The quantum communication power of indefinite causal order

Xuanqiang Zhao,[1] Benchi Zhao,[1] and Giulio Chiribella[1,2,3]

[1]*QICI Quantum Information and Computation Initiative, School of Computing and Data Science, The University of Hong Kong, Pokfulam Road, Hong Kong*

[2]*Quantum Group, Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom*

[3]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada*

(Dated: October 10, 2025)

Quantum theory is in principle compatible with scenarios where physical processes take place in an indefinite causal order, a possibility that was shown to yield advantages in several information processing tasks. However, advantages in communication, the most basic form of information processing, have so far remained controversial and hard to prove. Here we develop a framework that can be used to rigorously assess the role of causal order in a scenario where communication links are built by assembling multiple quantum devices. In this setting, we establish a clear-cut advantage of indefinite order in the one-shot transmission of classical messages. On the other hand, we also show that the advantage is not generic to all communication tasks. Notably, we find that indefinite order does not offer any advantage over shared entanglement in the asymptotic scenario where a large number of uses of the same communication device is employed. Overall, our results unveil non-trivial relations between communication, causal order, entanglement, and no-signaling resources in quantum mechanics.

I. INTRODUCTION

Most physical theories, with the notable exception of general relativity and its quantum extensions, describe the evolution of physical systems through time. In this picture, the evolution of a physical system typically consists of a sequence of physical processes taking place in a well-defined order. It has been observed, however, that the structure of quantum theory is in principle compatible with scenarios where the order becomes indefinite [1–3]. Part of the motivation to study these scenarios comes from the long-standing quest to quantize gravity, which naturally leads to the hypothesis that spacetime and causal structure could be subject to quantum indeterminacy [4, 5]. Further motivation comes from the attempt to characterize quantum theory in terms of information processing [6–8]. In this context, the in-principle possibility of indefinite order is a striking consequence of the quantum framework, and the study of its information-theoretic implications offers new insights into the nature of the framework itself.

Over the past two decades, advantages of indefinite order have been discovered in a variety of information processing tasks, including quantum processes discrimination [9–11], query complexity [12, 13], communication complexity [14, 15], quantum metrology [16–18], and inversion of unknown dynamics [19–21]. In all these tasks, quantum operations with indefinite order were proven to outperform all possible quantum operations compatible with the existence of a definite order. Quite strikingly, however, no such separation was found so far in the most basic information-processing task, upon which the whole of information theory was built, namely the communication of messages from a sender to a receiver [22, 23]. In the context of communication, the role of indefinite order has remained controversial so far. A body of papers [24–28] showed that combining noisy channels in an indefinite order leads to enhancements with respect to the standard quantum communication scenario, in which channels are combined in parallel or in sequence, without introducing additional systems that control their configuration. However, other works observed that the above enhancements only represent an advantage of indefinite causal order over a restricted set of scenarios with definite order, and therefore do not establish a separation between the performance of definite and indefinite order [29, 30]. Ref. [31] pointed out that establishing a separation requires a resource-theoretic framework where the basic resources are communication devices, assembled into communication networks by a restricted set of allowed operations. However, the specific choice of allowed operations remained open since then, and no communication advantage of indefinite causal order was proven so far.

Here, we provide the first rigorous demonstration of an advantage of indefinite order in quantum communication. To this purpose, we develop a resource theory of communication that treats definite and indefinite order in a unified way. In this resource theory, devices that can transmit signals are regarded as a resource, and operations that do not generate signaling from signaling-incapable devices are regarded as allowed. Among the allowed operations, we show that operations with indefinite order offer advantages in one-shot communication tasks. Specifically, we show that indefinite order enables a noiseless, single-shot transmission of a bit through two amplitude damping channels, whereas all the operations with definite order yield a noisy communication. This result establishes a clear-cut separation between the communication power of definite and indefinite causal order.

We also show, however, that the communication advantage of indefinite causal order is not generic to all quantum channels and to all communication tasks. In particular, we show that indefinite order does not offer any advantage for the class of Pauli channels, whose maximum classical capacity can already be reached by placing the channels in parallel and exploiting no-signaling resources. Furthermore, we show that every advantage of indefinite causal order in classical communication must necessarily disappear in the asymptotic limit of a large number of channel uses. In this asymptotic scenario, we prove that shared entanglement between the sender and receiver is sufficient to achieve the highest possible value of the classical channel capacity. Our results reveal that the landscape of relations between communication, causal order, entanglement, and no-signaling resources is more complex, and more profound than previously envisaged.

II. RESULTS

The resource theory of signaling. Here we develop a resource theory where the key resource is the ability to send signals. In this resource theory, the basic objects are communication devices, and the basic operations assemble communication devices into links connecting a sender to a receiver. Mathematically, the basic resources are represented by lists of quantum channels (completely positive, trace preserving maps transforming the density matrix of an input system into the density matrix of an output system.) For a list of N channels, we will use the vector notation $\vec{\mathcal{C}} = (\mathcal{C}_1, \dots, \mathcal{C}_N)$, and we will denote by X_i (Y_i) the input (output) system of the i-th channel in the list.

The resource theory is constructed by first specifying a set of devices which are regarded as non-resourceful, or free, and then by considering the largest set of operations that do not create resources [32]. Here, we focus on the resource of signaling: we regard as free all the devices that block any signal, by outputting a fixed state independent of their input. Mathematically, these devices are described by replacement channels, that is, quantum channels $\mathcal{C}$ satisfying the condition $\mathcal{C}(\rho) = \rho_0 \,\forall \rho$, where ρ is an arbitrary density matrix of the input system, and ρ_0 is a fixed density matrix of the output system. In general, free resources are described by lists of replacement channels.

Let us now specify the allowed operations that can be used to assemble quantum channels into communication links. Mathematically, operations on quantum channels are described by quantum supermaps [3, 33–35], that is, linear maps transforming (lists of) quantum channels into quantum channels. In general, a supermap can transform a list of N channels into another list of M channels. In the following, we will focus on the case of $M = 1$, corresponding to the case in which the N input channels in the initial list are assembled to build a communication channel connecting a sender to a receiver. The input and output systems of the channel produced by the supermap will be denoted by A and B, respectively.

The largest set of operations that do not generate resources is the set of signaling-non-generating supermaps, that is, quantum supermaps that transform lists of replacement channels into replacement channels. In our resource theory, signaling-non-generating supermaps will be regarded as free operations. In the special case where each list consists of a single channel, the set of signaling-non-generating supermaps has been known as "no-signaling resources" and has been fruitfully adopted in quantum Shannon theory [36–38]. No-signaling resources have been recently cast in a resource-theoretic framework [39, 40], and our resource theory can be regarded as the extension of this resource-theoretic framework to the case where multiple input channels can be assembled in a non-trivial order.

A complete characterization of the set of free operations is provided in Methods, where we show that signaling-non-generating supermaps are in one-to-one correspondence with a subset of multipartite channels satisfying a no-signaling condition, which we call the "No Forward Signaling Condition." Explicitly, a signaling-non-generating supermap $\mathscr{S}$ transforming a list of N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a channel with input/output pair (A, B) corresponds to a quantum channel with input system AY and output system BX (with the notation $X := X_1 \cdots X_N$ and $Y := Y_1 \cdots Y_N$), satisfying the condition that no signal can be sent from A to B. Informally, the No Forward Signaling Condition states that the supermap does not allow signals to travel directly from a sender to a receiver, bypassing the devices that are used to build the communication link. This condition guarantees that our theory satisfies a requirement known as no-side channel generation, previously put forward as a minimal requirement that any resource theory of communication should satisfy [31].

The resource theory where the free resources are replacement channels and the free operations are the signaling-non-generating supermaps will be called the "resource theory of signaling." With the set of free operations in place, we are now in position to rigorously analyze the role of indefinite causal order. To this purpose, we define three subsets of the set of free operations , ordered by mutual inclusion. The smallest set is the set of free supermaps that can be achieved by placing the N input channels in parallel inside a quantum circuit [34]. The second set is the set of free supermaps that can be achieved by placing the N input channel inside a quantum circuit in a fixed, pre-assigned order. These supermaps are known as quantum combs [34, 41]. The third set is the set of free supermaps that can be achieved by placing the N input channels inside a quantum circuit, in an order that is definite but possibly dynamically determined at subsequent steps of the circuit. These supermaps correspond to quantum circuits with classical control [42, 43], and represent the most general class of operations with definite causal order. In the following, the sets of free supermaps with parallel placement, fixed-order placement, and definite (dynamical) order placement will be denoted by $\mathbf{FreePar}$, $\mathbf{FreeFix}$, and $\mathbf{FreeDef}$, respectively. One has the (generally strict) inclusions $\mathbf{FreePar} \subseteq \mathbf{FreeFix} \subseteq \mathbf{FreeDef} \subseteq \mathbf{Free}$, where $\mathbf{Free}$ denotes the set of general supermaps that are free in our resource theory. Free supermaps that are in $\mathbf{Free}$ but not in $\mathbf{FreeDef}$ are free operations with indefinite causal order. In the following, we will show that using general free operations with indefinite causal order offers an advantage in one of the most fundamental communication tasks: the transmission of bits from a sender to a receiver.

Communication advantages of indefinite causal order. Suppose that Alice wants to transmit classical messages to Bob, using a communication link built from N communication devices. We now show that assembling the devices in an indefinite order can increase the number of bits Alice can reliably send to Bob.

Let us start from the ideal scenario in which the communication is free from errors. The task of zero-error classical communication can be formulated as the task of converting N communication devices into a noiseless classical channel by means of free operations. Mathematically, the conversion is implemented by a supermap $\mathcal{S}$ that transforms a list of N quantum channels $\vec{\mathcal{C}}$ into a noiseless classical channel Δ^m, of the form $\Delta^m : \rho \mapsto \sum_{j=1}^m \langle j|\,\rho\,|j\rangle\ |j\rangle\langle j|$, where m is the number of classical messages the channel can send and $\{|j\rangle\}_{j=1}^m$ is a fixed orthonormal basis for the vector space $\mathbb{C}^m$.

Now, the key question is: what is the maximum number of messages Alice can communicate to Bob in an error-free way? When the N communication devices are assembled using operations in the set $\mathbf{X}$ (where $\mathbf{X}$ can be $\mathbf{FreePar}$, or $\mathbf{FreeFix}$, or $\mathbf{FreeDef}$, or $\mathbf{Free}$), the maximum number of messages is $m_0^{\mathbf{X}}(\vec{\mathcal{C}}) := \max\{m \,|\, \exists \mathcal{S} \in \mathbf{X} : \mathcal{S}(\vec{\mathcal{C}}) = \Delta^m\}$. Due to the inclusion $\mathbf{FreeDef} \subset \mathbf{Free}$, the inequality $m_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) \leq m_0^{\mathbf{Free}}(\vec{\mathcal{C}})$ holds for every list of channels $\vec{\mathcal{C}}$. We now show that the inequality is generally strict: assembling the devices in an indefinite causal order generally increases the number of messages Alice can communicate to Bob in an error-free way. In fact, we show that there exists scenarios in which error-free communication is impossible when the channels are combined in definite order, and yet a non-zero number of bits can be perfectly transmitted when the channels are combined in an indefinite order.

Let us introduce the notion of one-shot zero-error capacity of the channels in the list $\vec{\mathcal{C}}$ assisted by signaling-non-generating operations in the set $\mathbf{X}$, defined as $C_0^{\mathbf{X}}(\vec{\mathcal{C}}) := \log_2 m_0^{\mathbf{X}}(\vec{\mathcal{C}})$. Then, consider the case of $N = 2$ amplitude damping channels, that is, channels of the form $\mathcal{A} : \rho \mapsto \mathcal{A}(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$, where ρ is an arbitrary density matrix of a single-qubit, and A_0 and A_1 are the 2-by-2 matrices given by $A_0 := \sqrt{\eta}|0\rangle\langle 1|$ and $A_1 := \sqrt{I - A_0^\dagger A_0}$, where $\eta \in [0,1]$ is the damping parameter. We then evaluate the maximum number of messages that Alice can transmit to Bob, setting $N = 2$ and $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{A}$. In Methods, we show that the evaluation can be reduced to a semidefinite program.

Using the semidefinite programs for $C_0^{\mathbf{FreeDef}}$ and $C_0^{\mathbf{Free}}$, we can now pin down the advantage of indefinite causal order both analytically and numerically. On the analytical side, we prove that for damping parameter $\eta = 0.1$, the zero-error capacity assisted by definite causal order is

$$C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = 0\,, \tag{1}$$

meaning that no information can be communicated without error whenever the two channels are combined in a definite order. In start contrast, we find that the zero-error capacity assisted by indefinite causal order is

$$C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = 1\,, \tag{2}$$

meaning that one bit of information can be communicated without any error if the two channels are combined in an indefinite order. Numerically, we determine the values of the zero-error capacity for values of damping parameter η between 0 and 0.5, showing that the advantage of indefinite causal order is present for all values of η in the interval between $\eta = 0$ (excluded) and $\eta \approx 0.2$, as illustrated in Fig. 1. Together, these results provide the first rigorous evidence of an advantage of indefinite causal order in quantum communication.

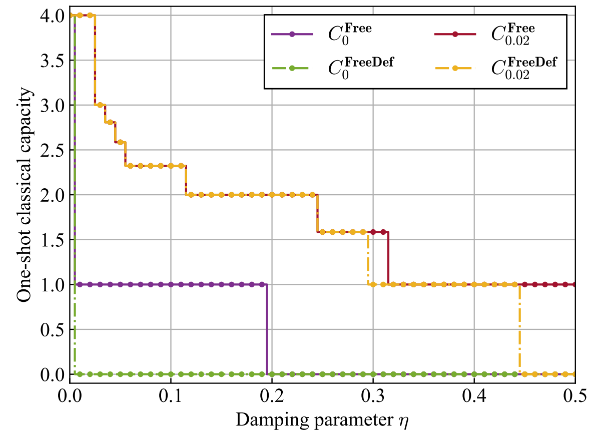

Fig 1. **Advantages of indefinite causal order in zero-error and bounded-error communications.** We compare the one-shot classical capacities of a pair of amplitude damping channels $(\mathcal{A}, \mathcal{A})$ assisted by $\mathbf{Free}$ and $\mathbf{FreeDef}$ in both the zero-error case and the case with an error tolerance of 0.02.

A similar approach can be used to derive advantages of indefinite causal order in classical communication beyond the zero-error scenario. The maximum number of messages that can be transmitted with error upper bounded by ϵ, using the channels $\vec{\mathcal{C}}$ assembled by operations in $\mathbf{X}$, is given by $m_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}) := \max\{m \,|\, \exists \mathcal{S} \in \mathbf{X} : \|\mathcal{S}(\vec{\mathcal{C}}) - \Delta^m\|_\diamond/2 \leq \epsilon\}$, where $\|\cdot\|_\diamond$ denotes the diamond norm [44]. We then introduce the one-shot capacity $C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}) := \log_2 m_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}})$, and show that, in general, the maximum one-shot capacity achieved by signaling-non-generating operations with definite causal order is strictly smaller than the one-shot capacity achievable by signaling-non-generating operations with indefinite causal order. For $N = 2$ amplitude damping channels, an illustration of this fact is provided in Fig. 1.

No advantage for Pauli channels. We now show that the advantage of indefinite causal order shown in the previous section is far from generic: no advantage is present for a large class of channels, known as Pauli channels [45].

