^n)_{\omega^{\otimes n}} \\ &\leqslant E_F(R^n;B^n)_\sigma + f(n,\varepsilon,\varepsilon',N_S) \\ &\leqslant E_F(R^nA^n\overline{A}_0;\overline{B}_0)_{\psi\otimes\Phi} + f(n,\varepsilon,\varepsilon',N_S) \\ &= E_F(\overline{A}_0;\overline{B}_0)_\Phi + f(n,\varepsilon,\varepsilon',N_S) \\ &= \log_2 M + f(n,\varepsilon,\varepsilon',N_S). \end{aligned} \tag{132}$$

The first inequality follows from the condition in (18), as well as from the continuity bound for entanglement of formation from Ref. [101], Proposition 5, noting that the total photon number of the reduced (thermal) state on systems R^n is equal to nN_S. The second inequality follows from the LOCC monotonicity of the entanglement of formation, here thinking of the person who possesses systems RA^n to be in the same laboratory as the one possessing the systems $\overline{A}_i$, while the person who possesses the $\overline{B}_i$ systems is in a different laboratory. The first equality follows from the fact that $(\phi_{RA}^{N_S})^{\otimes n}$ is in tensor product with $\Phi_{\overline{A}_0\overline{B}_0}$, so that by a local channel, one may remove $(\phi_{RA}^{N_S})^{\otimes n}$ or append it for free. The final equality follows because the entanglement of formation of the maximally entangled state is equal to the logarithm of its Schmidt rank. ■

A direct consequence of Proposition 4 is the following lower bound on the entanglement cost of the thermal, amplifier, and additive-noise channels.

Proposition 5. Let $\mathcal{N}_{A\to B}$ be a thermal, amplifier, or additive-noise channel, as defined in (106)–(108). Then the entanglement costs $E_C(\mathcal{N})$ and $E_C^{(p)}(\mathcal{N})$ are bounded from below by the entanglement cost of the state $\mathcal{N}_{A\to B}(\phi_{RA}^{N_S})$, where the two-mode squeezed vacuum state $\phi_{RA}^{N_S}$ has arbitrarily large squeezing strength:

$$E_C(\mathcal{N}) \geqslant E_C^{(p)}(\mathcal{N}) \tag{133}$$

$$\geqslant \sup_{N_S\geqslant 0} E_C\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big) \tag{134}$$

$$= \lim_{N_S\to\infty} E_C\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big). \tag{135}$$

Proof. The first inequality follows from definitions, as argued previously in (22). To arrive at the second inequality, in Proposition 4, set $\varepsilon' = \sqrt[4]{2\varepsilon}$, and take the limit as $n\to\infty$ and then as $\varepsilon\to 0$. Employing the fact that $\lim_{\xi\to 0}\xi H(H(\phi_R^{N_S/\xi})) = 0$ (Ref. [102], Proposition 1) and applying definitions, we find for all $N_S \geqslant 0$ that

$$E_C(\mathcal{N}) \geqslant E_C^{(p)}(\mathcal{N}) \tag{136}$$

$$\geqslant \lim_{n\to\infty}\frac{1}{n}E_F\big(\big[\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big]^{\otimes n}\big) \tag{137}$$

$$= E_C\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big). \tag{138}$$

Since the above bound holds for all $N_S \geqslant 0$, we conclude the bound in the statement of the proposition. The equality in (135) follows for the same reason as given for the equality in (122), and due to the fact that entanglement cost is nonincreasing with respect to an LOCC channel by definition. ■

F. Additivity of entanglement of formation for pure-loss and pure-amplifier channels

The bound in Proposition 5 is really only a formal statement, as it is not clear how to evaluate the lower bound explicitly. If it would however be possible to prove that

$$\frac{1}{n}E_F\big(\big[\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big]^{\otimes n}\big) \stackrel{?}{=} E_F\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big) \tag{139}$$

for all integer $n \geqslant 1$ and all $N_S \geqslant 0$, then we could conclude the following:

$$E_C(\mathcal{N}) \overset{?}{\geqslant} \lim_{N_S\to\infty} E_F\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big), \quad (140)$$

implying that this lower bound coincides with the upper bound from (127)–(129), due to the recent result of [99] recalled in (121)–(122).

In Proposition 6 below, I prove that the additivity relation in (139) indeed holds whenever the channel $\mathcal{N}_{A\to B}$ is a pure-loss channel $\mathcal{L}_\eta$ or pure-amplifier channel $\mathcal{A}_G$. The linchpin of the proof is the multimode bosonic minimum output entropy theorem from Ref. [103] and Ref. [104], Theorem 1.

Proposition 6. For $\mathcal{N}_{A\to B}$ a pure-loss channel $\mathcal{L}_\eta$ with transmissivity $\eta \in (0,1)$ or a pure-amplifier channel $\mathcal{A}_G$ with gain $G > 1$, the following additivity relation holds for all integer $n \geqslant 1$ and $N_S \in [0,\infty)$:

$$\frac{1}{n}E_F\big(\big[\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big]^{\otimes n}\big) = E_F\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big) \quad (141)$$

$$= E_F^g\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big), \quad (142)$$

where $\phi_{RA}^{N_S}$ is the two-mode squeezed vacuum state from (125) and E_F^g denotes the Gaussian entanglement of formation. Thus the entanglement cost of $\mathcal{N}_{A\to B}(\phi_{RA}^{N_S})$ is equal to its entanglement of formation:

$$E_C\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big) = E_F\big(\mathcal{N}_{A\to B}\big(\phi_{RA}^{N_S}\big)\big). \quad (143)$$

Proof. The proof of this proposition relies on three key prior results.

(1) The main result of [105] is that the entanglement of formation $E_F(A;B)_\psi$ is equal to the classically conditioned entropy $H(A|\overline{E})_\psi$ for a tripartite pure state ψ_{ABE}:

$$E_F(A;B)_\psi = H(A|\overline{E})_\psi, \quad (144)$$

where

$$H(A|\overline{E})_\psi = \inf_{\{\Lambda_E^x\}_x} \sum_x p_X(x) H(A)_{\sigma^x}, \quad (145)$$

with the optimization taken with respect to a positive operator-valued measure $\{\Lambda_E^x\}_x$ and

$$p_X(x) \equiv \mathrm{Tr}\big\{\Lambda_E^x \psi_E\big\}, \quad (146)$$

$$\sigma_A^x \equiv \frac{1}{p_X(x)} \mathrm{Tr}_E\big\{\big(I_A \otimes \Lambda_E^x\big)\psi_{AE}\big\}. \quad (147)$$

The sum in (145) can be replaced with an integral for continuous-outcome measurements. The equality in (144) can be understood as being a consequence of the quantum steering effect [106].

(2) The determination of and method of proof for the classically conditioned entropy $H(A|\overline{E})_\rho$ of an arbitrary two-mode Gaussian state ρ_{AE} with covariance matrix in certain standard forms [107]. [As remarked below, there is in fact a significant strengthening of the main result of [107], which relies on item (3) below.]

(3) The multimode bosonic minimum output entropy theorem from Ref. [103] and Ref. [104], Theorem 1 (see the related work in [82,108] also), which implies that the following identity holds for a phase-insensitive, single-mode bosonic Gaussian channel $\mathcal{G}$ and for all integer $n \geqslant 1$:

$$\inf_{\rho^{(n)}} H(\mathcal{G}^{\otimes n}(\rho^{(n)})) = H(\mathcal{G}^{\otimes n}([|0\rangle\langle 0|]^{\otimes n}))$$

$$= nH(\mathcal{G}(|0\rangle\langle 0|)), \quad (148)$$

where the optimization is with respect to an arbitrary n-mode input state $\rho^{(n)}$ and $|0\rangle\langle 0|$ denotes the bosonic vacuum state.