A Pauli channel acting on q qubits is a channel of the form $\mathcal{P}(\rho) = \sum_{i_1 \dots i_q} p_{i_1 \dots i_q} (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})\, \rho\, (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})$, where σ_0 is the identity matrix, $\sigma_1, \sigma_2, \sigma_3$ are the three Pauli matrices, and $(p_{i_1 \dots i_q})$ is a probability distribution, providing the probability that the q qubits undergo the product gate $\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q}$. Here we consider the situation where the initial communication devices are N, possibly different, Pauli channels $\vec{\mathcal{P}} := (\mathcal{P}_1, \dots, \mathcal{P}_N)$, possibly acting on different numbers of qubits. In this case, we prove that the ability to control the order of the channels does not enhance the classical capacity, even in the presence of an error tolerance. In fact, we prove an even stronger result: placing the N channels in parallel with the assistance of no-signaling resources is already sufficient to achieve the largest possible value of the capacity; in formula, $C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) = C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}})$ for every $\epsilon \geq 0$, where $\vec{\mathcal{P}} = (\mathcal{P}_1, \dots, \mathcal{P}_N)$ are the N initial channels (See Supplementary Note E for the proof). In other words, neither definite nor indefinite causal order can enhance the capacity beyond the maximum value achievable by placing the channels in parallel with the assistance of no-signaling resources.

No advantage for channel simulation. We now provide another no-go result highlighting a limitation of the power of indefinite causal order. Explicitly, we show that no advantage arises in the task of simulating a noisy channel through classical communication assisted by no-signaling resources [39, 46]. This task is the inverse of the classical communication task: while in communication one uses a noisy channel to simulate a noiseless channel for the transmission of bits, in simulation one uses the noiseless transmission of bits to simulate a noisy channel.

In our resource-theoretic framework, the task of channel simulation can be phrased as the problem of converting a set of ideal classical channels into a target channel $\mathcal{C}$ by means of signaling-non-generating operations. Let us denote the list of ideal classical channels as $\vec{\Delta}^m = (\Delta^{m_1}, \Delta^{m_2}, \dots)$, where m_i is the number of messages transmitted by the i-th channel, and $m := \sum_j m_j$ is the total number of messages transmitted. With this notation, the classical simulation cost for one realization of channel $\mathcal{C}$ assisted by signaling-non-generating operations in the set $\mathbf{X}$ is $S_\epsilon^{\mathbf{X}}(\mathcal{C}) := \min\{\log_2 m \,|\, \exists \mathcal{S} \in \mathbf{X} : \|\mathcal{S}(\vec{\Delta}^m) - \mathcal{C}\|_\diamond/2 \leq \epsilon\}$, which measures the minimum number of classical bits consumed by the simulation. This type of one-shot cost is operationally widely used in quantum resource theories [47, 48], and has played an important role in the study of entanglement [49, 50], coherence [51, 52], and communication [36, 39].

Since enlarging the set of allowed operations can in principle reduce the simulation cost, one has the inequalities $S_\epsilon^{\mathbf{FreePar}}(\mathcal{C}) \geq S_\epsilon^{\mathbf{FreeFix}}(\mathcal{C}) \geq S_\epsilon^{\mathbf{FreeDef}}(\mathcal{C}) \geq S_\epsilon^{\mathbf{Free}}(\mathcal{C})$ for every error tolerance $\epsilon \geq 0$. Surprisingly, however, we found out that this chain of inequalities collapses to a single equality: $S_\epsilon^{\mathbf{FreePar}}(\mathcal{C}) = S_\epsilon^{\mathbf{Free}}(\mathcal{C})$ for every $\epsilon \geq 0$ (see Supplementary Note F for the derivation). In other words, indefinite causal order does not offer any advantage over parallel operations in reducing the simulation cost: for the task of simulating a noisy channel with no-signaling resources, all the classical communication can be sent in a single round.

No classical communication advantage in the asymptotic scenario. We now consider the scenario in which Alice and Bob communicate classical messages through an asymptotically large number of uses of the same quantum channel $\mathcal{C}$. In our resource-theoretic framework, the problem can be formalized as follows: Alice and Bob communicate through a link established by combining N communication devices implementing the same quantum channel, namely $\vec{\mathcal{C}}_N = (\mathcal{C}_1, \dots, \mathcal{C}_N)$ with $\mathcal{C}_j = \mathcal{C}$ for every $j \in \{1, \dots, N\}$. Then, the classical capacity of the channel $\mathcal{C}$ assisted by operations in the set $\mathbf{X}$ is defined as the maximum number of bits communicated reliably per channel use, when the channels are assembled by signaling-non-generating operations in $\mathbf{X}$. In formula, $C^{\mathbf{X}} := \lim_{\epsilon\to 0} \lim_{N\to\infty} C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}_N)/N$.

We now show that the order of the channels, either definite or indefinite, does not affect the capacity in the asymptotic limit. The proof is based on the quantum reverse Shannon theorem [53, 54], which states the equality $C^{\mathbf{Ent}}(\mathcal{C}) = S^{\mathbf{Ent}}(\mathcal{C})$, where $C^{\mathbf{Ent}}(\mathcal{C})$ and $S^{\mathbf{Ent}}(\mathcal{C})$ denote the entanglement-assisted classical capacity and the entanglement-assisted classical simulation cost of the channel $\mathcal{C}$, respectively.

To prove that indefinite causal order does not offer any classical capacity advantage in the asymptotic scenario, we consider the asymptotic classical simulation cost assisted by supermaps in the set $\mathbf{X}$, defined as $S^{\mathbf{X}}(\mathcal{C}) := \lim_{\epsilon\to 0} \lim_{N\to\infty} S_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}_N)$. Since shared entanglement is a special case of a no-signaling resource, we have the inequalities $C^{\mathbf{Ent}}(\mathcal{C}) \leq C^{\mathbf{FreeDef}}(\mathcal{C}) \leq C^{\mathbf{Free}}(\mathcal{C})$ and $S^{\mathbf{Free}}(\mathcal{C}) \leq S^{\mathbf{FreeDef}}(\mathcal{C}) \leq S^{\mathbf{Ent}}(\mathcal{C})$. Moreover, the classical capacity of a channel cannot be larger than its simulation cost, whence we have the inequality $C^{\mathbf{Free}}(\mathcal{C}) \leq S^{\mathbf{Free}}(\mathcal{C})$. Combining the above inequality with the quantum reverse Shannon theorem, we then obtain the chain of equalities $C^{\mathbf{Ent}}(\mathcal{C}) = C^{\mathbf{Free}}(\mathcal{C}) = S^{\mathbf{Free}}(\mathcal{C}) = S^{\mathbf{Ent}}(\mathcal{C})$. This chain of equalities has two implications. The first implication is that channel transformations under signaling-non-generating supermaps are reversible even if indefinite causal order is allowed. The second implication is that neither definite nor indefinite causal order provide any advantage in the asymptotic regime: the assistance of entanglement is already sufficient to achieve the maximum possible classical capacity.

Note that the above argument also holds for the communication of quantum information. By teleportation [55] and superdense coding [56], both the entanglement-assisted quantum capacity $Q^{\mathbf{Ent}}$ and the entanglement-assisted quantum simulation cost $S_Q^{\mathbf{Ent}}$ are half of their classical counter part: $Q^{\mathbf{Ent}} = \frac{1}{2}C^{\mathbf{Ent}} = \frac{1}{2}S^{\mathbf{Ent}} = S_Q^{\mathbf{Ent}}$. Applying the same logic, this implies that neither definite nor indefinite causal order can enhance quantum communication in the asymptotic regime.

III. DISCUSSION

In this work, we introduced a rigorous framework for investigating the communication advantages of indefinite causal order. Our framework consists of resource theory of signaling, where the free resources are the replacement channels, and the set of free operations consists of the signaling-non-generating supermaps. In this framework, one can examine the performance of different causal structures within the set of free supermaps, and, when a gap arises, one can rigorously attribute it to the different causal structures.

Despite active research in quantum resource theories for various types of dynamical resources [40, 57–66] and general resource-independent approaches [67–73], supermaps with general causal structures has remain relatively unexplored in the resource-theoretic setting, with a few exception like [74]. In this respect, our framework provide a first study of the role of different causal structures within the set of free operations.

Although the focus of the present work was on communication, the approach developed here can also be applied to other quantum resources. For example, quantum SWITCH has been shown to enhance cooling a cold reservoir [75] and preserving nonstabilizerness [76]. However, whether such enhancements can be attributed to indefinite causal order requires further investigation. An interesting direction of future research is to employ our paradigm to explore the advantages of indefinite causal order in the context of other resource theories, such as thermodynamics and magic.

IV. METHODS

Characterization of supermaps. It is known that quantum supermaps are in one-to-one correspondence with quantum channels. Specifically, a supermap mapping N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a quantum channel with input A and output B corresponds to a multipartite quantum channel with input YA and output XB, where $X := X_1 \cdots X_N$ and $Y := Y_1 \cdots Y_N$. According to the Choi-Jamiołkowski isomorphism [77, 78], every channel can be represented by a unique linear operator, known as its Choi operator, which can also serve as a representation of the supermap corresponding to this channel.

The requirement of different classes of causal orders can be expressed as linear constraints imposed on Choi operators. In particular, the characterization of causally fixed supermaps (i.e., quantum combs) in terms of their Choi operators was introduced in Ref. [34], and the characterization of causally definite supermaps was studied in Ref. [43]. Such characterizations facilitate the formulation of various quantities studied in our work as optimization programs over linear operators. See Supplementary Note B for details of these characterizations.

The characterization of supermaps with general, possibly indefinite, causal structures in terms of their Choi operators was first studied in Ref. [2] in the language of "process matrices." A set of linear constraints describing the Choi operators of general supermaps was later given in Ref. [79]. In Supplementary Note B, we show that the aforementioned characterization is equivalent to a simpler one that has a clear physical interpretation. In particular, we prove that a linear operator J_{XYAB} is the Choi operator of some general supermap if and only if

$$J_{XYAB} \geq 0 \quad \text{and} \quad \mathcal{L}_{XY}^{\text{NS}}\left(J_{XYA}\right) = \frac{I_{XYA}}{d_X}, \tag{3}$$

where $J_{XYA} := \text{Tr}_B[J_{XYAB}]$, and $\mathcal{L}^{\text{NS}}$ is the orthogonal projection onto the subspace of Hermitian operators spanned by the Choi operators of no-signaling multipartite quantum channels.

Note that supermaps in $\mathbf{FreePar}$, $\mathbf{FreeFix}$, $\mathbf{FreeDef}$, and $\mathbf{Free}$ should not only satisfy the causal order constraints for their respective classes, but also be signaling-non-generating. This latter requirement enables us to fairly compare how well supermaps in different classes make use of the given communication channel. In Supplementary Note A, we prove that a supermap is signaling-non-generating if and only if it satisfies the No Forward Signaling Condition, namely, its corresponding channel is no-signaling from A to B. This No Forward Signaling Condition can be conveniently expressed as a constraint on the supermap's (also the corresponding channel's) Choi operator J_{XYAB}:

$$\text{Tr}_X\left[J_{XYAB}\right] = \text{Tr}_{XA}\left[J_{XYAB}\right] \otimes \frac{I_A}{d_A}, \tag{4}$$

where d_A denotes the dimension of system A. This constraint says that the output at system B (tracing out the other output system X) does not depend on the input at system A.

One-shot classical capacities as semidefinite programs. Given a list of quantum channels $\vec{\mathcal{C}}$, classical communication of m messages is to find an admissible code, represented by a supermap $\mathscr{S}$, such that $\mathscr{S}$ transforms the channels $\vec{\mathcal{C}}$ into a channel as close as possible to Δ^m, the m-dimensional noiseless classical channel. Specifically, the desired supermap $\mathscr{S}$ should achieve the minimum error $\omega^{\mathbf{X}}(\vec{\mathcal{C}}, \Delta^m)$ defined as $\omega^{\mathbf{X}}(\vec{\mathcal{C}}, \mathcal{M}) := \min\{\|\mathscr{S}(\vec{\mathcal{C}}) - \mathcal{M}\|_\diamond/2 \,|\, \mathscr{S} \in \mathbf{X}\}$, where $\|\cdot\|_\diamond$ denotes the diamond norm [44], measuring the closeness of two quantum channels, $\mathbf{X}$ is the set of allowed supermaps, and $\mathcal{M}$ refers to a target quantum channel. Based on this minimum error, we define the one-shot classical capacity of a given list of quantum channels $\vec{\mathcal{C}}$ assisted by supermaps in $\mathbf{X}$ as

$$C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}) := \max\{\log_2 m \,|\, \omega^{\mathbf{X}}(\vec{\mathcal{C}}, \Delta^m) \leq \epsilon\}, \tag{5}$$

where m is maximized over all positive integers, and $\epsilon \geq 0$ is a given error tolerance. This one-shot classical capacity quantifies the maximum number of classical bits that can be reliably transmitted through a single use of $\vec{\mathcal{C}}$ when assisted by supermaps in $\mathbf{X}$.

It is not straightforward to see if $C_\epsilon^{\mathbf{X}}$ can be formulated as a semidefinite program. One major obstacle is the involvement of the diamond norm. Instead of diamond distance, average probability of error has been commonly used in classical communication. Given a quantum channel $\mathcal{M}$, the average probability of error of transmitting m equiprobable classical messages through $\mathcal{M}$ is $p(\mathcal{M}, m) := 1 - \frac{1}{m}\sum_{j=1}^m \text{Tr}\left[\mathcal{M}\left(|j\rangle\!\langle j|\right)|j\rangle\!\langle j|\right]$. Then, we define the minimum average probability of error of a list of noisy channel $\vec{\mathcal{C}}$ assisted by supermaps in $\mathbf{X}$ as $p^{\mathbf{X}}(\vec{\mathcal{C}}, m) := \min\{p(\mathscr{S}(\vec{\mathcal{C}}), m) \,|\, \mathscr{S} \in \mathbf{X}\}$. Similar to Eq. (5), one can define the one-shot classical capacity using $p^{\mathbf{X}}(\vec{\mathcal{C}}, m)$, which can be formulated as a semidefinite program for $\mathbf{X} = \mathbf{FreePar}$ [80, 81].

In general, given a noisy channel $\mathcal{M}$, its diamond distance $\frac{1}{2}\left\|\mathcal{M} - \Delta^m\right\|_\diamond$ differs from its average probability of error $1 - \frac{1}{m}\sum_{j=0}^{m-1} \text{Tr}\left[\mathcal{M}\left(|j\rangle\!\langle j|\right)|j\rangle\!\langle j|\right]$. However, in Supplementary Note C, we show that these two quantities are actually equivalent after minimized over a set of supermaps if the set is convex and closed under simultaneous permutations on A and B. These generic conditions are satisfied by the sets of free supermaps considered in this work, including the no-signaling-assisted codes in the standard communication scenario. This equivalence, together with the characterizations of supermaps in terms of their Choi operators and the symmetry of the noiseless classical channel Δ^m, allows us to formulate $C_\epsilon^{\mathbf{X}}$ as semidefinite programs for $\mathbf{X}$ considered in this work. This also enables us to use a unified measure of error, the diamond distance, to address both communication and noisy channel simulation.