Indeed, these three key ingredients, with the third being the linchpin, lead to the statement of the proposition after making a few observations. Consider that a purification of the state $\rho_{AB} = (\mathrm{id}_{R\to A} \otimes \mathcal{L}_\eta)(\phi_{RA}^{N_S})$ is given by

$$\psi_{ABE} = (\mathrm{id}_{R\to A} \otimes \mathcal{B}_{AE\to BE}^{\eta})(\phi_{RA}^{N_S} \otimes |0\rangle\langle 0|_E), \quad (149)$$

where $\mathcal{B}_{AE\to BE}^{\eta}$ represents the unitary for a beam-splitter interaction [73] and $|0\rangle\langle 0|_E$ again denotes the vacuum state. Tracing over the system B gives the state $\psi_{AE} = (\mathrm{id}_{R\to A} \otimes \mathcal{L}_{1-\eta})(\phi_{RA}^{N_S})$, where $\mathcal{L}_{1-\eta}$ is a pure-loss channel of transmissivity $1-\eta$. The state ψ_{AE} is well known to have its covariance matrix in standard form [73] [see discussion surrounding Ref. [107], Eq. (5)] as

$$\begin{bmatrix} a & 0 & c & 0 \\ 0 & a & 0 & -c \\ c & 0 & b & 0 \\ 0 & -c & 0 & b \end{bmatrix} \quad (150)$$

and is also known as a two-mode squeezed thermal state [73]. As such, the main result of [107] applies, and we can conclude that heterodyne detection is the optimal measurement in (145), which in turn implies from (144) that the entanglement of formation of ρ_{AB} is equal to the Gaussian entanglement of formation.

However, what we require is that the same results hold for the multicopy state $\psi_{AE}^{\otimes n}$. Inspecting Eqs. (9)–(14) of [107], it is clear that the same steps hold, except that we replace Eq. (12) therein with (148). Thus it follows that n individual heterodyne detections on each E mode of $\psi_{AE}^{\otimes n}$ is the optimal measurement, so that

$$\frac{1}{n}H(A^n|\overline{E^n})_{\psi^{\otimes n}} = H(A|\overline{E})_\psi. \quad (151)$$

By applying (144) (as applied to the states $\rho_{AB}^{\otimes n}$ and $\psi_{AE}^{\otimes n}$), we conclude that

$$\frac{1}{n}E_F(A^n;B^n)_{\rho^{\otimes n}} = E_F(A;B)_\rho. \quad (152)$$

Furthermore, since the optimal measurement is given by heterodyne detection, performing it on mode E of ψ_{ABE} induces a Gaussian ensemble of pure states $\{p_X(x), \psi_{AB}^x\}$, which is the optimal decomposition of $\psi_{AB} = \rho_{AB}$, and thus we conclude that $E_F(A;B)_\rho = E_F^g(A;B)_\rho$.

A similar analysis applies for the quantum-limited amplifier channel. I give the argument for completeness. Consider that a purification of the state $\sigma_{AB} = (\mathrm{id}_{R\to A} \otimes \mathcal{A}_G)(\phi_{RA}^{N_S})$ is given by

$$\varphi_{ABE} = \big(\mathrm{id}_{R\to A} \otimes \mathcal{S}_{AE\to BE}^{G}\big)\big(\phi_{RA}^{N_S} \otimes |0\rangle\langle 0|_E\big), \quad (153)$$

where $\mathcal{S}^G_{AE\to BE}$ represents the unitary for a two-mode squeezer [73] and $|0\rangle\langle 0|_E$ again denotes the vacuum state. Tracing over the system B gives the state $\varphi_{AE} = (\mathrm{id}_{R\to A}\otimes\widetilde{\mathcal{A}}_G)(\phi^{N_S}_{RA})$, where $\widetilde{\mathcal{A}}_G$ denotes the channel conjugate to the quantum-limited amplifier. The state φ_{AE} has its covariance matrix in the form (see Mathematica files included with the arXiv posting or alternatively Ref. [73], Appendix D.4)

$$\begin{bmatrix} a & 0 & c & 0 \\ 0 & a & 0 & c \\ c & 0 & b & 0 \\ 0 & c & 0 & b \end{bmatrix}, \tag{154}$$

and so the same proof approach to get (151) can be used to conclude that

$$\frac{1}{n}H(A^n|\overline{E^n})_{\varphi^{\otimes n}} = H(A|\overline{E})_\varphi. \tag{155}$$

Indeed, this additionally follows from the discussion after Eqs. (17)–(19) in [107]. As such, we conclude in the same way that

$$\frac{1}{n}E_F(A^n;B^n)_{\sigma^{\otimes n}} = E_F(A;B)_\sigma = E^g_F(A;B)_\sigma. \tag{156}$$

The final statement about entanglement cost in (143) follows from the fact that it is equal to the regularized entanglement of formation. ■

Remark 2. As can be seen from the proof above, the multimode minimum output entropy theorem recalled in (148) provides a significant strengthening of the results from [107]. Indeed, for ρ_{AE} any two-mode Gaussian state considered in [107], the following equality holds:

$$\frac{1}{n}H(A^n|\overline{E^n})_{\rho^{\otimes n}} = H(A|\overline{E})_\rho, \tag{157}$$

implying that the measurement $\{\Lambda^x_E\}_x$ optimal for the right-hand side leads to a measurement $\{\Lambda^{x_1}_{E_1}\otimes\cdots\otimes\Lambda^{x_n}_{E_n}\}_{x_1,\dots,x_n}$ that is optimal for the left-hand side. Furthermore, by the relation in (144), for any purification ψ_{ABE} of the state ρ_{AE} mentioned above, we conclude that

$$\frac{1}{n}E_F(A^n;B^n)_{\psi^{\otimes n}} = E_F(A;B)_\psi, \tag{158}$$

for all integer $n\geqslant 1$, thus giving a whole host of two-mode Gaussian states for which their entanglement cost is equal to their entanglement of formation: $E_F(A;B)_\rho = E_C(\rho_{AB}) = E^g_F(A;B)_\rho$. For these examples of two-mode Gaussian states, the additivity relation in (158) has been explicitly shown.

Remark 3. One might wonder whether the same method of proof as given in Proposition 6 could be used to establish the equalities in (141) and (142) for general thermal, amplifier, and additive-noise channels. At the moment, it is not clear how to do so. The issue is that the state $(\mathrm{id}_R\otimes\mathcal{L}_{\eta,N_B})(\phi^{N_S}_{RA})$ for $N_B>0$ is a faithful state, meaning that it is positive definite and thus has two symplectic eigenvalues >1. This means that any purification of it requires at least four modes (Ref. [109], Sec. III-D). Then tracing over the B system leaves a three-mode state, of which we should be measuring two of them, and so it is not clear how to apply the methods of [107] to such a state. The same issues apply to the states $(\mathrm{id}_R\otimes\mathcal{A}_{G,N_B})(\phi^{N_S}_{RA})$ for $N_B>0$ and $(\mathrm{id}_R\otimes\mathcal{T}_\xi)(\phi^{N_S}_{RA})$ for $\xi>0$, which are the states resulting from the amplifier and additive-noise channels, respectively.