In particular, the one-shot classical capacity assisted by supermaps in $\mathbf{Free}$ can be formulated as the following semidefinite program:

$$\begin{aligned} C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \max\ & \lfloor m \rfloor \\ \text{s.t. } & E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \geq m(1-\epsilon), \\ & E_Y = I_Y,\, 0 \leq E_{XY} \leq F_{XY}, \\ & \mathcal{L}_{XY}^{\mathrm{NS}}(F_{XY}) = \frac{m}{d_X} I_{XY}, \end{aligned} \tag{6}$$

where m is maximized over real numbers, and Hermitian matrices E_{XY} and F_{XY} are optimization variables. The $\star$ operation, known as link product, acting on two linear operators P_{AB} and Q_{BC} is defined as $P_{AB} \star Q_{BC} := \mathrm{Tr}_B\left[(P_{AB} \otimes I_C)\left(I_A \otimes Q_{BC}^{\mathrm{T}_B}\right)\right]$, where T_B is the partial transpose over system B. It gives the Choi operator of the quantum process (channel or supermap) resulted from composing two quantum processes with Choi operators P_{AB} and Q_{BC} [34, 41]. Similarly, the one-shot classical capacity assisted by supermaps in $\mathbf{FreeDef}$ can also be formulated as a semidefinite program. The derivation of both semidefinite programs is presented in Supplementary Note C. These semidefinite program formulations allow us to rigorously prove the advantage of indefinite causal order (see Supplementary Note D for the proof) as well as to numerically compute the capacities for different $\vec{\mathcal{C}}$, as we did in Fig. 1.

ACKNOWLEDGMENTS

We thank Jessica Bavaresco, Cyril Branciard, Timothée Hoffreumon, Marco Túlio Quintino, Bartosz Reguła, Xin Wang, and Chengkai Zhu for helpful discussions. This work is supported by the Chinese Ministry of Science and Technology (MOST) through grant 2023ZD0300600, by the Hong Kong Research Grant Council (RGC) through grants SRFS2021-7S02 and R7035-21F, and by the State Key Laboratory of Quantum Information Technologies and Materials, Chinese University of Hong Kong. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron, "Beyond quantum computers," arXiv:0912.0195 (2009).
[2] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner, "Quantum correlations with no causal order," Nature Communications **3**, 1092 (2012).
[3] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron, "Quantum computations without definite causal structure," Physical Review A **88**, 022318 (2013).
[4] Lucien Hardy, "Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure," Journal of Physics A: Mathematical and Theoretical **40**, 3081 (2007).
[5] Lucien Hardy, "Quantum gravity computers: On the theory of computation with indefinite causal structure," in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony*, edited by Wayne C. Myrvold and Joy Christian (Springer Netherlands, Dordrecht, 2009) pp. 379–401.
[6] Christopher A Fuchs, "Quantum mechanics as quantum information, mostly," Journal of Modern Optics **50**, 987–1023 (2003).
[7] Gilles Brassard, "Is information the key?" Nature Physics **1**, 2–4 (2005).
[8] Giulio Chiribella and Robert W. Spekkens, eds., *Quantum theory: informational foundations and foils*, Fundamental theories of physics, volume 181 (Springer, Dordrecht, 2016).
[9] Giulio Chiribella, "Perfect discrimination of no-signalling channels via quantum superposition of causal structures," Physical Review A **86**, 040301 (2012).
[10] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino, "Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination," Physical Review Letters **127**, 200504 (2021).
[11] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino, "Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies," Journal of Mathematical Physics **63**, 042203 (2022).
[12] Alastair A Abbott, Mehdi Mhalla, and Pierre Pocreau, "Quantum query complexity of boolean functions under indefinite causal order," Physical Review Research **6**, L032020 (2024).
[13] Alastair A Abbott, Mehdi Mhalla, and Pierre Pocreau, "Classical and quantum query complexity of boolean functions under indefinite causal order," in *Proceedings of the 22nd International Conference on Quantum Physics and Logic* (2025).
[14] Philippe Allard Guérin, Adrien Feix, Mateus Araújo, and Časlav Brukner, "Exponential communication complexity advantage from quantum superposition of the direction of communication," Physical Review Letters **117**, 100502 (2016).
[15] Kejin Wei, Nora Tischler, Si-Ran Zhao, Yu-Huai Li, Juan Miguel Arrazola, Yang Liu, Weijun Zhang, Hao Li, Lixing You, Zhen Wang, *et al.*, "Experimental quantum switching for exponentially superior quantum communication complexity," Physical Review Letters **122**, 120504 (2019).
[16] Xiaobin Zhao, Yuxiang Yang, and Giulio Chiribella, "Quantum metrology with indefinite causal order," Physical Review Letters **124**, 190503 (2020).
[17] Peng Yin, Xiaobin Zhao, Yuxiang Yang, Yu Guo, Wen-Hao Zhang, Gong-Chu Li, Yong-Jian Han, Bi-Heng Liu, Jin-Shi Xu, Giulio Chiribella, *et al.*, "Experimental super-heisenberg quan-

tum metrology with indefinite gate order," Nature Physics **19**, 1122–1127 (2023).

[18] Qiushi Liu, Zihao Hu, Haidong Yuan, and Yuxiang Yang, "Optimal strategies of quantum metrology with a strict hierarchy," Physical Review Letters **130**, 070803 (2023).

[19] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao, "Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations," Physical Review Letters **123**, 210502 (2019).

[20] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao, "Probabilistic exact universal quantum circuits for transforming unitary operations," Physical Review A **100**, 062339 (2019).

[21] Marco Túlio Quintino and Daniel Ebler, "Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality," Quantum **6**, 679 (2022).

[22] Claude E Shannon, "A mathematical theory of communication," The Bell System Technical Journal **27**, 379–423 (1948).

[23] Claude Shannon, "The zero error capacity of a noisy channel," IRE Transactions on Information Theory **2**, 8–19 (1956).

[24] Daniel Ebler, Sina Salek, and Giulio Chiribella, "Enhanced communication with the assistance of indefinite causal order," Physical Review Letters **120**, 120502 (2018).

[25] Sina Salek, Daniel Ebler, and Giulio Chiribella, "Quantum communication in a superposition of causal orders," arXiv:1809.06655 (2018).

[26] K Goswami, Y Cao, GA Paz-Silva, J Romero, and AG White, "Increasing communication capacity via superposition of order," Physical Review Research **2**, 033292 (2020).

[27] Giulio Chiribella, Manik Banik, Some Sankar Bhattacharya, Tamal Guha, Mir Alimuddin, Arup Roy, Sutapa Saha, Sristy Agrawal, and Guruprasad Kar, "Indefinite causal order enables perfect quantum communication with zero capacity channels," New Journal of Physics **23**, 033039 (2021).

[28] Giulio Chiribella, Matt Wilson, and HF Chau, "Quantum and classical data transmission through completely depolarizing channels in a superposition of cyclic orders," Physical Review Letters **127**, 190502 (2021).

[29] Alastair A Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, and Cyril Branciard, "Communication through coherent control of quantum channels," Quantum **4**, 333 (2020).

[30] Philippe Allard Guérin, Giulia Rubino, and Časlav Brukner, "Communication through quantum-controlled noise," Physical Review A **99**, 062317 (2019).

[31] Hlér Kristjánsson, Giulio Chiribella, Sina Salek, Daniel Ebler, and Matthew Wilson, "Resource theories of communication," New Journal of Physics **22**, 073014 (2020).

[32] Eric Chitambar and Gilad Gour, "Quantum resource theories," Reviews of Modern Physics **91**, 025001 (2019).

[33] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti, "Transforming quantum operations: Quantum supermaps," Europhysics Letters **83**, 30004 (2008).

[34] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti, "Quantum circuit architecture," Physical Review Letters **101**, 060401 (2008).

[35] Alessandro Bisio and Paolo Perinotti, "Theoretical framework for higher-order quantum theory," Proceedings of the Royal Society A **475**, 20180706 (2019).

[36] Runyao Duan and Andreas Winter, "No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the lovász number," IEEE Transactions on Information Theory **62**, 891–914 (2016).

[37] Xin Wang and Runyao Duan, "On the quantum no-signalling assisted zero-error classical simulation cost of non-commutative bipartite graphs," in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2016) pp. 2244–2248.

[38] Xin Wang and Runyao Duan, "Separation between quantum lovász number and entanglement-assisted zero-error classical capacity," IEEE Transactions on Information Theory **64**, 1454–1460 (2018).

[39] Kun Fang, Xin Wang, Marco Tomamichel, and Mario Berta, "Quantum channel simulation and the channel's smooth max-information," IEEE Transactions on Information Theory **66**, 2129–2140 (2020).

[40] Ryuji Takagi, Kun Wang, and Masahito Hayashi, "Application of the resource theory of channels to communication scenarios," Physical Review Letters **124**, 120502 (2020).

[41] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, "Theoretical framework for quantum networks," Physical Review A **80**, 022339 (2009).

[42] Ognyan Oreshkov and Christina Giarmatzi, "Causal and causally separable processes," New Journal of Physics **18**, 093020 (2016).

[43] Julian Wechs, Hippolyte Dourdent, Alastair A Abbott, and Cyril Branciard, "Quantum circuits with classical versus quantum control of causal order," PRX Quantum **2**, 030335 (2021).

[44] A Yu Kitaev, "Quantum computations: algorithms and error correction," Russian Mathematical Surveys **52**, 1191 (1997).

[45] Mark M Wilde, *Quantum information theory* (Cambridge university press, 2013).

[46] Charles H Bennett, Peter W Shor, John A Smolin, and Ashish V Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem," IEEE Transactions on Information Theory **48**, 2637–2655 (2002).

[47] Zi-Wen Liu, Kaifeng Bu, and Ryuji Takagi, "One-shot operational quantum resource theory," Physical Review Letters **123**, 020401 (2019).

[48] Ryuji Takagi, Bartosz Regula, and Mark M Wilde, "One-shot yield-cost relations in general quantum resource theories," PRX Quantum **3**, 010348 (2022).

[49] Francesco Buscemi and Nilanjana Datta, "Entanglement cost in practical scenarios," Physical Review Letters **106**, 130503 (2011).

[50] Xin Wang and Mark M Wilde, "Cost of quantum entanglement simplified," Physical Review Letters **125**, 040502 (2020).

[51] Qi Zhao, Yunchao Liu, Xiao Yuan, Eric Chitambar, and Xiongfeng Ma, "One-shot coherence dilution," Physical Review Letters **120**, 070403 (2018).

[52] María García Díaz, Kun Fang, Xin Wang, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Andreas Winter, "Using and reusing coherence to realize quantum processes," Quantum **2**, 100 (2018).

[53] Charles H Bennett, Igor Devetak, Aram W Harrow, Peter W Shor, and Andreas Winter, "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels," IEEE Transactions on Information Theory **60**, 2926–2959 (2014).

[54] Mario Berta, Matthias Christandl, and Renato Renner, "The quantum reverse shannon theorem based on one-shot information theory," Communications in Mathematical Physics **306**, 579–615 (2011).

[55] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," Physical Review Letters **70**, 1895

(1993).
[56] Charles H Bennett and Stephen J Wiesner, "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states," Physical Review Letters **69**, 2881 (1992).
[57] Khaled Ben Dana, María García Díaz, Mohamed Mejatty, and Andreas Winter, "Resource theory of coherence: Beyond states," Physical Review A **95**, 062327 (2017).
[58] Denis Rosset, Francesco Buscemi, and Yeong-Cherng Liang, "Resource theory of quantum memories and their faithful verification with minimal assumptions," Physical Review X **8**, 021033 (2018).
[59] Thomas Theurer, Dario Egloff, Lijian Zhang, and Martin B Plenio, "Quantifying operations with an application to coherence," Physical Review Letters **122**, 190405 (2019).
[60] Gaurav Saxena, Eric Chitambar, and Gilad Gour, "Dynamical resource theory of quantum coherence," Physical Review Research **2**, 023298 (2020).
[61] Gilad Gour and Carlo Maria Scandolo, "Dynamical entanglement," Physical Review Letters **125**, 180505 (2020).
[62] Ho-Joon Kim, Soojoon Lee, Ludovico Lami, and Martin B Plenio, "One-shot manipulation of entanglement for quantum channels," IEEE Transactions on Information Theory **67**, 5339–5351 (2021).
[63] Gaurav Saxena and Gilad Gour, "Quantifying multiqubit magic channels with completely stabilizer-preserving operations," Physical Review A **106**, 042422 (2022).
[64] Benjamin Stratton, Chung-Yun Hsieh, and Paul Skrzypczyk, "Dynamical resource theory of informational nonequilibrium preservability," Physical Review Letters **132**, 110202 (2024).
[65] Benchi Zhao, Kosuke Ito, and Keisuke Fujii, "Probabilistic channel simulation using coherence," Physical Review Research **6**, 043316 (2024).
[66] Chengkai Zhu, Xuanqiang Zhao, and Xin Wang, "Classical communication cost of a bipartite quantum channel assisted by non-signaling correlations," IEEE Transactions on Information Theory **71**, 6041–6060 (2025).
[67] Zi-Wen Liu and Andreas Winter, "Resource theories of quantum channels and the universal role of resource erasure," arXiv:1904.04201 (2019).
[68] Gilad Gour and Andreas Winter, "How to quantify a dynamical quantum resource," Physical Review Letters **123**, 150401 (2019).
[69] Yunchao Liu and Xiao Yuan, "Operational resource theory of quantum channels," Physical Review Research **2**, 012035 (2020).
[70] Xiao Yuan, Pei Zeng, Minbo Gao, and Qi Zhao, "One-shot dynamical resource theory," arXiv:2012.02781 (2020).
[71] Bartosz Regula and Ryuji Takagi, "One-shot manipulation of dynamical quantum resources," Physical Review Letters **127**, 060402 (2021).
[72] Bartosz Regula and Ryuji Takagi, "Fundamental limitations on distillation of quantum channel resources," Nature Communications **12**, 4411 (2021).
[73] Kaiyuan Ji and Eric Chitambar, "Entropic and operational characterizations of dynamic quantum resources," in *2023 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2023) pp. 755–760.
[74] Simon Milz, Jessica Bavaresco, and Giulio Chiribella, "Resource theory of causal connection," Quantum **6**, 788 (2022).
[75] David Felce and Vlatko Vedral, "Quantum refrigeration with indefinite causal order," Physical Review Letters **125**, 070603 (2020).
[76] Yin Mo, Chengkai Zhu, Zhiping Liu, Mingrui Jing, and Xin Wang, "Enhancement of nonstabilizerness within indefinite causal order," Physical Review A **109**, 062428 (2024).
[77] Andrzej Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," Reports on Mathematical Physics **3**, 275–278 (1972).
[78] Man-Duen Choi, "Completely positive linear maps on complex matrices," Linear Algebra and its Applications **10**, 285–290 (1975).
[79] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner, "Witnessing causal nonseparability," New Journal of Physics **17**, 102001 (2015).
[80] Debbie Leung and William Matthews, "On the power of PPT-preserving and non-signalling codes," IEEE Transactions on Information Theory **61**, 4486–4499 (2015).
[81] Xin Wang, Wei Xie, and Runyao Duan, "Semidefinite programming strong converse bounds for classical capacity," IEEE Transactions on Information Theory **64**, 640–653 (2018).
[82] John Watrous, "Semidefinite programs for completely bounded norms," Theory of Computing **5**, 217–238 (2009).