G. Entanglement cost of pure-loss and pure-amplifier channels

Based on the results in the previous sections, we conclude the following theorem, which gives simple formulas for the entanglement cost of two fundamental bosonic Gaussian channels.

Theorem 2. For a pure-loss channel $\mathcal{L}_\eta$ with transmissivity $\eta\in(0,1)$ or a pure-amplifier channel $\mathcal{A}_G$ with gain $G>1$, the following formulas characterize the entanglement costs of these channels:

$$E_C(\mathcal{L}_\eta) = E^{(p)}_C(\mathcal{L}_\eta) = \frac{h_2(1-\eta)}{1-\eta}, \tag{159}$$

$$E_C(\mathcal{A}_G) = E^{(p)}_C(\mathcal{A}_G) = \frac{g_2(G-1)}{G-1}, \tag{160}$$

where $h_2(\cdot)$ is the binary entropy defined in (66) and $g_2(\cdot)$ is the bosonic entropy function defined in (25).

Proof. Recalling the discussion in Sec. V D 1, for a pure-loss and pure-amplifier channel, there exist respective resource states $\omega^\eta_{A'B'}$ and $\omega^G_{A'B'}$ such that

$$E_C(\mathcal{L}_\eta) \leqslant E_F(A';B')_{\omega^\eta} \tag{161}$$

$$= \lim_{N_S\to\infty} E_F(R;B)_{\sigma^\eta(N_S)}, \tag{162}$$

$$E_C(\mathcal{A}_G) \leqslant E_F(A';B')_{\omega^G} \tag{163}$$

$$= \lim_{N_S\to\infty} E_F(R;B)_{\sigma^G(N_S)}, \tag{164}$$

where

$$\sigma^\eta(N_S)_{RB} \equiv (\mathrm{id}_R\otimes\mathcal{L}_\eta)(\phi^{N_S}_{RA}), \tag{165}$$

$$\sigma^G(N_S)_{RB} \equiv (\mathrm{id}_R\otimes\mathcal{A}_G)(\phi^{N_S}_{RA}), \tag{166}$$

with the equalities in (162) and (164) being one of the main results of [99]. Furthermore, explicit formulas for $E_F(A';B')_{\omega^\eta}$ and $E_F(A';B')_{\omega^G}$ have been given in Ref. [99], Eqs. (4)–(6), and evaluating these formulas leads to the expressions in (159)–(160) (supplemental Mathematica files that automate these calculations are available with the arXiv posting of this paper).

On the other hand, Propositions 5 and 6 imply that

$$E_C(\mathcal{L}_\eta) \geqslant E^{(p)}_C(\mathcal{L}_\eta) \tag{167}$$

$$\geqslant \lim_{N_S\to\infty} E_C(\sigma^\eta(N_S)_{RB}) \tag{168}$$

$$= \lim_{N_S\to\infty} E_F(R;B)_{\sigma^\eta(N_S)}, \tag{169}$$

$$E_C(\mathcal{A}_G) \geqslant E^{(p)}_C(\mathcal{A}_G) \tag{170}$$

$$\geqslant \lim_{N_S\to\infty} E_C(\sigma^G(N_S)_{RB}) \tag{171}$$

$$= \lim_{N_S\to\infty} E_F(R;B)_{\sigma^G(N_S)}. \tag{172}$$

Combining the inequalities above, we conclude the statement of the theorem. ■

It is interesting to consider various limits of the formulas in (159) and (160):

$$\lim_{\eta\to 1}\frac{h_2(1-\eta)}{1-\eta}=\lim_{G\to 1}\frac{g_2(G-1)}{G-1}=\infty, \tag{173}$$

$$\lim_{\eta\to 0}\frac{h_2(1-\eta)}{1-\eta}=\lim_{G\to \infty}\frac{g_2(G-1)}{G-1}=0. \tag{174}$$

We expect these to hold because the channels approach the ideal channel in the limits $\eta, G \to 1$, which we previously argued has infinite entanglement cost, while they both approach the completely depolarizing (useless) channel in the no-transmission limit $\eta \to 0$ and infinite-amplification limit $G \to \infty$. Furthermore, these formulas obey the symmetry

$$\frac{h_2(1-\eta)}{1-\eta}=\frac{g_2(1/\eta-1)}{1/\eta-1}, \tag{175}$$

which is consistent with the idea that the transformation $\eta \to 1/\eta$ takes a channel of transmissivity $\eta \in [0,1]$ and produces a channel of gain $1/\eta$. Finally, we have the Taylor expansions:

$$\frac{h_2(1-\eta)}{1-\eta}=\frac{\eta}{\ln 2}[1-\ln(\eta)]+O(\eta^2), \tag{176}$$

$$\frac{g_2(G-1)}{G-1}=\frac{1+\ln(G)}{G\ln 2}+O(1/G^2), \tag{177}$$

which are relevant in the low-transmissivity and high-gain regimes.

In [110], simple formulas for the distillable entanglement of these channels were determined and given by

$$E_D(\mathcal{L}_\eta)=-\log_2(1-\eta), \tag{178}$$

$$E_D(\mathcal{A}_G)=-\log_2(1-1/G). \tag{179}$$

Thus the prior results and the formulas in Theorem 2 demonstrate that the resource theory of entanglement for these channels is irreversible. That is, if one started from a pure-loss channel of transmissivity η and distilled ebits from it at the ideal rate of $-\log_2(1-\eta)$, and then subsequently wanted to use these ebits to simulate a pure-loss channel with the same transmissivity, this is not possible, because the rate at which ebits are distilled is not sufficient to simulate the channel again. The same statement applies to the pure-amplifier channel. Figures 6 and 7 compare the formulas for entanglement cost and distillable entanglement of these channels, demonstrating that there is a noticeable gap between them. I note here that the differences are given by

$$E_C(\mathcal{L}_\eta)-E_D(\mathcal{L}_\eta)=\frac{-\eta\log_2\eta}{1-\eta}, \tag{180}$$

$$E_C(\mathcal{A}_G)-E_D(\mathcal{A}_G)=\frac{\log_2 G}{G-1}, \tag{181}$$

implying that these differences are strictly greater than zero for all the relevant channel parameter values $\eta \in (0,1)$ and $G > 1$.

VI. EXTENSION TO OTHER RESOURCE THEORIES

Let us now consider how to extend many of the concepts in this paper to other resource theories (see [6] for a review of quantum resource theories). In fact, this can be accomplished on a simple conceptual level by replacing "LOCC channel" with "free channel," "separable state" with "free state," and (roughly) "maximally entangled state" with resource state throughout the paper. To be precise, let F denote the set of free states for a given resource theory, and let $\mathcal{F}$ be a free channel, which takes a free state to a free state. In Ref. [36], Sec. 7, a general notion of distillation of a resource from n uses of a channel was given (see Fig. 4 therein). In particular, one interleaves n uses of a given channel by free channels, and the goal is to distill resource from the n channels. As a generalization of a teleportation-simulable channel with an associated resource state, the notion of a ν-freely simulable channel was introduced as a channel $\mathcal{N}$ that can be simulated as

$$\mathcal{N}_{A\to B}(\rho_A)=\mathcal{F}^{\text{sim}}_{AE\to B}(\rho_A\otimes\nu_E), \tag{182}$$

where $\mathcal{F}^{\text{sim}}$ is a free channel and ν is some resource state. The implications of this for distillation protocols was discussed in Ref. [36], Sec. 7, which is merely that the rate at which

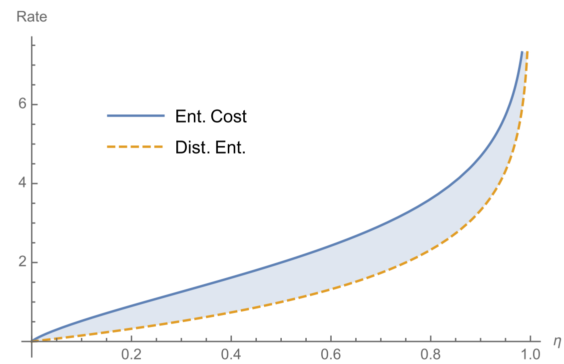

FIG. 6. Plot of the entanglement cost $E_C(\mathcal{L}_\eta)=\frac{h_2(1-\eta)}{1-\eta}$ and the distillable entanglement $E_D(\mathcal{L}_\eta)=-\log_2(1-\eta)$ of the pure-loss channel $\mathcal{L}_\eta$ as a function of the transmissivity $\eta \in [0,1]$, with the shaded area demonstrating the gap between them. The units for rate on the vertical axis are ebits per channel use and η on the horizontal axis is dimensionless.

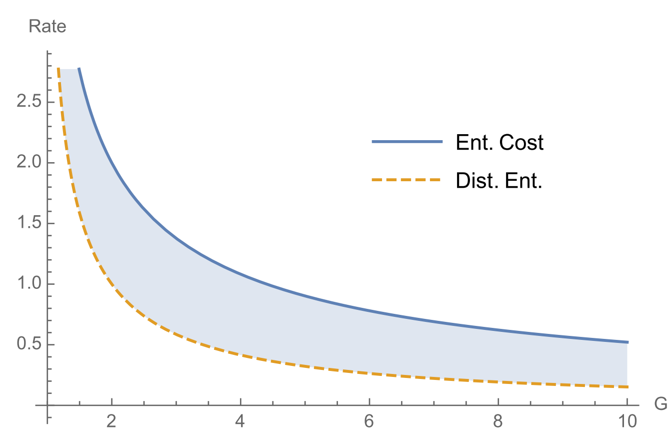

FIG. 7. Plot of the entanglement cost $E_C(\mathcal{A}_G)=\frac{g_2(G-1)}{G-1}$ and the distillable entanglement $E_D(\mathcal{A}_G)=-\log_2(1-1/G)$ of the pure-amplifier channel $\mathcal{A}_G$ as a function of the gain $G \geqslant 1$, with the shaded area demonstrating the gap between them. The units for rate on the vertical axis are ebits per channel use and G on the horizontal axis is dimensionless.

resource can be distilled is limited by the resourcefulness of the underlying resource state ν.

Going forward, we can also consider a resource-seizable channel in a general resource theory to be a ν-freely simulable channel for which, by pre- and postprocessing, one can seize the underlying resource state ν as

$$\mathcal{F}^{\text{post}}_{RB\to E}\big(\mathcal{N}_{A\to B}\big(\kappa^{\text{pre}}_{RA}\big)\big) = \nu_E, \tag{183}$$

where κ^{pre}_{RA} is a free state and $\mathcal{F}^{\text{post}}_{RB\to E}$ is a free channel, extending Definition 2.

The general notion of channel simulation, as presented in Sec. II B, can be considered in any resource theory also. Again, the main idea is really to replace "LOCC channel" with "free channel" and "maximally entangled state" with "resourceful state" in the protocol depicted in Fig. 1, and the goal is to determine the minimum rate at which resourcefulness is needed in order to simulate n uses of a given channel. If the channels are resource seizable as discussed above, then the theory should significantly simplify, as has occurred in this paper for the entanglement theory of channels (see Theorem 1). Furthermore, along the lines of the discussion in Sec. II C (and related to Ref. [6], Sec. III-D-5), suppose that a channel $\mathcal{N}_{A\to B}$ can be realized from another channel $\mathcal{M}_{A'\to B'}$ via a preprocessing free channel $\mathcal{F}^{\text{pre}}_{A\to A'M}$ and a postprocessing free channel $\mathcal{F}^{\text{post}}_{B'M\to B}$ as

$$\mathcal{N}_{A\to B} = \mathcal{F}^{\text{post}}_{B'M\to B}\circ\mathcal{M}_{A'\to B'}\circ\mathcal{F}^{\text{pre}}_{A\to A'M}. \tag{184}$$

Then for the same reasons given there, the simulation cost of $\mathcal{N}$ should never exceed the simulation cost of $\mathcal{M}$.

Finally, let us note that some discussions about channel simulation for the resource theory of coherence have appeared in the last paragraph of [111], as well as the last paragraphs of [112]. It is clear from the findings of the present paper that identifying interesting resource-seizable channels could be a useful first step for understanding interconversion costs of simulating one channel from another in the resource theory of coherence. It could also be helpful in further understanding channel simulation in the resource theory of thermodynamics [113].

VII. CONCLUSION

In summary, this paper has provided a definition for the entanglement cost of a channel, in terms of the most general strategy that a discriminator could use to distinguish n uses of the channel from its simulation. I established an upper bound on the entanglement cost of a teleportation-simulable channel in terms of the entanglement cost of the underlying resource state, and I proved that the bound is saturated in the case that the channel is resource seizable (Definition 2). I then established single-letter formulas for the entanglement cost of erasure, dephasing, three-dimensional Werner-Holevo channels, and epolarizing channels (complements of depolarizing channels), by leveraging existing results about the entanglement cost of their Choi states. I finally considered single-mode bosonic Gaussian channels, establishing bounds on the entanglement cost of the thermal, amplifier, and additive-noise channels, while giving simple formulas for the entanglement cost of pure-loss and pure-amplifier channels. By relating to prior work on the distillable entanglement of these channels, it became clear that the resource theory of entanglement for quantum channels is irreversible.

Going forward from here, there are many directions to pursue. The discrimination protocols considered in Sec. II B do not impose any realistic energy constraint on the states that can be used in discriminating the actual n uses of the channel from the simulation. We could certainly do so by imposing that the average energy of all the states input to the actual channel or its simulation should be less than a threshold, and the result is to demand only that the *energy-constrained strategy norm* (defined naturally as an extension of both the strategy norm [26,27,33] and the energy-constrained diamond norm [76,77]) is less than $\varepsilon \in (0,1)$. To be specific, let H_A be a (positive semidefinite) Hamiltonian acting on the input of the channel $\mathcal{N}_{A\to B}$ and let $E \in [0,\infty)$ be an energy constraint. Then, demanding that the supremum in (18) is taken over all strategies such that

$$\frac{1}{n}\sum_{i=1}^{n}\text{Tr}\left\{H_A\rho_{A_i}\right\},\quad \frac{1}{n}\sum_{i=1}^{n}\text{Tr}\left\{H_A\tau_{A_i}\right\} \leqslant E, \tag{185}$$

the resulting quantity is an energy-constrained strategy norm. With an energy constraint in place, one would expect that less entanglement is required to simulate the channel than if there is no constraint at all, and the resulting entanglement cost would depend on the given energy constraint. For example, Proposition 4 leads to a lower bound on entanglement cost for an energy-constrained sequential simulation, but it remains open to determine if there is a matching upper bound.