Supplementary Material for
The quantum communication power of indefinite causal order

Supplementary Note A: Characterization of signaling-non-generating supermaps

Proposition 1 *A general supermap transforming N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a channel with input A and output B is signaling-non-generating if and only if it satisfies the No Forward Signaling Condition: its corresponding quantum channel, with input YA and output XB, is no-signaling from A to B.*

Proof Given a general supermap $\mathscr{S}$, suppose it satisfies the No Forward Signaling Condition, i.e., its corresponding quantum channel $\mathcal{S}$ being no-signaling from A to B. Assume, for the sake of contradiction, that there exists a list of replacement channels $\vec{\mathcal{C}}$ such that $\mathcal{M} := \mathscr{S}(\vec{\mathcal{C}})$ is not a replacement channel. That is, there exists a pair of quantum states ρ, ρ' such that $\mathcal{M}(\rho) \neq \mathcal{M}(\rho')$. Because $\vec{\mathcal{C}}$ is a list of replacement channel, when inserting $\vec{\mathcal{C}}$ into the supermap $\mathscr{S}$, the state sent into the system Y is always fixed, and we denote this fixed state by σ_Y. Hence, $\mathcal{M}(\omega_A) = \mathrm{Tr}_X\left[\mathcal{S}\left(\omega_A \otimes \sigma_Y\right)\right]$ for every state ω_A. Then, $\mathcal{M}(\rho) \neq \mathcal{M}(\rho')$ implies that

$$\mathrm{Tr}_X\left[\mathcal{S}\left(\rho_A \otimes \sigma_Y\right)\right] \neq \mathrm{Tr}_X\left[\mathcal{S}\left(\rho'_A \otimes \sigma_Y\right)\right], \tag{A1}$$

which contradicts with our assumption that $\mathcal{S}$ is no-signaling from A to B. Therefore, $\mathscr{S}$ must map any list of N replacement channels to a replacement channel. That is, every supermap satisfying the No Forward Signaling Condition must be signaling-non-generating.

For the other direction, let us assume that $\mathscr{S}$ does not satisfy the No Forward Signaling Condition. As a result, there exist states ρ_A, ρ'_A, and σ_Y such that

$$\mathrm{Tr}_X\left[\mathcal{S}(\rho_A \otimes \sigma_Y)\right] \neq \mathrm{Tr}_X\left[\mathcal{S}(\rho'_A \otimes \sigma_Y)\right]. \tag{A2}$$

Define $\mathcal{M} := \mathscr{S}(\vec{\mathcal{C}})$, where $\vec{\mathcal{C}}$ is a list of N replacement channels that replaces every state on system X with the state σ_Y. Then, Eq. (A2) implies that $\mathcal{M}(\rho) \neq \mathcal{M}(\rho')$, and thus $\mathscr{S}$ is not signaling-non-generating. Therefore, every signaling-non-generating supermap must satisfy the No Forward Signaling Condition. Together with what we proved in the previous paragraph, we conclude that a supermap is signaling-non-generating if and only if it satisfies the No Forward Signaling Condition. ■

Supplementary Note B: Characterizations of supermaps with different causal structures

In Methods, we mentioned that each supermap mapping N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a quantum channel with input A and output B can be represented by a unique linear operator J_{XYAB}, known as its Choi operator. In this note, we specify the characterizations of supermaps with different causal structures in terms of linear constraints on their Choi operators. These linear constraints, together with the the constraint

$$\mathrm{Tr}_X\left[J_{XYAB}\right] = \mathrm{Tr}_{XA}\left[J_{XYAB}\right] \otimes \frac{I_A}{d_A}, \tag{B1}$$

which imposes the No Forward Signaling Condition, offer convenient characterizations of the sets of free supermaps considered in this work ($\mathbf{FreePar}$, $\mathbf{FreeFix}$, $\mathbf{FreeDef}$, and $\mathbf{Free}$).

1. Causally Fixed Supermaps

Causally fixed supermaps, also known as quantum combs, are transformations between quantum channels that can be implemented by inserting input quantum channels into quantum circuits with open slots in a fixed order. These supermaps consists of both parallel and sequential strategies, where sequential strategies are also known as adaptive strategies sometimes. Such supermaps are introduced in Refs. [33, 34, 41] with the following characterization.

Proposition 2 (Characterization of causally fixed supermaps [34]) *A linear operator J_{XYAB} is the Choi operator of some causally fixed supermap mapping N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a channel with input A and output B if and only if*

$$J_{XYAB} \geq 0, \quad {}_{[1-Y_n]}J_{X_1Y_1\ldots X_nY_nA} = 0 \quad \forall\, n = 1, \ldots, N, \quad J_A = d_Y I_A. \tag{B2}$$

Parallel strategies correspond to the simplest type of quantum combs, the single-slot ones:

$$J_{XYAB} \geq 0, \quad {}_{[1-Y]}J_{XYA} = 0, \quad J_A = d_Y I_A. \tag{B3}$$

2. Causally Definite Supermaps

Causally fixed supermaps are not the most general type of supermaps allowed by definite causal order. By dynamically control the order of input channels in a classical manner, one can realize supermaps that are causally non-fixed but still causally definite [42, 43].

Proposition 3 (Characterization of causally definite supermaps [43]) *A linear operator J_{XYAB} is the Choi operator of some causally definite supermap mapping N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a channel with input A and output B if and only if*

$$\begin{gathered} J_{XYAB} = \sum_{(k_1,\ldots,k_N)} J_{XYAB}^{(k_1,\ldots,k_N)}, \qquad \sum_{(k_1,\ldots,k_N)} J_{A}^{(k_1,\ldots,k_N)} = d_Y I_A, \\ J_{XYAB}^{(k_1,\ldots,k_N)} \geq 0 \quad \forall\,(k_1,\ldots,k_N), \\ {}_{[1-Y_{k_n}]}\left(\sum_{(k_{n+1},\ldots,k_N)} J_{X_{k_1}Y_{k_1}\ldots X_{k_n}Y_{k_n}A}^{(k_1,\ldots,k_n,k_{n+1},\ldots,k_N)}\right) = 0 \quad \forall\, n = 1,\ldots,N-1 \quad \textit{and} \quad \forall\,(k_1,\ldots,k_n), \\ {}_{[1-Y_{k_N}]}J_{XYA}^{(k_1,\ldots,k_N)} = 0 \quad \forall\,(k_1,\ldots,k_N), \end{gathered} \tag{B4}$$

where $(k_1, \ldots, k_N)$ enumerates all ordered sequences consisting of numbers 1 to N with each number appearing once in each sequence.

3. General Supermaps

The most general causal structure allows indefinite causal order [2, 3]. The characterization of general, possibly causally indefinite supermaps was briefly discussed in Methods. The formalism of process matrices was first proposed in Ref. [2] to describe all possible transformations, including causally indefinite ones, from multipartite local quantum operations to probability distributions. Such transformations are named as processes. In this work, a general supermap mapping N channels to a new channel can be regarded as a process involving $N+2$ parties. The first N parties have input/output pairs $(X_i, Y_i)_{i=1}^N$. The $(N+1)$-th party has an output system A and a trivial input system. The $(N+2)$-th party has an input system B and a trivial output system. According to Ref. [79], a linear operator J_{XYAB} is the Choi operator of such a general supermap if and only if

$$J_{XYAB} \geq 0, \quad \mathrm{Tr}\left[J_{XYAB}\right] = d_{YA}, \quad \mathcal{L}\left(J_{XYAB}\right) = J_{XYAB}, \tag{B5}$$

where $\mathcal{L}$ is a projection onto a linear subspace, defined by $\mathcal{L}\left(J_{XYAB}\right) := {}_{[1-(\prod_{j=1}^N(1-Y_j+X_jY_j))B+XYAB]}J_{XYAB}$. In the definition of $\mathcal{L}$ and the rest of the paper, we use the notation ${}_XJ := \mathrm{Tr}_X\left[J\right] \otimes I_X/d_X$, ${}_1J := J$, and ${}_{[\sum_j s_jX_j]}J := \sum_j s_j \cdot {}_{X_j}J$ for $s_j \in \{-1, 1\}$. We now show that this above characterization is equivalent to a simpler one that has a clear physical interpretation.

Proposition 4 (Characterization of general supermaps) *A linear operator J_{XYAB} is the Choi operator of some general supermap mapping N channels with input/output pairs $(X_i, Y_i)_{i=1}^N$ into a channel with input A and output B if and only if*

$$J_{XYAB} \geq 0 \quad \textit{and} \quad \mathcal{L}_{XY}^{\mathrm{NS}}\left(J_{XYA}\right) = \frac{I_{XYA}}{d_X}. \tag{B6}$$

Proof Consider that

$$\mathcal{L}\left(J_{XYAB}\right) = J_{XYAB} - {}_{[\prod_{j=1}^N(1-Y_j+X_jY_j)]}J_{XYA} \otimes \frac{I_B}{d_B} + \frac{\mathrm{Tr}\left[J_{XYAB}\right] I_{XYAB}}{d_{XYAB}} \tag{B7}$$

$$= J_{XYAB} - \left({}_{[\prod_{j=1}^N(1-Y_j+X_jY_j)]}J_{XYA} - \frac{I_{XYA}}{d_X}\right) \otimes \frac{I_B}{d_B}, \tag{B8}$$

where the second equality follows from $\mathrm{Tr}\left[J_{XYAB}\right] = d_{YA}$ as given in Eq. (B5). Then, to require $\mathcal{L}\left(J_{XYAB}\right) = J_{XYAB}$ is equivalent to requiring

$$\mathcal{L}^{\mathrm{NS}}_{XY}\left(J_{XYA}\right) \coloneqq {}_{\left[\prod_{j=1}^{N}(1-Y_j+X_jY_j)\right]}J_{XYA} = \frac{I_{XYA}}{d_X}. \tag{B9}$$

Note that the projector $\mathcal{L}^{\mathrm{NS}}_{XY}$ is trace-preserving. Therefore, the above constraint implies

$$\mathrm{Tr}\left[J_{XYAB}\right] = \mathrm{Tr}\left[\mathcal{L}^{\mathrm{NS}}_{XY}\left(J_{XYA}\right)\right] = d_{YA}, \tag{B10}$$

which completes the proof. ■

This characterization also applies to general processes involving N local parties as long as we set both A and B to trivial one-dimensional systems. Note that we use the superscript NS in $\mathcal{L}^{\mathrm{NS}}$ to emphasize the fact that $\mathcal{L}^{\mathrm{NS}}$ is actually the orthogonal projection onto the subspace of Hermitian operators spanned by the Choi operators of no-signaling multipartite quantum channels, as we show below.

Proposition 5 (Projection onto no-signaling subspace) $\mathcal{L}^{\mathrm{NS}}$ *is the orthogonal projection of Hermitian operators onto the subspace spanned by the Choi operators of no-signaling channels.*

Proof We call an N-partite channel no-signaling if the input state at any one local input system does not affect the output state at the rest $N-1$ local output systems. In terms of its Choi operator J_{XY}, it is equivalent to require

$$\mathrm{Tr}_{Y_j}\left[J_{XY}\right] = \mathrm{Tr}_{X_jY_j}\left[J_{XY}\right] \otimes \frac{I_{X_j}}{d_{X_j}} \quad \forall\, j = 1, \dots, N. \tag{B11}$$

These no-signaling constraints can be equivalently written as

$${}_{[Y_j - X_jY_j]}J_{XY} = 0 \quad \forall\, j = 1, \dots, N. \tag{B12}$$

Note that the no-signaling constraints implies ${}_{Y}J_{XY} = \mathrm{Tr}\left[J_{XY}\right] I_{XY}/d_{XY}$, which says that J_{XY} is proportional to the Choi operator of some trace-preserving map. Due to the linearity of this set of constraints, the subspace spanned by Choi operators of no-signaling channels consists of all Hermitian operators satisfying the same set of constraints. For each j, the constraint ${}_{[Y_j - X_jY_j]}J_{XY} = 0$ is satisfied if and only if

$$\mathcal{L}^{\mathrm{NS}}_{X_jY_j}\left(J_{XY}\right) \coloneqq {}_{[1-Y_j+X_jY_j]}J_{XY} = J_{XY}. \tag{B13}$$

Indeed, by checking that $\mathcal{L}^{\mathrm{NS}}_{X_jY_j} \circ \mathcal{L}^{\mathrm{NS}}_{X_jY_j} = L^{\mathrm{NS}}_{X_jY_j}$ and it is self-adjoint, one can verify that $\mathcal{L}^{\mathrm{NS}}_{X_jY_j}$ is the orthogonal projection onto the subspace of Hermitian operators satisfying ${}_{[Y_j - X_jY_j]}J_{X_jY_j} = 0$. The intersection of subspaces associated with all $\mathcal{L}^{\mathrm{NS}}_{X_jY_j}$ is exactly the no-signaling subspace. Since all the projections $\mathcal{L}^{\mathrm{NS}}_{X_jY_j}$ commute as they all act on different systems, the orthogonal projection onto the no-signaling subspace is given by

$$\mathcal{L}^{\mathrm{NS}}_{X_1Y_1} \circ \cdots \circ \mathcal{L}^{\mathrm{NS}}_{X_NY_N} = \mathcal{L}^{\mathrm{NS}}_{XY}, \tag{B14}$$

which completes the proof. ■

Supplementary Note C: One-shot classical communication capacity

1. Equivalence between minimum error and minimum average probability of error

Proposition 6 *Let* $\mathbf{X}$ *be a convex set of general supermaps mapping* N *channels with input/output pairs* $(X_i, Y_i)_{i=1}^{N}$ *into a channel with input* A *and output* B*, where both* A *and* B *are* m*-dimensional systems. Given a list* $\vec{\mathcal{C}}$ *of* N *channels, it holds that*

$$\omega^{\mathbf{X}}(\mathcal{C}, \Delta^m) = p^{\mathbf{X}}\left(\mathcal{C}, m\right) \tag{C1}$$

if $\mathbf{X}$ *is closed under simultaneous permutations on* A *and* B*, i.e., for every supermap* $\mathscr{S} \in \mathbf{X}$ *and for every permutation* τ *in the symmetric group* S_m *of degree* m*, the supermap* $\mathscr{W}^{\tau} \circ \mathscr{S} \in \mathbf{X}$*, where* $\mathscr{W}^{\tau}$ *acts as* $\mathscr{W}^{\tau}(\cdot) \coloneqq \mathcal{W}^{\tau}_{B\to B} \circ (\cdot) \circ \mathcal{W}^{\tau^{\dagger}}_{A\to A}$ *with* $\mathcal{W}^{\tau}$ *being the permutation channel associated with* τ*.*

Proof For $m=1$, it is clear that $\omega^{\mathbf{X}}(\mathcal{C},\Delta^1)=p^{\mathbf{X}}(\mathcal{C},1)=0$ because there is only one possible state for a one-dimensional quantum system, which makes the communication task trivial. Hence, Eq. (C1) does hold for $m=1$.

Now, we consider cases where $m\geq 2$. For a list of N quantum channels $\vec{\mathcal{C}}=(\mathcal{C}_1,\ldots,\mathcal{C}_N)$, we can express the minimum average error probability $p^{\mathbf{X}}(\mathcal{C},m)$ in terms of its Choi operator $J^{\vec{\mathcal{C}}}_{XY}:=\bigotimes_{i=1}^N J^{\mathcal{C}_i}_{X_iY_i}$ ($J^{\mathcal{C}_i}_{X_iY_i}$ is the Choi operator of $\mathcal{C}_i$) as

$$\begin{aligned}p^{\mathbf{X}}\left(\vec{\mathcal{C}},m\right)&=\min 1-\frac{1}{m}\sum_{j=1}^m \operatorname{Tr}\left[J^{\mathcal{M}}_{AB}\left(|j\rangle\langle j|_A\otimes|j\rangle\langle j|_B\right)\right]\\&\quad\text{s.t. } J^{\mathcal{M}}_{AB}=J^{\mathcal{S}}_{XYAB}\star J^{\vec{\mathcal{C}}}_{XY},\ J^{\mathcal{S}}_{XYAB}\in\widetilde{\mathbf{X}}\end{aligned}\tag{C2}$$

$$\begin{aligned}&=\min 1-\frac{1}{m}\operatorname{Tr}\left[\left(J^{\mathcal{S}}_{XYAB}\star J^{\vec{\mathcal{C}}}_{XY}\right)J^{\Delta^m}_{AB}\right]\\&\quad\text{s.t. } J^{\mathcal{S}}_{XYAB}\in\widetilde{\mathbf{X}},\end{aligned}\tag{C3}$$

where $J^{\Delta^m}_{AB}:=\sum_{j=1}^m|j\rangle\langle j|_A\otimes|j\rangle\langle j|_B$ is the Choi operator of an m-dimensional classical noiseless channel Δ^m, and $\widetilde{\mathbf{X}}$ denotes the set of all Choi operators corresponding to supermaps in the given set $\mathbf{X}$.