Similar to how measures like squashed entanglement [12] and relative entropy of entanglement [114] allow for obtaining converse bounds or fundamental limitations on the distillation rates of quantum states or channels, simply by making a clever choice of a squashing channel or separable state, it would be useful to have a measure like this for bounding entanglement cost from below. That is, it would be desirable for the measure to involve a supremum over a given set of test states or channels rather than an infimum as is the case for squashed entanglement and relative entropy of entanglement. For example, it would be useful to be able to bound the entanglement cost of thermal, amplifier, and additive-noise channels from below, in order to determine how tight are the upper bounds in (127)–(129). Progress on this front is available in [8], but more results in this area would be beneficial.

One of the main tools used in the analysis of the (parallel) entanglement cost of channels from [19] is a de Finetti–style approach, consisting of the postselection technique [115]. In particular, the problem of asymptotic (parallel) channel simulation was reduced to simulating the channel on a single state, called the universal de Finetti state. For the asymptotic theory of (sequential) entanglement cost of channels, could there be a single universal adaptive channel discrimination protocol to consider, such that simulating the channel well for such a protocol would imply that it has been simulated well for all protocols?

For the task of entanglement cost, one could modify the set of free channels to be either those that completely preserve the positivity of partial transpose [10,11] or are k extendible in the sense of [116]. Could we find simpler lower bounds on entanglement cost of channels in this way? The semidefinite programming quantity from [8] could be helpful here also. Most recently, the exact entanglement cost has been solved in [117] for the case of *exact* channel simulation, with the set of free channels taken to be those that completely preserve the positivity of partial transpose.

Another way to think about quantum channel simulation is to allow the entanglement to be free but count the cost of classical communication. This was the approach taken for the reverse Shannon theorem [118,119], and these works also considered only parallel channel simulation. How are the results there affected if the goal is sequential channel simulation instead? Is the previous answer from [118,119], the mutual information of the channel, robust under this change? How do prior results on simulation of quantum measurements [120–122] hold up under this change? A comprehensive summary of results on parallel simulation of quantum channels, including the quantum reverse Shannon theorem, measurement simulation, and entanglement cost, is available in [123].

Finally, is there an example of a channel for which its sequential entanglement cost is strictly greater than its parallel entanglement cost? The examples discussed here are those for which either there are equalities or no conclusion could be drawn. Evidence from quantum channel discrimination [29] and related evidence from [124] suggests the possibility. One concrete example to examine in this context is the channel presented in Ref. [43], Appendix A, given that it is not implementable from its image, as discussed there.

ACKNOWLEDGMENTS

I thank Gerardo Adesso, Eneet Kaur, Debbie Leung, Alexander Müller-Hermes, Kaushik Seshadreesan, Xin Wang, John Watrous, and Andreas Winter for helpful discussions, especially Andreas Winter for inspiring discussions regarding a general notion of channel simulation. I acknowledge support from the National Science Foundation under Grant No. 1350397.

APPENDIX A: RELATION BETWEEN RESOURCE-SEIZABLE CHANNELS AND THOSE THAT ARE IMPLEMENTABLE FROM THEIR IMAGE

Definition 2 introduced the notion of a resource-seizable channel and Sec. VI discussed how this notion can play a role in an arbitrary resource theory. In Ref. [43], Appendix A, a channel $\mathcal{N}_{A\to B}$ was defined to be implementable from its image if there exists a state $\sigma_{A'A}$ and an LOCC channel $\mathcal{L}_{AA'B'\to B}$ such that the following equality holds for all input states ρ_A:

$$\mathcal{N}_{A\to B}(\rho_A) = \mathcal{L}_{AA'B'\to B}(\rho_A \otimes \mathcal{N}_{A''\to B'}(\sigma_{A'A''})), \tag{A1}$$

where system A'' is isomorphic to system A and system B' is isomorphic to system B. An example of a channel that is not implementable from its image was discussed at length in Ref. [43], Appendix A.

Here, I prove that a channel is resource seizable in the resource theory of entanglement if and only if it is implementable from its image. To see this, suppose that a channel is implementable from its image. Then, given the above structure in (A1), it is clear that $\mathcal{N}_{A\to B}$ is teleportation simulable with associated resource state given by $\omega_{A'B'} = \mathcal{N}_{A''\to B'}(\sigma_{A'A''})$. Thus one can trivially seize the resource state $\omega_{A'B'}$ by sending in the input state $\sigma_{A'A''}$, which is clearly separable between Alice and Bob, given that Bob's "system" here is trivial.

Now suppose that a teleportation-simulable channel is resource seizable, as in Definition 2. This means that

$$\mathcal{N}_{A\to B}(\rho_A) = \mathcal{M}_{AA'B'\to B}(\rho_A \otimes \omega_{A'B'}), \tag{A2}$$

where $\omega_{A'B'}$ is the resource state and $\mathcal{M}_{AA'B'\to B}$ is an LOCC channel. Furthermore, since it is resource seizable, this means that there exists a separable state $\rho_{A_M A B_M}$ and a postprocessing LOCC channel $\mathcal{D}_{A_M B B_M\to A'B'}$ such that

$$\mathcal{D}_{A_M B B_M\to A'B'}(\mathcal{N}_{A\to B}(\rho_{A_M A B_M})) = \omega_{A'B'}. \tag{A3}$$

To see that the channel is implementable from its image, consider that $\rho_{A_M A B_M}$ has a decomposition as follows, given that it is separable:

$$\sum_x p_X(x)\psi^x_{A_M A} \otimes \phi^x_{B_M}, \tag{A4}$$

for p_X a probability distribution and $\{\psi^x_{A_M A}\}_x$ and $\{\phi^x_{B_M}\}_x$ sets of pure states. Now define the input state $\sigma_{A_M A X_A}$ as

$$\sigma_{A_M A X_A} \equiv \sum_x p_X(x)\psi^x_{A_M A} \otimes |x\rangle\langle x|_{X_A}, \tag{A5}$$

and note that this is the state we can use for implementing the channel's image. Define the LOCC measure-prepare channel $\mathcal{P}_{X_A\to B_M}$ as

$$\mathcal{P}_{X_A\to B_M}(\cdot) \equiv \sum_x \langle x|_{X_A}(\cdot)|x\rangle_{X_A}\, \phi^x_{B_M}, \tag{A6}$$

which is understood to be implemented via LOCC by measuring Alice's system X_A and communicating the outcome x to Bob, who then prepares the state $\phi^x_{B_M}$ based on the outcome. We find that

$$\big(\mathcal{D}_{A_M B B_M\to A'B'} \circ \mathcal{P}_{X_A\to B_M} \circ \mathcal{N}_{A\to B}\big)\big(\sigma_{A_M A X_A}\big) = \omega_{A'B'}. \tag{A7}$$

We finally conclude that

$$\mathcal{N}_{A\to B}(\rho_A) = \mathcal{M}_{AA'B'\to B}(\rho_A \otimes \omega_{A'B'}) \tag{A8}$$

$$= \mathcal{L}_{AA_M X_A\bar{B}\to B}\big(\rho_A \otimes \mathcal{N}_{\bar{A}\to\bar{B}}\big(\sigma_{A_M\bar{A}X_A}\big)\big), \tag{A9}$$

where

$$\mathcal{L}_{AA_M X_A\bar{B}\to B} \equiv \mathcal{M}_{AA'B'\to B} \circ \mathcal{D}_{A_M\bar{B}B_M\to A'B'} \circ \mathcal{P}_{X_A\to B_M}, \tag{A10}$$

so that the channel is implementable from its image by inputting the state $\sigma_{A_M A X_A}$ and postprocessing with the LOCC channel $\mathcal{M}_{AA'B'\to B} \circ \mathcal{D}_{A_M B B_M\to A'B'} \circ \mathcal{P}_{X_A\to B_M}$.