The optimization program in Eq. (C3) can be further simplified by exploiting the symmetry of $J^{\Delta^m}_{AB}$. Specifically, if $J^{\mathcal{S}}_{XYAB}$ is optimal with the corresponding supermap $\mathcal{S}\in\mathbf{X}$, then the corresponding Choi operator $J^{\mathcal{S}^\tau}_{XYAB}$ of $\mathcal{S}^\tau:=\mathscr{W}^\tau\circ\mathcal{S}$ is feasible for any permutation τ in the symmetric group S_m of degree m because $\mathcal{S}^\tau\in\mathbf{X}$. Given that

$$J^{\mathscr{W}^\tau}_{ABA'B'}=\left(\sum_{j,k=1}^m|j\rangle\langle k|_{A'}\otimes|\tau^\dagger(j)\rangle\langle\tau^\dagger(k)|_A\right)\otimes\left(\sum_{j',k'=1}^m|j'\rangle\langle k'|_B\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\right)\tag{C4}$$

$$=\left(\sum_{j,k=1}^m|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|j\rangle\langle k|_A\right)\otimes\left(\sum_{j',k'=1}^m|j'\rangle\langle k'|_B\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\right)\tag{C5}$$

$$=\sum_{j,k,j',k'=1}^m|j\rangle\langle k|_A\otimes|j'\rangle\langle k'|_B\otimes|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|\tau(j')\rangle\langle\tau(k')|_{B'},\tag{C6}$$

where we mark the output system as $A'B'$ for clarity, the Choi operator $J^{\mathcal{S}^\tau}_{XYAB}$ can be written as

$$J^{\mathcal{S}^\tau}_{XYAB}=J^{\mathcal{S}}_{XYAB}\star J^{\mathscr{W}^\tau}_{ABA'B'}\tag{C7}$$

$$=\operatorname{Tr}_{AB}\left[\left(J^{\mathcal{S}}_{XYAB}\otimes I_{A'B'}\right)\left(I_{XY}\otimes\sum_{j,k,j',k'=1}^m|k\rangle\langle j|_A\otimes|k'\rangle\langle j'|_B\otimes|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\right)\right]\tag{C8}$$

$$=\sum_{j,k,j',k'=1}^m\operatorname{Tr}_{AB}\left[\left(J^{\mathcal{S}}_{XYAB}\otimes I_{A'B'}\right)\left(I_{XY}\otimes|k\rangle\langle j|_A\otimes|k'\rangle\langle j'|_B\otimes|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\right)\right]\tag{C9}$$

$$=\sum_{j,k,j',k'=1}^m\operatorname{Tr}_{AB}\left[J^{\mathcal{S}}_{XYAB}\left(I_{XY}\otimes|k\rangle\langle j|_A\otimes|k'\rangle\langle j'|_B\right)\right]\otimes|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\tag{C10}$$

$$=\sum_{j,k,j',k'=1}^m\left(I_{XY}\otimes\langle j|_A\otimes\langle j'|_B\right)J^{\mathcal{S}}_{XYAB}\left(I_{XY}\otimes|k\rangle_A\otimes|k'\rangle_B\right)\otimes|\tau(j)\rangle\langle\tau(k)|_{A'}\otimes|\tau(j')\rangle\langle\tau(k')|_{B'}\tag{C11}$$

$$=\sum_{j,k,j',k'=1}^m\left(I_A\otimes|\tau(j)\rangle_{A'}\langle j|_A\otimes|\tau(j')\rangle_{B'}\langle j'|_B\right)J^{\mathcal{S}}_{XYAB}\left(I_A\otimes|k\rangle_A\langle\tau(k)|_{A'}\otimes|k'\rangle_B\langle\tau(k')|_{B'}\right)\tag{C12}$$

$$\begin{aligned}&=\left(I_{XY}\otimes\sum_{j=0}^{m-1}|\tau(j)\rangle_{A'}\langle j|_A\otimes\sum_{j'=0}^{m-1}|\tau(j')\rangle_{B'}\langle j'|_B\right)J^{\mathcal{S}}_{XYAB}\\&\qquad\left(I_{XY}\otimes\sum_{k=0}^{m-1}|k\rangle_A\langle\tau(k)|_{A'}\otimes\sum_{k'=0}^{m-1}|k'\rangle_B\langle\tau(k')|_{B'}\right)\end{aligned}\tag{C13}$$

$$=\left(I_{XY}\otimes\tau_A\otimes\tau_B\right)J^{\mathcal{S}}_{XYAB}\left(I_{XY}\otimes\tau_A\otimes\tau_B\right)^\dagger.\tag{C14}$$

More importantly, $J^{\mathscr{S}^\tau}_{XYAB}$ achieves the same average probability of error as $J^{\mathscr{S}}_{XYAB}$ since

$$\begin{aligned}\operatorname{Tr}\left[\left(J^{\mathscr{S}^\tau}_{XYAB} \star J^{\vec{\mathcal{C}}}_{XY}\right) J^{\Delta^m}_{AB}\right] &= \operatorname{Tr}\left[\left(\left(I_{XY} \otimes \tau_A \otimes \tau_B\right) J^{\mathscr{S}}_{XYAB} \left(I_{XY} \otimes \tau_A \otimes \tau_B\right)^\dagger \star J^{\vec{\mathcal{C}}}_{XY}\right) J^{\Delta^m}_{AB}\right] &\text{(C15)}\\ &= \operatorname{Tr}\left[\left(J^{\mathscr{S}}_{XYAB} \star J^{\vec{\mathcal{C}}}_{XY}\right) \left(\tau_A \otimes \tau_B\right)^\dagger J^{\Delta^m}_{AB} \left(\tau_A \otimes \tau_B\right)\right] &\text{(C16)}\\ &= \operatorname{Tr}\left[\left(J^{\mathscr{S}}_{XYAB} \star J^{\vec{\mathcal{C}}}_{XY}\right) J^{\Delta^m}_{AB}\right], &\text{(C17)}\end{aligned}$$

where the last equality holds due to $\left(\tau_A \otimes \tau_B\right)^\dagger J^{\Delta^m}_{AB} \left(\tau_A \otimes \tau_B\right) = J^{\Delta^m}_{AB}$. Thus, the optimality of $J^{\mathscr{S}}_{XYAB}$ implies the optimality of $J^{\mathscr{S}^\tau}_{XYAB}$. Moreover, any convex combination of optimal Choi operators is still optimal due to the convexity of $\mathbf{X}$ and the linearity of the average error probability, implying that

$$J^{\mathscr{S}'}_{XYAB} \coloneqq \frac{1}{m!} \sum_{\tau \in S_m} J^{\mathscr{S}^\tau}_{XYAB} \tag{C18}$$

is also optimal. Then, by Schur's Lemma, we can restrict the optimization range of the operator $J^{\mathscr{S}}_{XYAB}$ to operators of the form $E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right)$, where E_{XY} and F_{XY} are arbitrary Hermitian operators. For that

$$\left(E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right)\right) \star J^{\vec{\mathcal{C}}}_{XY} = \left(E_{XY} \star J^{\vec{\mathcal{C}}}_{XY}\right) J^{\Delta^m}_{AB} + \left(F_{XY} \star J^{\vec{\mathcal{C}}}_{XY}\right) \left(I_{AB} - J^{\Delta^m}_{AB}\right), \tag{C19}$$

the optimization program can be written as

$$\begin{aligned} p^{\mathbf{X}}\left(\vec{\mathcal{C}}, m\right) &= \min\, 1 - E_{XY} \star J^{\vec{\mathcal{C}}}_{XY} \\ &\quad \text{s.t. } E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right) \in \widetilde{\mathbf{X}}. \end{aligned} \tag{C20}$$

For the minimum error $\omega^{\mathbf{X}}(\vec{\mathcal{C}}, \Delta^m)$, we can use the semidefinite program (SDP) [82] for the diamond distance between quantum channels to write it as the following optimization program:

$$\begin{aligned} \omega^{\mathbf{X}}(\vec{\mathcal{C}}, \Delta^m) &= \min\, \epsilon \\ &\quad \text{s.t. } Y_{AB} \geq J^{\mathscr{S}}_{XYAB} \star J^{\vec{\mathcal{C}}}_{XY} - J^{\Delta^m}_{AB},\ Y_{AB} \geq 0,\ Y_A \leq \epsilon I_A,\ J^{\mathscr{S}}_{XYAB} \in \widetilde{\mathbf{X}}. \end{aligned} \tag{C21}$$

Similarly, we can restrict the optimization range of the operator $J^{\mathscr{S}}_{XYAB}$ to operators of the form $E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right)$ so that

$$\begin{aligned} \omega^{\mathbf{X}}(\vec{\mathcal{C}}, \Delta^m) &= \min\, \epsilon \\ &\quad \text{s.t. } Y_{AB} \geq \left(E_{XY} \star J^{\vec{\mathcal{C}}}_{XY} - 1\right) J^{\Delta^m}_{AB} + \left(F_{XY} \star J^{\vec{\mathcal{C}}}_{XY}\right) \left(I_{AB} - J^{\Delta^m}_{AB}\right),\ Y_{AB} \geq 0,\ Y_A \leq \epsilon I_A, &\text{(C22)} \\ &\quad\quad E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right) \in \widetilde{\mathbf{X}} \\ &= \min\, \frac{1}{2} \left\|\mathcal{M} - \Delta^m\right\|_\diamond \\ &\quad \text{s.t. } E_{XY} \otimes J^{\Delta^m}_{AB} + F_{XY} \otimes \left(I_{AB} - J^{\Delta^m}_{AB}\right) \in \widetilde{\mathbf{X}}, &\text{(C23)} \end{aligned}$$

where $\mathcal{M}$ denotes the quantum channel whose Choi operator is

$$J^{\mathcal{M}}_{AB} \coloneqq \left(E_{XY} \star J^{\vec{\mathcal{C}}}_{XY}\right) J^{\Delta^m}_{AB} + \left(F_{XY} \star J^{\vec{\mathcal{C}}}_{XY}\right) \left(I_{AB} - J^{\Delta^m}_{AB}\right). \tag{C24}$$

One constraint that holds for the Choi operator $J^{\mathscr{S}}_{XYAB}$ of every general supermap $\mathscr{S}$ is $\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathscr{S}}_{XYA}\right) = \frac{I_{XYA}}{d_X}$ (see Proposition 4). In terms of operators E_{XY} and F_{XY}, this constraint can be expressed as

$$\mathcal{L}^{\mathrm{NS}}_{XY}\left(E_{XY} + (m-1)F_{XY}\right) = \frac{I_{XY}}{d_X}, \tag{C25}$$

which implies that

$$F_{XY} \star J_{XY}^{\vec{\mathcal{C}}} = F_{XY} \star \mathcal{L}_{XY}^{\mathrm{NS}} \left(J_{XY}^{\vec{\mathcal{C}}} \right) \tag{C26}$$

$$= \mathcal{L}_{XY}^{\mathrm{NS}} \left(F_{XY} \right) \star J_{XY}^{\vec{\mathcal{C}}} \tag{C27}$$

$$= \frac{1}{m-1} \left(\frac{I_{XY}}{d_X} - \mathcal{L}_{XY}^{\mathrm{NS}} \left(E_{XY} \right) \right) \star J_{XY}^{\vec{\mathcal{C}}} \tag{C28}$$

$$= \frac{1}{m-1} \left(1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \right), \tag{C29}$$

where the first equality holds because $\vec{\mathcal{C}}$ is no-signaling and the second equality is due to that $\mathcal{L}_{XY}^{\mathrm{NS}}$ is self-dual. It then follows that

$$J_{AB}^{\mathcal{M}} - J_{AB}^{\Delta^m} = \left(E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} - 1 \right) J_{AB}^{\Delta^m} + \frac{1}{m-1} \left(1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \right) \left(I_{AB} - J_{AB}^{\Delta^m} \right) \tag{C30}$$

$$= \frac{1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}}}{m-1} \left(I_{AB} - m J_{AB}^{\Delta^m} \right). \tag{C31}$$

According to the definition of the diamond norm, we have

$$\left\| \mathcal{M} - \Delta^m \right\|_{\diamond} = \max_{\rho_{RA}} \left\| \mathcal{M} \left(\rho_{RA} \right) - \Delta^m \left(\rho_{RA} \right) \right\|_1, \tag{C32}$$

where R is an m-dimensional quantum system. Denote $1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}}$ by α. For any quantum state ρ_{RA}, consider that

$$\mathcal{M} \left(\rho_{RA} \right) - \Delta^m \left(\rho_{RA} \right) = \frac{\alpha}{m-1} \left(I_{AB} \star \rho_{RA} - m J_{AB}^{\Delta^m} \star \rho_{RA} \right) \tag{C33}$$

$$= \frac{\alpha}{m-1} \left(\rho_R \otimes I_B - m \operatorname{Tr}_A \left[\left(\sum_{j=1}^{m} I_R \otimes |j\rangle\langle j|_A \otimes |j\rangle\langle j|_B \right) \left(\rho_{RA} \otimes I_B \right) \right] \right) \tag{C34}$$

$$= \frac{\alpha}{m-1} \left(\sum_{j=1}^{m} \rho_R \otimes |j\rangle\langle j|_B - m \sum_{j=1}^{m} \operatorname{Tr}_A \left[\left(I_R \otimes |j\rangle\langle j|_A \right) \rho_{RA} \right] \otimes |j\rangle\langle j|_B \right) \tag{C35}$$

$$= \frac{\alpha}{m-1} \sum_{j=1}^{m} \left(\rho_R - m \operatorname{Tr}_A \left[\left(I_R \otimes |j\rangle\langle j|_A \right) \rho_{RA} \right] \right) \otimes |j\rangle\langle j|_B \tag{C36}$$

$$= \frac{\alpha}{m-1} \sum_{j=1}^{m} \left(\rho_R - m \widetilde{\rho}_R^j \right) \otimes |j\rangle\langle j|_B, \tag{C37}$$

where we define $\widetilde{\rho}_R^j := \operatorname{Tr}_A \left[\left(I_R \otimes |j\rangle\langle j|_A \right) \rho_{RA} \right]$, which is positive semidefinite for every j. Since $\rho_R = \operatorname{Tr}_A \left[\left(I_R \otimes I_A \right) \rho_{RA} \right] = \sum_{j=1}^{m} \operatorname{Tr}_A \left[\left(I_R \otimes |j\rangle\langle j|_A \right) \rho_{RA} \right] = \sum_{j=1}^{m} \widetilde{\rho}_R^j$, we have

$$\mathcal{M} \left(\rho_{RA} \right) - \Delta^m \left(\rho_{RA} \right) = \frac{\alpha}{m-1} \sum_{j=1}^{m} \left(\rho_R - \widetilde{\rho}_R^j - (m-1) \widetilde{\rho}_R^j \right) \otimes |j\rangle\langle j|_B, \tag{C38}$$

where $\rho_R - \widetilde{\rho}_R^j = \sum_{k=1,\, k \neq j}^{m} \widetilde{\rho}_R^k$ is positive semidefinite. Then,

$$\left\| \mathcal{M} \left(\rho_{RA} \right) - \Delta^m \left(\rho_{RA} \right) \right\|_1 = \left\| \frac{\alpha}{m-1} \sum_{j=1}^{m} \left(\rho_R - \widetilde{\rho}_R^j - (m-1) \widetilde{\rho}_R^j \right) \otimes |j\rangle\langle j|_B \right\|_1 \tag{C39}$$

$$= \frac{\alpha}{m-1} \sum_{j=1}^{m} \left\| \rho_R - \widetilde{\rho}_R^j - (m-1) \widetilde{\rho}_R^j \right\|_1 \tag{C40}$$

$$
\begin{aligned}
&\leq \frac{\alpha}{m-1}\sum_{j=1}^{m}\left(\left\|\rho_R-\widetilde{\rho}_R^j\right\|_1+(m-1)\left\|\widetilde{\rho}_R^j\right\|_1\right) && \text{(C41)}\\
&=\frac{\alpha}{m-1}\sum_{j=1}^{m}\left(1+(m-2)\operatorname{Tr}\left[\widetilde{\rho}_R^j\right]\right) && \text{(C42)}\\
&=\frac{\alpha}{m-1}\left(m+(m-2)\operatorname{Tr}\left[\rho_R\right]\right) && \text{(C43)}\\
&=2\alpha. && \text{(C44)}
\end{aligned}
$$

Since this holds for any state ρ_{RA}, we obtain an upper bound on the diamond distance:

$$
\frac{1}{2}\left\|\mathcal{M}-\Delta^m\right\|_\diamond=\frac{1}{2}\max_{\rho_{RA}}\left\|\mathcal{M}\left(\rho_{RA}\right)-\Delta^m\left(\rho_{RA}\right)\right\|_1\leq\alpha. \tag{C45}
$$

On the other hand, $\frac{1}{2}\left\|\mathcal{M}-\Delta^m\right\|_\diamond$ is lower bounded by α because, for $\rho_{RA}=|0\rangle\langle 0|_R\otimes|0\rangle\langle 0|_A$, $\rho_R=\widetilde{\rho}_R^0=|0\rangle\langle 0|_R$ and $\widetilde{\rho}_R^j=0$ for all $j\neq 0$. As a result, Eq. (C40) implies that

$$
\begin{aligned}
\left\|\mathcal{M}\left(\rho_{RA}\right)-\Delta^m\left(\rho_{RA}\right)\right\|_1&=\frac{\alpha}{m-1}\left(\left\|-(m-1)\widetilde{\rho}_R^0\right\|_1+(m-1)\left\|\rho_R\right\|_1\right) && \text{(C46)}\\
&=2\alpha. && \text{(C47)}
\end{aligned}
$$

Since α is both an upper bound and a lower bound on $\frac{1}{2}\left\|\mathcal{M}-\Delta^m\right\|_\diamond$, we obtain

$$
\frac{1}{2}\left\|\mathcal{M}-\Delta^m\right\|_\diamond=\alpha=1-E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}. \tag{C48}
$$

Therefore, we can write the optimization program of the minimum error as

$$
\begin{aligned}
\omega^{\mathbf{X}}(\vec{\mathcal{C}},\Delta^m)=\min\ &1-E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}\\
\text{s.t. }&E_{XY}\otimes J_{AB}^{\Delta^m}+F_{XY}\otimes\left(I_{AB}-J_{AB}^{\Delta^m}\right)\in\widetilde{\mathbf{X}}
\end{aligned} \tag{C49}
$$

$$
=p^{\mathbf{X}}\left(\vec{\mathcal{C}},m\right), \tag{C50}
$$

as we set out to prove. ■

2. SDP for one-shot classical capacities assisted by free general supermaps

Proposition 7 *Given a list of N quantum channels $\vec{\mathcal{C}}$, its one-shot classical capacity assisted by free general supermaps with error tolerance $\epsilon\geq 0$ can be formulated as*

$$
\begin{aligned}
C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{C}})=\log_2\max\ &\lfloor m\rfloor\\
\text{s.t. }&E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}\geq m(1-\epsilon),\ E_Y=I_Y,\\
&0\leq E_{XY}\leq F_{XY},\ \mathcal{L}_{XY}^{\mathrm{NS}}\left(F_{XY}\right)=\frac{m}{d_X}I_{XY},
\end{aligned} \tag{C51}
$$

where m is maximized over real numbers, and E_{XY} and F_{XY} are optimization variables.

Proof Recall that

$$
\begin{aligned}
C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}})\coloneqq\log_2\max\ &m\\
\text{s.t. }&\omega^{\mathbf{X}}(\vec{\mathcal{C}},\Delta^m)\leq\epsilon,
\end{aligned} \tag{C52}
$$

where m is maximized over all positive integers. According to Eq. (C49), we can write $C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}})$ as

$$\begin{aligned} C_\epsilon^{\mathbf{X}}(\vec{\mathcal{C}}) = \log_2 \max\, & m \\ \text{s.t. } & 1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \leq \epsilon, \\ & E_{XY} \otimes J_{AB}^{\Delta^m} + F_{XY} \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right) \in \widetilde{\mathbf{X}}. \end{aligned} \tag{C53}$$

Choi operators of signaling-non-generating supermaps should satisfy constraint (B1), which can be compactly written as

$$ {}_{[1-A]}J_{YAB} = 0 \tag{C54}$$

for the Choi operator J_{XYAB} of a supermap. Applying this constraint to $E_{XY} \otimes J_{AB}^{\Delta^m} + F_{XY} \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right)$ results in

$$\begin{aligned} {}_{[1-A]}\left(E_Y \otimes J_{AB}^{\Delta^m} + F_Y \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right)\right) &= E_Y \otimes {}_{[1-A]}J_{AB}^{\Delta^m} + F_Y \otimes {}_{[1-A]}\left(I_{AB} - J_{AB}^{\Delta^m}\right) && \text{(C55)} \\ &= (E_Y - F_Y) \otimes \left(J_{AB}^{\Delta^m} - \frac{1}{m} I_{AB}\right) && \text{(C56)} \\ &= 0, && \text{(C57)} \end{aligned}$$

which is true if and only if $E_Y = F_Y$.

The case where $\mathbf{X} = \mathbf{Free}$ corresponds to the constraints given in Proposition 4, in addition to the forward no-signaling constraint (C54). The first constraint in Proposition 4 is requiring both $E_{XY} \geq 0$ and $F_{XY} \geq 0$. The second constraint gives us

$$\mathcal{L}_{XY}^{\mathrm{NS}}\left(E_{XY} + (m-1)F_{XY}\right) = \frac{I_{XY}}{d_X}, \tag{C58}$$

which, as we show for Eq. (F15), implies

$$\mathrm{Tr}_X\left[\mathcal{L}_{XY}^{\mathrm{NS}}\left(E_{XY} + (m-1)F_{XY}\right)\right] = E_Y + (m-1)F_Y = I_Y. \tag{C59}$$

Under the above equality, the forward no-signaling constraint $E_Y = F_Y$ is equivalent to $mE_Y = I_Y$. Thus, we obtain

$$\begin{aligned} C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \max\, & m \\ \text{s.t. } & 1 - E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \leq \epsilon,\ mE_Y = I_Y, \\ & E_{XY} \geq 0,\ F_{XY} \geq 0,\ \mathcal{L}_{XY}^{\mathrm{NS}}\left(E_{XY} + (m-1)F_{XY}\right) = \frac{I_{XY}}{d_X}. \end{aligned} \tag{C60}$$

Relabeling mE_{XY} as E_{XY} and $mE_{XY} + m(m-1)F_{XY}$ as F_{XY} leads to

$$\begin{aligned} C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \max\, & m \\ \text{s.t. } & E_{XY} \star J_{XY}^{\vec{\mathcal{C}}} \geq m(1-\epsilon),\ E_Y = I_Y, \\ & 0 \leq E_{XY} \leq F_{XY},\ \mathcal{L}_{XY}^{\mathrm{NS}}\left(F_{XY}\right) = \frac{m}{d_X} I_{XY}. \end{aligned} \tag{C61}$$

While m is maximized over integers, it can actually be maximized over real numbers and then take the largest integer smaller than it. To see this, suppose E_{XY}^*, F_{XY}^*, m^* are optimal when m is maximized over real numbers, and m^* is not an integer. Then, denoting $p := \frac{\lfloor m^* \rfloor - 1}{m^* - 1}$, one can verify that $pE_{XY}^* + \frac{1-p}{d_X} I_{XY}$, $pF_{XY}^* + \frac{1-p}{d_X} I_{XY}$, and $\lfloor m^* \rfloor$ are feasible. This allows us to formulate $C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{C}})$ as an SDP. ■

3. SDP for one-shot classical capacities assisted by free causally definite supermaps

The one-shot classical capacity assisted by free causally definite supermaps can be formulated as an SDP following a similar procedure. Here, we consider the bipartite case for demonstration.

Proposition 8 *Given a list of two quantum channels $\vec{\mathcal{C}}$, its one-shot classical capacity assisted by free causally definite su-*

permaps with error tolerance $\epsilon \geq 0$ can be formulated as

$$\begin{aligned} C_\epsilon^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = \log_2 \max\ & \lfloor m \rfloor \\ \text{s.t. } & (E_{XY} + G_{XY}) \star J_{XY}^{\vec{\mathcal{C}}} \geq m(1-\epsilon),\ E_Y + G_Y = I_Y, \\ & 0 \leq E_{XY} \leq F_{XY},\ 0 \leq G_{XY} \leq H_{XY},\ \mathrm{Tr}\left[F_{XY} + H_{XY}\right] = m d_Y, \\ & {}_{[1-Y_2]}F_{XY} = 0,\ {}_{[1-Y_1]}F_{X_1Y_1} = 0,\ {}_{[1-Y_1]}H_{XY} = 0,\ {}_{[1-Y_2]}H_{X_2Y_2} = 0, \end{aligned} \tag{C62}$$

where m is maximized over real numbers, and E_{XY}, F_{XY}, G_{XY}, and H_{XY} are optimization variables.

Proof In the bipartite case, the constraints given in Proposition 3 on the Choi operator J_{XYAB} of a causally definite supermap reduce to

$$\begin{gathered} J_{XYAB} = J_{XYAB}^{(1,2)} + J_{XYAB}^{(2,1)}, \quad J_A^{(1,2)} + J_A^{(2,1)} = d_Y I_A, \\ J_{XYAB}^{(1,2)} \geq 0, \quad J_{XYAB}^{(2,1)} \geq 0, \\ {}_{[1-Y_1]}J_{X_1Y_1A}^{(1,2)} = 0, \quad {}_{[1-Y_2]}J_{X_2Y_2A}^{(2,1)} = 0, \\ {}_{[1-Y_2]}J_{XYA}^{(1,2)} = 0, \quad {}_{[1-Y_1]}J_{XYA}^{(2,1)} = 0. \end{gathered} \tag{C63}$$

Since simultaneous permutation at A and B on J_{XYAB} does not change the optimal value when solving the optimization program of the minimum error, we can restrict $J_{XYAB}^{(1,2)}$ and $J_{XYAB}^{(2,1)}$ to operators of the form

$$J_{XYAB}^{(1,2)} = E_{XY} \otimes J_{AB}^{\Delta^m} + F_{XY} \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right) \quad \text{and} \quad J_{XYAB}^{(2,1)} = G_{XY} \otimes J_{AB}^{\Delta^m} + H_{XY} \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right), \tag{C64}$$

where E_{XY}, F_{XY}, G_{XY}, and H_{XY} are arbitrary Hermitian operators. Then, we can rewrite constraints (C63) in terms of E_{XY}, F_{XY}, G_{XY}, and H_{XY}.

The constraint $J_A^{(1,2)} + J_A^{(2,1)} = d_Y I_A$ is equivalent to

$$\mathrm{Tr}\left[E_{XY} + (m-1)F_{XY} + G_{XY} + (m-1)H_{XY}\right] = d_Y. \tag{C65}$$

The constraints $J_{XYAB}^{(1,2)} \geq 0$ and $J_{XYAB}^{(2,1)} \geq 0$ correspond to

$$E_{XY} \geq 0, \quad F_{XY} \geq 0, \quad G_{XY} \geq 0, \quad \text{and} \quad H_{XY} \geq 0. \tag{C66}$$

Since $J_{XYA}^{(1,2)} = (E_{XY} + (m-1)F_{XY}) \otimes I_A$, the constraints ${}_{[1-Y_2]}J_{XYA}^{(1,2)} = 0$ and ${}_{[1-Y_1]}J_{X_1Y_1A}^{(1,2)} = 0$ correspond to

$${}_{[1-Y_2]}\left(E_{XY} + (m-1)F_{XY}\right) = 0 \quad \text{and} \quad {}_{[1-Y_1]}\left(E_{X_1Y_1} + (m-1)F_{X_1Y_1}\right) = 0. \tag{C67}$$

Similarly, ${}_{[1-Y_1]}J_{XYA}^{(2,1)} = 0$ and ${}_{[1-Y_2]}J_{X_2Y_2A}^{(2,1)} = 0$ correspond to

$${}_{[1-Y_1]}\left(G_{XY} + (m-1)H_{XY}\right) = 0 \quad \text{and} \quad {}_{[1-Y_2]}\left(G_{X_2Y_2} + (m-1)H_{X_2Y_2}\right) = 0. \tag{C68}$$

Finally, comparing $J_{XYAB} = (E_{XY} + G_{XY}) \otimes J_{AB}^{\Delta^m} + F_{XY} + H_{XY} \otimes \left(I_{AB} - J_{AB}^{\Delta^m}\right)$ with Eq. (C53), we note that E_{XY} and F_{XY} in Eq. (C53) correspond to $E_{XY} + G_{XY}$ and $F_{XY} + H_{XY}$, respectively, in this case. Hence, we arrive at

$$\begin{aligned} C_\epsilon^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = \log_2 \max\ & m \\ \text{s.t. } & (E_{XY} + G_{XY}) \star J_{XY}^{\vec{\mathcal{C}}} \geq 1-\epsilon,\ E_Y + G_Y = F_Y + H_Y, \\ & \mathrm{Tr}\left[E_{XY} + (m-1)F_{XY} + G_{XY} + (m-1)H_{XY}\right] = d_Y, \\ & E_{XY} \geq 0,\ F_{XY} \geq 0,\ G_{XY} \geq 0,\ H_{XY} \geq 0, \\ & {}_{[1-Y_2]}\left(E_{XY} + (m-1)F_{XY}\right) = 0,\ {}_{[1-Y_1]}\left(E_{X_1Y_1} + (m-1)F_{X_1Y_1}\right) = 0, \\ & {}_{[1-Y_1]}\left(G_{XY} + (m-1)H_{XY}\right) = 0,\ {}_{[1-Y_2]}\left(G_{X_2Y_2} + (m-1)H_{X_2Y_2}\right) = 0. \end{aligned} \tag{C69}$$

Since J_{XYAB} also satisfy the causality constraint $\mathcal{L}_{XY}^{\mathrm{NS}}\left(J_{XYA}\right) = \frac{I_{XYA}}{d_X}$ for a general supermap, we can equivalently write the forward no-signaling constraint $E_Y + G_Y = F_Y + H_Y$ as

$$m\left(E_Y + G_Y\right) = I_Y, \tag{C70}$$

as we did in the case with general supermaps. Relabeling mE_{XY} as E_{XY}, mG_{XY} as G_{XY}, $mE_{XY}+m(m-1)F_{XY}$ as F_{XY}, and $mG_{XY}+m(m-1)H_{XY}$ as H_{XY}, we obtain

$$\begin{aligned}C_\epsilon^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = \log_2 \max\ & m \\ \text{s.t. } & (E_{XY}+G_{XY})\star J_{XY}^{\vec{\mathcal{C}}} \geq m(1-\epsilon),\ E_Y+G_Y=I_Y, \\ & 0\leq E_{XY}\leq F_{XY},\ 0\leq G_{XY}\leq H_{XY},\ \mathrm{Tr}\left[F_{XY}+H_{XY}\right]=md_Y, \\ & {}_{[1-Y_2]}F_{XY}=0,\ {}_{[1-Y_1]}F_{X_1Y_1}=0,\ {}_{[1-Y_1]}H_{XY}=0,\ {}_{[1-Y_2]}H_{X_2Y_2}=0.\end{aligned} \tag{C71}$$

By the same argument as in the case with general supermaps, we can relax m to be maximized over real numbers and then take the largest integer smaller than it, which completes the proof. ■

4. SDP for one-shot classical capacities assisted by free parallel supermaps

The one-shot classical capacity assisted by free parallel supermaps is the same as that assisted by no-signaling correlations in the standard communication scenario, which can be formulated as the following SDP [80, 81]:

$$\begin{aligned}C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{C}}) = \log_2 \max\ & \lfloor m\rfloor \\ \text{s.t. } & E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}\geq m(1-\epsilon),\ E_Y=I_Y, \\ & 0\leq E_{XY}\leq F_X\otimes I_Y,\ \mathrm{Tr}\left[F_X\right]=m,\end{aligned} \tag{C72}$$

where m is maximized over real numbers, and E_{XY} and F_X are optimization variables.