APPENDIX B: RELATION BETWEEN CHOI STATE OF A COMPLEMENTARY CHANNEL AND MAXIMALLY MIXED STATE SENT THROUGH ISOMETRIC EXTENSION

The purpose of this appendix is to prove the equality in (98). Consider a d-dimensional depolarizing channel

$$\rho \to (1-p)\rho + p\frac{I}{d}. \tag{B1}$$

As noted in Ref. [65], Eq. (3.2), a Kraus representation for this channel is as follows:

$$\{\sqrt{1-p}I, \{\sqrt{p/d}|i\rangle\langle j|\}_{i,j}\}. \tag{B2}$$

This is because

$$[\sqrt{1-p}I]\rho[\sqrt{1-p}I] + \sum_{i,j}[\sqrt{p/d}|i\rangle\langle j|]\rho[\sqrt{p/d}|j\rangle\langle i|]$$

$$= (1-p)\rho + \frac{p}{d}\sum_i |i\rangle\langle i| \sum_j \langle j|\rho|j\rangle \tag{B3}$$

$$= (1-p)\rho + p\,\mathrm{Tr}\{\rho\}\frac{I}{d}. \tag{B4}$$

Now consider a generic channel $\mathcal{N}_{A\to B}$ with Kraus operators $\{N^i\}_i$ so that an isometric extension is given by $\sum_i N^i \otimes |i\rangle_E$. Send the maximally mixed state $\pi = I/d$ through the isometric extension $\sum_i N^i \otimes |i\rangle_E$. This leads to the state

$$\frac{1}{d}\sum_{i,j} N^i N^{j\dagger} \otimes |i\rangle\langle j|_E. \tag{B5}$$

Furthermore, a complementary channel of the original channel, resulting from the isometric extension $\sum_i N^i \otimes |i\rangle_E$, is then

$$\rho \to \mathcal{N}^c_{A\to E}(\rho) = \sum_{i,j} \mathrm{Tr}\{N^i \rho N^{j\dagger}\}|i\rangle\langle j|_E. \tag{B6}$$

The Choi state for this complementary channel is given by

$$\mathcal{N}^c_{A\to E}(\Phi_{RA}) = \frac{1}{d}\sum_{k,l,i,j} |k\rangle\langle l|_R \otimes \mathrm{Tr}\{N^i|k\rangle\langle l|_A N^{j\dagger}\}|i\rangle\langle j|_E$$

$$= \frac{1}{d}\sum_{k,l,i,j} |k\rangle\langle l|_R \otimes \langle l|_A N^{j\dagger}N^i|k\rangle|i\rangle\langle j|_E$$

$$= \frac{1}{d}\sum_{k,l,i,j} |k\rangle\langle l|_A N^{j\dagger}N^i|k\rangle\langle l|_R \otimes |i\rangle\langle j|_E$$

$$= \frac{1}{d}\sum_{i,j} T(N^{j\dagger}N^i) \otimes |i\rangle\langle j|_E, \tag{B7}$$

where $T(N^{j\dagger}N^i)$ denotes the transpose of $N^{j\dagger}N^i$. If it holds that $N^iN^{j\dagger} = T(N^{j\dagger}N^i)$, then we conclude that the state resulting from sending in the maximally mixed state to the isometric extension of the channel is the same as the Choi state of the complementary channel. This is the case for the depolarizing channel with the Kraus operators in (B2). Since all complementary channels and isometric extensions of a channel are related by an isometry acting on the environment system, we arrive at the same conclusion for any isometric extension and the corresponding complementary channel to which it leads.

APPENDIX C: MATLAB CODE FOR COMPUTING RAINS RELATIVE ENTROPY

This Appendix provides a brief listing of Matlab code that can be used to compute the Rains relative entropy of a bipartite state ρ_{AB} [11,70]. The code requires the QuantInf package in order to generate a random state [125], the CVX package for semidefinite programming optimization [126], and the CVXQuad package [127] for relative entropy optimization [71,72].

Listing 1. Matlab code for calculating the Rains relative entropy of a random bipartite state ρ_{AB}.

```
na = 2;  nb = 2;
rho = randRho (na*nb); % Generate a random bipartite state rho

cvx_begin  sdp
    variable  tau (na*nb,na*nb)  hermitian;
    minimize (quantum_rel_entr (rho, tau)/ log(2));
    tau >= 0;
    norm_nuc (Tx(tau, 2, [ na nb ])) <= 1;
cvx_end

rains_rel_ent = cvx_optval;
```


[1] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[2] A. S. Holevo, *Quantum Systems, Channels, Information*, de Gruyter Studies in Mathematical Physics (Book 16) (de Gruyter, Berlin, 2012).

[3] M. Hayashi, *Quantum Information: An Introduction* (Springer, Berlin, 2006).

[4] M. M. Wilde, *Quantum Information Theory*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2017).

[5] J. Watrous, *Theory of Quantum Information* (Cambridge University Press, Cambridge, UK, 2018).

[6] E. Chitambar and G. Gour, Quantum resource theories, arXiv:1806.06107.

[7] P. M. Hayden, M. Horodecki, and B. M. Terhal, The asymptotic entanglement cost of preparing a quantum state, J. Phys. A: Math. Gen. **34**, 6891 (2001).

[8] X. Wang and R. Duan, Irreversibility of Asymptotic Entanglement Manipulation Under Quantum Operations Completely Preserving Positivity of Partial Transpose, Phys. Rev. Lett. **119**, 180506 (2017).

[9] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, Proc. R. Soc. A **461**, 207 (2005).

[10] E. M. Rains, Bound on distillable entanglement, Phys. Rev. A **60**, 179 (1999).

[11] E. M. Rains, A semidefinite program for distillable entanglement, IEEE Trans. Inf. Theory **47**, 2921 (2001).

[12] M. Christandl and A. Winter, "Squashed entanglement": An additive entanglement measure, J. Math. Phys. **45**, 829 (2004).

[13] X. Wang and R. Duan, An improved semidefinite programming upper bound on distillable entanglement, Phys. Rev. A **94**, 050301 (2016).

[14] 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. **70**, 1895 (1993).

[15] M. Takeoka, S. Guha, and M. M. Wilde, Fundamental rate-loss tradeoff for optical quantum key distribution, Nat. Commun. **5**, 5235 (2014).

[16] M. Takeoka, S. Guha, and M. M. Wilde, The squashed entanglement of a quantum channel, IEEE Trans. Inf. Theory **60**, 4987 (2014).

[17] M. M. Wilde, Squashed entanglement and approximate private states, Quantum Inf. Process. **15**, 4563 (2016).

[18] M. Berta and M. M. Wilde, Amortization does not enhance the max-Rains information of a quantum channel, New J. Phys. **20**, 053044 (2018).

[19] M. Berta, F. G. S. L. Brandão, M. Christandl, and S. Wehner, Entanglement cost of quantum channels, IEEE Trans. Inf. Theory **59**, 6779 (2013).

[20] 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. **328**, 303 (2014).

[21] A. Kitaev, Quantum computations: Algorithms and error correction, Russian Math. Surveys **52**, 1191 (1997).