Supplementary Note D: Advantage of indefinite causal order in zero-error communication

Corollary 1 (SDPs for one-shot zero-error classical capacity assisted by free general supermaps) *Given a list of N quantum channels $\vec{\mathcal{C}}$, its one-shot zero-error classical capacity assisted by free general supermaps can be formulated as*

$$\begin{aligned}C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \max\ & \lfloor m\rfloor \\ \text{s.t. } & E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}=m,\ 0\leq E_{XY}\leq F_{XY},\ \mathcal{L}_{XY}^{\mathrm{NS}}(F_{XY})=\frac{m}{d_X}I_{XY},\ E_Y=I_Y\end{aligned} \tag{D1}$$

$$\begin{aligned}\leq \log_2 \min\ & \lfloor \mathrm{Tr}\left[N_Y\right]\rfloor \\ \text{s.t. } & \mathcal{L}_{XY}^{\mathrm{NS}}\left(M_{XY}\right)\geq 0,\ \mathcal{L}_{XY}^{\mathrm{NS}}\left(M_{XY}\right)+I_X\otimes N_Y\geq \left(1+\frac{\mathrm{Tr}\left[M_{XY}\right]}{d_X}\right)J_{XY}^{\vec{\mathcal{C}}}\end{aligned} \tag{D2}$$

with primal variables m, E_{XY}, F_{XY} and dual variables M_{XY}, N_Y.

Proof Since zero-error is a special case where $\epsilon=0$, we can obtain an SDP of $C_0^{\mathbf{Free}}(\vec{\mathcal{C}})$ from Proposition 7:

$$\begin{aligned}C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \max\ & \lfloor m\rfloor \\ \text{s.t. } & E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}\geq m,\ 0\leq E_{XY}\leq F_{XY},\ \mathcal{L}_{XY}^{\mathrm{NS}}(F_{XY})=\frac{m}{d_X}I_{XY},\ E_Y=I_Y.\end{aligned} \tag{D3}$$

The inequality constraint $E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}\geq m$ can be changed to the equality constraint $E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}=m$. To see this, suppose $m':=E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}>m$. Then, we can replace F_{XY} by $F_{XY}+\frac{m'-m}{d_X}I_{XY}$ so that m', E_{XY}, and $F_{XY}+\frac{m'-m}{d_X}I_{XY}$ form a feasible solution. Since this works for every m, we arrive at the claimed SDP in Eq. (D1).

To derive the dual SDP (D2), we first note that

$$\begin{aligned}-C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = \log_2 \min\ & \lceil -m\rceil \\ \text{s.t. } & E_{XY}\star J_{XY}^{\vec{\mathcal{C}}}=m,\ 0\leq E_{XY}\leq F_{XY},\ \mathcal{L}_{XY}^{\mathrm{NS}}(F_{XY})=\frac{m}{d_X}I_{XY},\ E_Y=I_Y.\end{aligned} \tag{D4}$$

The Lagrangian associated with the above SDP is

$$\begin{aligned}
&L\left(m, E_{XY}, F_{XY}, \lambda, K^{(1)}_{XY}, K^{(2)}_{XY}, K^{(3)}_{XY}, K^{(4)}_{Y}\right) \\
\coloneqq & -m + \lambda\left(E_{XY} \star J^{\vec{\mathcal{C}}}_{XY} - m\right) - \left\langle K^{(1)}_{XY}, E_{XY}\right\rangle + \left\langle K^{(2)}_{XY}, E_{XY} - F_{XY}\right\rangle \\
&+ \left\langle K^{(3)}_{XY}, \mathcal{L}^{\mathrm{NS}}_{XY}(F_{XY}) - \frac{m}{d_X} I_{XY}\right\rangle + \left\langle K^{(4)}_{Y}, E_Y - I_Y\right\rangle
\end{aligned} \tag{D5}$$

$$\begin{aligned}
= & -m\left(1 + \lambda + \frac{1}{d_X}\operatorname{Tr}\left[K^{(3)}_{XY}\right]\right) + \left\langle \lambda\left(J^{\vec{\mathcal{C}}}_{XY}\right)^{\mathrm{T}} - K^{(1)}_{XY} + K^{(2)}_{XY} + I_X \otimes K^{(4)}_{Y}, E_{XY}\right\rangle \\
&+ \left\langle \mathcal{L}^{\mathrm{NS}}_{XY}\left(K^{(3)}_{XY}\right) - K^{(2)}_{XY}, F_{XY}\right\rangle - \operatorname{Tr}\left[K^{(4)}_{Y}\right],
\end{aligned} \tag{D6}$$

where $\lambda, K^{(1)}_{XY}, K^{(2)}_{XY}, K^{(3)}_{XY}, K^{(4)}_{Y}$ are dual variables. For the corresponding Lagrange dual function

$$g\left(\lambda, K^{(1)}_{XY}, K^{(2)}_{XY}, K^{(3)}_{XY}, K^{(4)}_{Y}\right) \coloneqq \inf_{m, E_{XY}, F_{XY}} L\left(m, E_{XY}, F_{XY}, \lambda, K^{(1)}_{XY}, K^{(2)}_{XY}, K^{(3)}_{XY}, K^{(4)}_{Y}\right) \tag{D7}$$

to be bounded, the dual variables must satisfy

$$\begin{aligned}
&K^{(1)}_{XY} \geq 0, \quad K^{(2)}_{XY} \geq 0, \quad 1 + \frac{1}{d_X}\operatorname{Tr}\left[K^{(3)}_{XY}\right] = -\lambda, \\
&\lambda\left(J^{\vec{\mathcal{C}}}_{XY}\right)^{\mathrm{T}} + K^{(2)}_{XY} + I_X \otimes K^{(4)}_{Y} = K^{(1)}_{XY}, \quad \mathcal{L}^{\mathrm{NS}}_{XY}\left(K^{(3)}_{XY}\right) = K^{(2)}_{XY}.
\end{aligned} \tag{D8}$$

Note that variables λ ,$K^{(1)}_{XY}$ and $K^{(2)}_{XY}$ can be removed from the above constraints, resulting in the simplified constraints

$$\mathcal{L}^{\mathrm{NS}}_{XY}\left(K^{(3)}_{XY}\right) + I_X \otimes K^{(4)}_{Y} \geq \left(1 + \frac{\operatorname{Tr}\left[K^{(3)}_{XY}\right]}{d_X}\right)\left(J^{\vec{\mathcal{C}}}_{XY}\right)^{\mathrm{T}}, \quad \mathcal{L}^{\mathrm{NS}}_{XY}\left(K^{(3)}_{XY}\right) \geq 0. \tag{D9}$$

Defining $M_{XY} \coloneqq K^{(3)}_{XY}$ and $N_Y \coloneqq K^{(4)}_{Y}$, we arrive at the dual SDP of maximizing Lagrange dual function g with the above constraints:

$$\begin{aligned}
&-C^{\mathbf{Free}}_0(\vec{\mathcal{C}}) \geq \log_2 \max\ \lceil -\operatorname{Tr}[N_Y] \rceil \\
&\qquad \text{s.t. } \mathcal{L}^{\mathrm{NS}}_{XY}(M_{XY}) \geq 0,\ \mathcal{L}^{\mathrm{NS}}_{XY}(M_{XY}) + I_X \otimes N_Y \geq \left(1 + \frac{\operatorname{Tr}[M_{XY}]}{d_X}\right)\left(J^{\vec{\mathcal{C}}}_{XY}\right)^{\mathrm{T}},
\end{aligned} \tag{D10}$$

where the inequality is due to weak duality. Then, $C^{\mathbf{Free}}_0(\vec{\mathcal{C}}) \leq \log_2 \min\ \lfloor \operatorname{Tr}[N_Y] \rfloor$ with the same constraints, resulting in the claimed SDP in Eq. (D2) with the additional observation that dropping the transpose acting on $J^{\vec{\mathcal{C}}}_{XY}$ does not affect the optimization result. ■

Following similar steps, we can derive the SDP for the capacity assisted by causally definite supermaps and its dual.

Corollary 2 (SDPs for one-shot zero-error classical capacity assisted by free causally definite supermaps) *Given a list of two quantum channels $\vec{\mathcal{C}}$, its one-shot zero-error classical capacity assisted by free causally definite supermaps can be formulated as*

$$\begin{aligned}
&C^{\mathbf{FreeDef}}_0(\vec{\mathcal{C}}) = \log_2 \max\ \lfloor m \rfloor \\
&\qquad \text{s.t. } (E_{XY} + G_{XY}) \star J^{\vec{\mathcal{C}}}_{XY} = m,\ 0 \leq E_{XY} \leq F_{XY},\ 0 \leq G_{XY} \leq H_{XY}, \\
&\qquad\quad {}_{[1-Y_2]}F_{XY} = 0,\ {}_{[1-Y_1]}F_{X_1Y_1} = 0,\ {}_{[1-Y_1]}H_{XY} = 0,\ {}_{[1-Y_2]}H_{X_2Y_2} = 0, \\
&\qquad\quad \operatorname{Tr}[F_{XY} + H_{XY}] = m d_Y,\ E_Y + G_Y = I_Y.
\end{aligned} \tag{D11}$$

$$
\begin{aligned}
\leq \log_2 \min\ & \lfloor \mathrm{Tr}\left[K_Y\right] \rfloor \\
\text{s.t. } & M_{XY} \geq 0,\ {}_{[1-X_2]}M_{XY_1} = 0,\ M_{X_1} = d_{X_2}\lambda I_{X_1}, \\
& M_{XY} + I_X \otimes K_Y \geq (\lambda+1) J_{XY}^{\vec{\mathcal{C}}}, \\
& N_{XY} \geq 0,\ {}_{[1-X_1]}N_{XY_2} = 0,\ N_{X_2} = d_{X_1}\lambda I_{X_2}, \\
& N_{XY} + I_X \otimes K_Y \geq (\lambda+1) J_{XY}^{\vec{\mathcal{C}}}
\end{aligned}
\tag{D12}
$$

with primal variables $m, E_{XY}, F_{XY}, G_{XY}, H_{XY}$ and dual variables $\lambda, M_{XY}, N_{XY}, K_Y$.

Theorem 1 *Let $\vec{\mathcal{C}} \coloneqq (\mathcal{A}, \mathcal{A})$ be a pair of identical amplitude damping channels $\mathcal{A}$ with damping rates 0.1. The one-shot zero-error classical capacity of $\vec{\mathcal{C}}$ assisted by causally indefinite supermaps is strictly larger than the capacity assisted by causally definite supermaps:*

$$
C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = 1 > C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = 0. \tag{D13}
$$

Proof We first prove $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = 1$ by showing that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \geq 1$ using the primal SDP (D1) and $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \leq 1$ using the dual SDP (D2). To see that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \geq 1$, we note that

$$
F_{XY} = \mathrm{diag}\left(\frac{13}{20}, \frac{3}{5}, \frac{11}{20}, 0, \frac{3}{5}, \frac{11}{20}, \frac{3}{5}, \frac{1}{20}, \frac{11}{20}, \frac{3}{5}, \frac{1}{4}, \frac{4}{5}, 0, \frac{1}{20}, \frac{4}{5}, \frac{27}{20}\right),
$$

$$
E_{XY} = \begin{pmatrix}
\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{9} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{10}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{233}{4860} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{233}{4860} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{\sqrt{10}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{10}} & 0 & 0 & \frac{7}{15} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{233}{4860} & 0 & 0 \\
0 & 0 & \frac{1}{\sqrt{10}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7}{15} & 0 \\
\frac{4}{9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1387}{1620}
\end{pmatrix} \tag{D14}
$$

form a feasible solution to SDP (D1) with $m = 2$, implying that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \geq \log_2 2 = 1$. To see that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \leq 1$, we note that

$$N_Y = \frac{5}{2}|00\rangle\!\langle 00|_{A_1B_1},$$

$$M_{XY} = \begin{pmatrix}
160 & 0 & 0 & \frac{3808}{25} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3808}{25} & 0 & 0 & \frac{289}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{737}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{737}{50} & 0 \\
\frac{3808}{25} & 0 & 0 & \frac{7263}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 145 & 0 & 0 & \frac{6879}{50} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{737}{50} & 0 & 0 & \frac{737}{50} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{737}{50} & 0 & 0 & \frac{737}{50} & 0 & 0 & 0 & 0 \\
\frac{3808}{25} & 0 & 0 & 145 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7263}{50} & 0 & 0 & \frac{6879}{50} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{737}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{737}{50} & 0 \\
\frac{289}{2} & 0 & 0 & \frac{6879}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{6879}{50} & 0 & 0 & \frac{3263}{25}
\end{pmatrix} \tag{D15}$$

form a feasible solution to SDP (D2) with $\mathrm{Tr}\left[N_Y\right] = \frac{5}{2}$, implying that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) \leq \log_2 \left\lfloor \frac{5}{2} \right\rfloor = \log_2 2 = 1$. Hence, we conclude that $C_0^{\mathbf{Free}}(\vec{\mathcal{C}}) = 1$.