[22] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[23] A. S. Holevo, Statistical decision theory for quantum systems, J. Multivar. Anal. **3**, 337 (1973).

[24] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[25] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, Phys. Rev. Lett. **101**, 060401 (2008).

[26] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory Effects in Quantum Channel Discrimination, Phys. Rev. Lett. **101**, 180501 (2008).

[27] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[28] R. Duan, Y. Feng, and M. Ying, Perfect Distinguishability of Quantum Operations, Phys. Rev. Lett. **103**, 210501 (2009).

[29] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Adaptive versus nonadaptive strategies for quantum channel discrimination, Phys. Rev. A **81**, 032339 (2010).

[30] 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. **344**, 797 (2016).

[31] G. Gutoski and J. Watrous, Toward a general theory of quantum games, Proc. STOC **2007**, 565 (2007).

[32] G. Gutoski, Ph.D. thesis, University of Waterloo, 2009, arXiv:1003.0038.

[33] G. Gutoski, On a measure of distance for quantum strategies, J. Math. Phys. **53**, 032202 (2012).

[34] G. Gutoski, A. Rosmanis, and J. Sikora, Fidelity of quantum strategies with applications to cryptography, in *12th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2017)*, Leibniz International Proceedings in Informatics (LIPIcs) Vol. 73 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2018), pp. 8:1–8:13.

[35] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Quantum nonlocality without entanglement, Phys. Rev. A **59**, 1070 (1999).

[36] E. Kaur and M. M. Wilde, Amortized entanglement of a quantum channel and approximately teleportation-simulable channels, J. Phys. A **51**, 035303 (2018).

[37] A. Winter, Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints, Commun. Math. Phys. **347**, 291 (2016).

[38] M. Horodecki, P. W. Shor, and M. B. Ruskai, Entanglement breaking channels, Rev. Math. Phys. **15**, 629 (2003).

[39] A. S. Holevo, Entanglement-breaking channels in infinite dimensions, Probl. Inf. Transm. (Engl. Transl.) **44**, 171 (2008).

[40] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke, Irreversibility for All Bound Entangled States, Phys. Rev. Lett. **95**, 190501 (2005).

[41] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Phys. Rev. A **60**, 1888 (1999).

[42] F. Buscemi and N. Datta, Entanglement Cost in Practical Scenarios, Phys. Rev. Lett. **106**, 130503 (2011).

[43] M. Christandl and A. Müller-Hermes, Relative entropy bounds on quantum, private and repeater capacities, Commun. Math. Phys. **353**, 821 (2017).

[44] M. B. Hastings, Superadditivity of communication capacity using entangled inputs, Nat. Phys. **5**, 255 (2009).

[45] A. S. Holevo, Remarks on the classical capacity of quantum channel, arXiv:quant-ph/0212025.

[46] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Realization schemes for quantum instruments in finite dimensions, J. Math. Phys. **50**, 042101 (2009).

[47] M. M. Wilde, M. Tomamichel, and M. Berta, Converse bounds for private communication over quantum channels, IEEE Trans. Inf. Theory **63**, 1792 (2017).

[48] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[49] M. Horodecki and P. Horodecki, Reduction criterion of separability and limits for a class of distillation protocols, Phys. Rev. A **59**, 4206 (1999).

[50] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, Phys. Rev. A **56**, 33 (1997).

[51] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Capacities of Quantum Erasure Channels, Phys. Rev. Lett. **78**, 3217 (1997).

[52] M. Tomamichel, M. M. Wilde, and A. Winter, Strong converse rates for quantum communication, IEEE Trans. Inf. Theory **63**, 715 (2017).

[53] G. Vidal, W. Dür, and J. I. Cirac, Entanglement Cost of Bipartite Mixed States, Phys. Rev. Lett. **89**, 027901 (2002).

[54] M. Horodecki, A. Sen(De), and U. Sen, Rates of asymptotic entanglement transformations for bipartite mixed states:

Maximally entangled states are not special, Phys. Rev. A **67**, 062314 (2003).
[55] W. K. Wootters, Entanglement of Formation of an Arbitrary State of Two Qubits, Phys. Rev. Lett. **80**, 2245 (1998).
[56] R. F. Werner and A. S. Holevo, Counterexample to an additivity conjecture for output purity of quantum channels, J. Math. Phys. **43**, 4353 (2002).
[57] D. Leung and W. Matthews, On the power of PPT-preserving and non-signalling codes, IEEE Trans. Inf. Theory **61**, 4486 (2015).
[58] M. Christandl, N. Schuch, and A. Winter, Entanglement of the antisymmetric state, Commun. Math. Phys. **311**, 397 (2012).
[59] K. G. H. Vollbrecht and R. F. Werner, Entanglement measures under symmetry, Phys. Rev. A **64**, 062307 (2001).
[60] F. Yura, Entanglement cost of three-level antisymmetric states, J. Phys. A: Math. Gen. **36**, L237 (2003).
[61] A. Winter (private communication).
[62] W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).
[63] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, Properties of conjugate channels with applications to additivity and multiplicativity, Markov Process. Relat. Fields **13**, 391 (2007).
[64] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Trans. Inf. Theory **51**, 44 (2005).
[65] N. Datta, M. Fukuda, and A. S. Holevo, Complementarity and additivity for covariant channels, Quantum Inf. Process. **5**, 179 (2006); N. Datta and A. S. Holevo, Complementarity and additivity for depolarizing channels, arXiv:quant-ph/0510145; M. Fukuda and A. S. Holevo, On Weyl-covariant channels, arXiv:quant-ph/0510148.
[66] D. Leung and J. Watrous, On the complementary quantum capacity of the depolarizing channel, Quantum **1**, 28 (2017).
[67] K. Matsumoto, T. Shimono, and A. Winter, Remarks on additivity of the Holevo channel capacity and of the entanglement of formation, Commun. Math. Phys. **246**, 427 (2004).
[68] C. King, The capacity of the quantum depolarizing channel, IEEE Trans. Inf. Theory **49**, 221 (2003).
[69] H. Umegaki, Conditional expectations in an operator algebra IV (entropy and information), Kodai Math. Seminar Rep. **14**, 59 (1962).
[70] 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 **66**, 032310 (2002).
[71] H. Fawzi, J. Saunderson, and P. A. Parrilo, Semidefinite approximations of the matrix logarithm, Found. Comput. Math. (2018).
[72] H. Fawzi and O. Fawzi, Efficient optimization of the quantum relative entropy, J. Phys. A: Math. Theor. **51**, 154003 (2018).
[73] A. Serafini, *Quantum Continuous Variables* (CRC Press, Boca Raton, FL, 2017).
[74] A. S. Holevo and V. Giovannetti, Quantum channels and their entropic characteristics, Rep. Prog. Phys. **75**, 046001 (2012).
[75] M. E. Shirokov and A. S. Holevo, On approximation of infinite-dimensional quantum channels, Probl. Inf. Transm. (Engl. Transl.) **44**, 73 (2008).
[76] M. E. Shirokov, Energy-constrained diamond norms and their use in quantum information theory, Probl. Inf. Transm. (Engl. Transl.) **54**, 20 (2018).
[77] A. Winter, Energy-constrained diamond norm with applications to the uniform continuity of continuous variable channel capacities, arXiv:1712.10267.
[78] M. M. Wilde, Strong and uniform convergence in the teleportation simulation of bosonic Gaussian channels, Phys. Rev. A **97**, 062305 (2018).
[79] M. E. Shirokov, On properties of the space of quantum states and their application to the construction of entanglement monotones, Izv.: Math. **74**, 849 (2010).
[80] A. S. Holevo, One-mode quantum Gaussian channels: Structure and quantum capacity, Probl. Inf. Transm. (Engl. Transl.) **43**, 1 (2007).
[81] J. S. Ivan, K. K. Sabapathy, and R. Simon, Operator-sum representation for bosonic Gaussian channels, Phys. Rev. A **84**, 042311 (2011).
[82] V. Giovannetti, R. Garcia-Patron, N. J. Cerf, and A. S. Holevo, Ultimate classical communication rates of quantum optical channels, Nat. Photon. **8**, 796 (2014).
[83] S. L. Braunstein and H. J. Kimble, Teleportation of Continuous Quantum Variables, Phys. Rev. Lett. **80**, 869 (1998).
[84] M. Ban, M. Sasaki, and M. Takeoka, Continuous variable teleportation as a generalized thermalizing quantum channel, J. Phys. A: Math. Gen. **35**, L401 (2002).
[85] M. Takeoka, M. Ban, and M. Sasaki, Quantum channel of continuous variable teleportation and nonclassicality of quantum states, J. Opt. B: Quantum Semiclass. Opt. **4**, 114 (2002).
[86] G. Giedke and J. I. Cirac, Characterization of Gaussian operations and distillation of Gaussian states, Phys. Rev. A **66**, 032316 (2002).
[87] M. M. Wolf, D. Pérez-García, and G. Giedke, Quantum Capacities of Bosonic Channels, Phys. Rev. Lett. **98**, 130501 (2007).
[88] J. Niset, J. Fiurasek, and N. J. Cerf, No-go Theorem for Gaussian Quantum Error Correction, Phys. Rev. Lett. **102**, 120501 (2009).
[89] M. M. Wolf, G. Giedke, O. Krüger, R. F. Werner, and J. I. Cirac, Gaussian entanglement of formation, Phys. Rev. A **69**, 052320 (2004).
[90] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).
[91] J. Williamson, On the algebraic problem concerning the normal forms of linear dynamical systems, Am. J. Math. **58**, 141 (1936).
[92] M. M. Wilde, J. M. Renes, and S. Guha, Second-order coding rates for pure-loss bosonic channels, Quantum Inf. Process. **15**, 1289 (2016).
[93] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer, Berlin, 1993).
[94] B. Schumacher, Quantum coding, Phys. Rev. A **51**, 2738 (1995).
[95] A. Winter, Coding theorem and strong converse for quantum channels, IEEE Trans. Inf. Theory **45**, 2481 (1999).
[96] T. Ogawa and H. Nagaoka, Making good codes for classical-quantum channel coding via quantum hypothesis testing, IEEE Trans. Inf. Theory **53**, 2261 (2007).
[97] P. Liuzzo-Scorpo, A. Mari, V. Giovannetti, and G. Adesso, Optimal Continuous Variable Quantum Teleportation with Limited Resources, Phys. Rev. Lett. **119**, 120503 (2017).

[98] E. Kaur and M. M. Wilde, Upper bounds on secret key agreement over lossy thermal bosonic channels, Phys. Rev. A **96**, 062318 (2017).

[99] S. Tserkis, J. Dias, and T. C. Ralph, Simulation of Gaussian channels via teleportation and error correction of Gaussian states, arXiv:1803.03516.

[100] G. Giedke, J. Eisert, J. I. Cirac, and M. B. Plenio, Entanglement transformations of pure Gaussian states, Quantum Inf. Comput. **3**, 211 (2003).

[101] M. E. Shirokov, Squashed entanglement in infinite dimensions, J. Math. Phys. **57**, 032203 (2016).

[102] M. E. Shirokov, Entropy characteristics of subsets of states. I, Izv.: Math. **70**, 1265 (2006).

[103] V. Giovannetti, A. S. Holevo, and R. Garcia-Patron, A solution of Gaussian optimizer conjecture for quantum channels, Commun. Math. Phys. **334**, 1553 (2015).

[104] V. Giovannetti, A. S. Holevo, and A. Mari, Majorization and additivity for multimode bosonic Gaussian channels, Theor. Math. Phys. **182**, 284 (2015).

[105] M. Koashi and A. Winter, Monogamy of quantum entanglement and other correlations, Phys. Rev. A **69**, 022309 (2004).

[106] E. Schrödinger, Discussion of probability relations between separated systems, Math. Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[107] S. Pirandola, G. Spedalieri, S. L. Braunstein, N. J. Cerf, and S. Lloyd, Optimality of Gaussian Discord, Phys. Rev. Lett. **113**, 140405 (2014).

[108] A. Mari, V. Giovannetti, and A. S. Holevo, Quantum state majorization at the output of bosonic Gaussian channels, Nat. Commun. **5**, 3826 (2014).

[109] A. S. Holevo and R. F. Werner, Evaluating capacities of bosonic Gaussian channels, Phys. Rev. A **63**, 032312 (2001).

[110] S. Pirandola *et al.*, Nat. Commun. **8**, 15043 (2017).

[111] K. B. Dana, M. García Díaz, M. Mejatty, and A. Winter, Resource theory of coherence: Beyond states, Phys. Rev. A **95**, 062327 (2017).

[112] M. García Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, Using and reusing coherence to realize quantum processes, arXiv:1805.04045.

[113] P. Faist, M. Berta, and F. G. S. L. Brandao, Thermodynamic capacity of quantum processes, arXiv:1807.05610.

[114] V. Vedral and M. B. Plenio, Entanglement measures and purification procedures, Phys. Rev. A **57**, 1619 (1998).

[115] M. Christandl, R. König, and R. Renner, Post-Selection Technique for Quantum Channels with Applications to Quantum Cryptography, Phys. Rev. Lett. **102**, 020504 (2009).

[116] E. Kaur, S. Das, M. M. Wilde, and A. Winter, Extendibility limits the performance of quantum processors, arXiv:1803.10710.

[117] X. Wang and M. M. Wilde, Exact entanglement cost of quantum states and channels under PPT-preserving operations, arXiv:1809.09592.

[118] 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 **60**, 2926 (2014).

[119] M. Berta, M. Christandl, and R. Renner, The quantum reverse Shannon theorem based on one-shot information theory, Commun. Math. Phys. **306**, 579 (2011).

[120] A. Winter, "Extrinsic" and "intrinsic" data in quantum measurements: asymptotic convex decomposition of positive operator valued measures, Commun. Math. Phys. **244**, 157 (2004).

[121] M. M. Wilde, P. Hayden, F. Buscemi, and M.-H. Hsieh, The information-theoretic costs of simulating quantum measurements, J. Phys. A: Math. Theor. **45**, 453001 (2012).

[122] M. Berta, J. M. Renes, and M. M. Wilde, Identifying the information gain of a quantum measurement, IEEE Trans. Inf. Theory **60**, 7987 (2014).

[123] M. Berta, Ph.D. thesis, ETH Zurich, 2013, arXiv:1310.4581.

[124] S. Das and M. M. Wilde, Quantum reading capacity: General definition and bounds, arXiv:1703.03706.

[125] T. Cubitt, Quantinf matlab package (v0.5.1), http://www.dr-qubit.org/matlab.html.

[126] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014 http://cvxr.com/cvx.

[127] H. Fawzi, CVXQuad, https://github.com/hfawzi/cvxquad.