Now, we prove $C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = 0$ using the dual SDP (D12). We note that

$$\lambda = 200, \quad K_Y = \frac{39}{20}|00\rangle\langle 00|_Y,$$

$$M_{XY} = \begin{pmatrix}
200 & 0 & 0 & \frac{1903}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{952}{5} & 0 & 0 & \frac{361}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{189}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{931}{50} & 0 \\
\frac{1903}{10} & 0 & 0 & \frac{1811}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4528}{25} & 0 & 0 & \frac{8589}{50} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{373}{20} & 0 & 0 & \frac{3719}{200} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{50} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3719}{200} & 0 & 0 & \frac{1859}{100} & 0 & 0 & 0 & 0 \\
\frac{952}{5} & 0 & 0 & \frac{4528}{25} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3627}{20} & 0 & 0 & \frac{4294}{25} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{931}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{459}{25} & 0 \\
\frac{361}{2} & 0 & 0 & \frac{8589}{50} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4294}{25} & 0 & 0 & \frac{16299}{100}
\end{pmatrix},$$

$$N_{XY} = \begin{pmatrix}
200 & 0 & 0 & \frac{952}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1903}{10} & 0 & 0 & \frac{361}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{93}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{37189}{2000} & 0 \\
\frac{952}{5} & 0 & 0 & \frac{3627}{20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4528}{25} & 0 & 0 & \frac{4294}{25} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{189}{10} & 0 & 0 & \frac{931}{50} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{50} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{931}{50} & 0 & 0 & \frac{459}{25} & 0 & 0 & 0 & 0 \\
\frac{1903}{10} & 0 & 0 & \frac{4528}{25} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1811}{10} & 0 & 0 & \frac{8589}{50} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{37189}{2000} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1859}{100} & 0 \\
\frac{361}{2} & 0 & 0 & \frac{4294}{25} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8589}{50} & 0 & 0 & \frac{16299}{100}
\end{pmatrix} \tag{D16}$$

form a feasible solution to SDP (D12) with $\mathrm{Tr}\,[K_Y] = \frac{39}{20}$, implying that $C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) \leq \log_2 \left\lfloor \frac{39}{20} \right\rfloor = \log_2 1 = 0$. Since, by definition, $C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) \geq 0$, we conclude that $C_0^{\mathbf{FreeDef}}(\vec{\mathcal{C}}) = 0$. ■

Supplementary Note E: No advantage for Pauli channels

Theorem 2 *Given a list of N, possibly different, Pauli channels $\vec{\mathcal{P}}$, it holds that*

$$C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}}) = C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) \tag{E1}$$

for every error tolerance $\epsilon \geq 0$.

Proof For a q-qubit Pauli channel $\mathcal{P}(\cdot) = \sum_{i_1 \dots i_q} p_{i_1 \dots i_q} (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})(\cdot)(\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})$, its Choi operator is

$$J_{XY}^{\mathcal{P}} = \sum_{i_1 \dots i_q} p_{i_1 \dots i_q} \left(I_X \otimes (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})_Y\right) |\Gamma\rangle\langle\Gamma|_{XY} \left(I_X \otimes (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})_Y\right) = \sum_{i_1 \dots i_q} p_{i_1 \dots i_q} \Gamma_{XY}^{i_1 \dots i_q}, \tag{E2}$$

where $\Gamma_{XY}^{i_1 \dots i_q} := \left(I_X \otimes (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})_Y\right) |\Gamma\rangle\langle\Gamma|_{XY} \left(I_X \otimes (\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_q})_Y\right)$ and $|\Gamma\rangle_{XY} := \sum_{j=1}^{2^q} |j\rangle_X \otimes |j\rangle_Y$. The Choi operator of N Pauli channels $\vec{\mathcal{P}}=(\mathcal{P}_1,\ldots,\mathcal{P}_N)$ with the j-th Pauli channel $\mathcal{P}_j$ acting on q_j number of qubits is

$$J_{XY}^{\vec{\mathcal{P}}}=\bigotimes_{j=1}^{N}\left(\sum_{i_1\ldots i_{q_j}}p_{i_1\ldots i_{q_j}}\Gamma_{X_jY_j}^{i_1\ldots i_{q_j}}\right)=\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}},\tag{E3}$$

where $q_{\rm tot}=\sum_{j=1}^{N}q_j$ is the total number of qubits $\vec{\mathcal{P}}$ acts on, and each $p_{i_1\ldots i_{q_{\rm tot}}}$ is a probability determined by all $p_{i_1\ldots i_{q_j}}$. Then, according to Proposition 7, the one-shot ϵ-error classical capacity of $\vec{\mathcal{P}}$ assisted by general supermaps can be formulated as

$$\begin{aligned}C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}})=\log_2\max\lfloor m\rfloor&\\ \text{s.t. } E_{XY}\star\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\geq m(1-\epsilon),\ E_Y=I_Y,&\\ 0\leq E_{XY}\leq F_{XY},\ \mathcal{L}_{XY}^{\rm NS}\left(F_{XY}\right)=\frac{mI_{XY}}{2^{q_{\rm tot}}}.&\end{aligned}\tag{E4}$$

Assume that m^*, E_{XY}^*, and F_{XY}^* form an optimal solution to the above maximization program. For every $i_1\ldots i_{q_{\rm tot}}$, define $\alpha_{i_1\ldots i_{q_{\rm tot}}}:=\frac{1}{4^{q_{\rm tot}}}\operatorname{Tr}\left[E_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]$ and $\beta_{i_1\ldots i_{q_{\rm tot}}}:=\frac{1}{4^{q_{\rm tot}}}\operatorname{Tr}\left[F_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]$. Then, m^*, $\widetilde{E}_{XY}:=\sum_{i_1\ldots i_{q_{\rm tot}}}\alpha_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}$, and $\widetilde{F}_{XY}:=\sum_{i_1\ldots i_{q_{\rm tot}}}\beta_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}$ also form an optimal solution, as we shall show now.

For the error tolerance constraint, observe that

$$\widetilde{E}_{XY}\star\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}=\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\widetilde{E}_{XY}\star\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\tag{E5}$$

$$=\sum_{i_1\ldots i_{q_{\rm tot}}}\sum_{j_1\ldots j_{q_{\rm tot}}}\alpha_{j_1\ldots j_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{j_1\ldots j_{q_{\rm tot}}}\star\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\tag{E6}$$

$$=\sum_{i_1\ldots i_{q_{\rm tot}}}\alpha_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\cdot 4^{q_{\rm tot}}\tag{E7}$$

$$=\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\operatorname{Tr}\left[E_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]\tag{E8}$$

$$=E_{XY}^*\star\sum_{i_1\ldots i_{q_{\rm tot}}}p_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\tag{E9}$$

$$\geq m^*(1-\epsilon),\tag{E10}$$

where in the third equality, we use the identity $\operatorname{Tr}\left[\Gamma_{XY}^{j_1\ldots j_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]=4^{q_{\rm tot}}\delta_{i_1\ldots i_{q_{\rm tot}},j_1\ldots j_{q_{\rm tot}}}$. Since $E_{XY}^*\geq 0$, we know $\operatorname{Tr}\left[E_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]$ and thus $\alpha_{i_1\ldots i_{q_{\rm tot}}}$ are nonnegative for every $i_1\ldots i_{q_{\rm tot}}$, implying that $\widetilde{E}_{XY}\geq 0$. Similarly, because $E_{XY}^*\leq F_{XY}^*$, we have $\operatorname{Tr}\left[E_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]\leq\operatorname{Tr}\left[F_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]$, and thus $\alpha_{i_1\ldots i_{q_{\rm tot}}}\leq\beta_{i_1\ldots i_{q_{\rm tot}}}$ for every $i_1\ldots i_{q_{\rm tot}}$, implying that $\widetilde{E}_{XY}\leq\widetilde{F}_{XY}$. For the causality constraint, first note that, for every $i_1\ldots i_{q_{\rm tot}}$,

$$\operatorname{Tr}\left[F_{XY}^*\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]=\operatorname{Tr}\left[F_{XY}^*\mathcal{L}_{XY}^{\rm NS}\left(\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right)\right]\tag{E11}$$

$$=\operatorname{Tr}\left[\mathcal{L}_{XY}^{\rm NS}\left(F_{XY}^*\right)\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]\tag{E12}$$

$$=\frac{m^*}{2^{q_{\rm tot}}}\operatorname{Tr}\left[I_{XY}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}\right]\tag{E13}$$

$$=m^*,\tag{E14}$$

where the first equality is due to that $\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}$ is the Choi operator of a no-signaling N-partite channel, and the second equality holds because the projection $\mathcal{L}_{XY}^{\rm NS}$ is self-dual. Then, $\beta_{i_1\ldots i_{q_{\rm tot}}}=\frac{m^*}{4^{q_{\rm tot}}}$ for every $i_1\ldots i_{q_{\rm tot}}$, resulting in

$$\widetilde{F}_{XY}=\frac{m^*}{4^{q_{\rm tot}}}\sum_{i_1\ldots i_{q_{\rm tot}}}\Gamma_{XY}^{i_1\ldots i_{q_{\rm tot}}}=\frac{m^*I_{XY}}{2^{q_{\rm tot}}},\tag{E15}$$

where the second equality is due to the identity $\sum_{i_1\dots i_{q_{\rm tot}}} \Gamma_{XY}^{i_1\dots i_{q_{\rm tot}}} = 2^{q_{\rm tot}} I_{XY}$. By Eq. (E15), it is straightforward to see that the causality constraint $\mathcal{L}_{XY}^{\rm NS}\left(\widetilde{F}_{XY}\right) = \frac{m^* I_{XY}}{2^{q_{\rm tot}}}$ holds. Finally, consider that

$$\widetilde{E}_Y = \sum_{i_1\dots i_{q_{\rm tot}}} \alpha_{i_1\dots i_{q_{\rm tot}}} \operatorname{Tr}_X\left[\Gamma_{XY}^{i_1\dots i_{q_{\rm tot}}}\right] \tag{E16}$$

$$= \frac{1}{4^{q_{\rm tot}}} \sum_{i_1\dots i_{q_{\rm tot}}} \operatorname{Tr}\left[E_{XY}^* \Gamma_{XY}^{i_1\dots i_{q_{\rm tot}}}\right] I_Y \tag{E17}$$

$$= \frac{1}{4^{q_{\rm tot}}} \operatorname{Tr}\left[E_{XY}^* \left(2^{q_{\rm tot}} I_{XY}\right)\right] I_Y \tag{E18}$$

$$= \frac{\operatorname{Tr}\left[E_{XY}^*\right]}{2^{q_{\rm tot}}} I_Y \tag{E19}$$

$$= I_Y, \tag{E20}$$

where the last equality is due to $\operatorname{Tr}\left[E_{XY}^*\right] = \operatorname{Tr}\left[E_Y^*\right] = \operatorname{Tr}\left[I_Y\right] = 2^{q_{\rm tot}}$.

So far, we have shown that m^*, $\widetilde{E}_{XY}$, and $\widetilde{F}_{XY}$ satisfy all the constraints in SDP (E4). Since m^* is optimal, they form an optimal solution. Now, by Eq. (C72),

$$\begin{aligned} C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) = \log_2 \max\ & \lfloor m \rfloor \\ \text{s.t. } & E_{XY} \star \sum_{i_1\dots i_{q_{\rm tot}}} p_{i_1\dots i_{q_{\rm tot}}} \Gamma_{XY}^{i_1\dots i_{q_{\rm tot}}} \geq m(1-\epsilon),\ E_Y = I_Y, \\ & 0 \leq E_{XY} \leq F_X \otimes I_Y,\ \operatorname{Tr}\left[F_X\right] = m. \end{aligned} \tag{E21}$$

Clearly, m^*, $\widetilde{E}_{XY}$, and $\frac{1}{2^{q_{\rm tot}}}\widetilde{F}_X$ form a feasible solution as $\widetilde{F}_{XY} = \frac{m^*}{2^{q_{\rm tot}}} I_X \otimes I_Y = \frac{1}{2^{q_{\rm tot}}}\widetilde{F}_X \otimes I_Y$. Thus, we know $C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) \geq \log_2\lfloor m^*\rfloor$. Since $C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) \leq C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}})$ and $C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}}) = \log_2\lfloor m^*\rfloor$, we conclude that $C_\epsilon^{\mathbf{FreePar}}(\vec{\mathcal{P}}) = C_\epsilon^{\mathbf{Free}}(\vec{\mathcal{P}})$. ■

Supplementary Note F: No advantage for channel simulation

Theorem 3 *For any quantum channel $\mathcal{C}$ and error tolerance $\epsilon \geq 0$, the one-shot classical simulation cost assisted by free general supermaps is equal to that assisted by free parallel supermaps, i.e.,*

$$S_\epsilon^{\mathbf{Free}}(\mathcal{C}) = S_\epsilon^{\mathbf{FreePar}}(\mathcal{C}). \tag{F1}$$

Proof By definition, it is straightforward to see that $S_\epsilon^{\mathbf{Free}}(\mathcal{C}) \leq S_\epsilon^{\mathbf{FreePar}}(\mathcal{C})$ since $\mathbf{FreePar} \subseteq \mathbf{Free}$. Now, we show that $S_\epsilon^{\mathbf{FreePar}}(\mathcal{C})$ is also a lower bound to $S_\epsilon^{\mathbf{Free}}(\mathcal{C})$.

The first step is to express $S_\epsilon^{\mathbf{Free}}(\mathcal{C})$ as an optimization program in terms of the Choi operator of channel $\mathcal{C}$ with input A and output B. The diamond distance between two quantum channels $\mathcal{C}_1$ and $\mathcal{C}_2$ can be written as an SDP [82]:

$$\begin{aligned} \frac{1}{2}\left\|\mathcal{C}_1 - \mathcal{C}_2\right\|_\diamond = \min\ & \mu \\ \text{s.t. } & Y_{AB} \geq 0,\ Y_{AB} \geq J_{AB}^{\mathcal{C}_1} - J_{AB}^{\mathcal{C}_2},\ \operatorname{Tr}_B\left[Y_{AB}\right] \leq \mu I_A, \end{aligned} \tag{F2}$$

where $J_{AB}^{\mathcal{C}_1}$ and $J_{AB}^{\mathcal{C}_2}$ are the Choi operators of the corresponding channels. It then follows that the simulation cost $S_\epsilon^{\mathbf{Free}}(\mathcal{C})$ can be formulated as

$$\begin{aligned} S_\epsilon^{\mathbf{Free}}(\mathcal{C}) = \log_2 \min\ & m \\ \text{s.t. } & Y_{AB} \geq 0,\ Y_{AB} \geq J_{AB}^{\mathcal{M}} - J_{AB}^{\mathcal{C}},\ Y_A \leq \epsilon I_A,\ J_{AB}^{\mathcal{M}} = J_{XYAB}^{\mathcal{S}} \star J_{XY}^{\vec{\Delta}^m}, \\ & J_{XYAB}^{\mathcal{S}} \geq 0,\ \mathcal{L}_{XY}^{\rm NS}\left(J_{XYA}^{\mathcal{S}}\right) = \frac{I_{XYA}}{d_X},\ {}_{[1-A]}J_{YAB}^{\mathcal{S}} = 0, \end{aligned} \tag{F3}$$

where m is minimized over positive integers. The constraints in the last line is requiring $\mathcal{S}$ to be a supermap in $\mathbf{Free}$. We obtain a lower bound to $S_\epsilon^{\mathbf{Free}}(\mathcal{C})$ by relaxing the constraint $\mathcal{L}_{XY}^{\rm NS}\left(J_{XYA}^{\mathcal{S}}\right) = \frac{I_{XYA}}{d_X}$ into constraints $J_A^{\mathcal{M}} = I_A$ and $J_{YA}^{\mathcal{S}} = I_{YA}$. To show that this is indeed a relaxation, we now prove that these two relaxed constraints are implications of $\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathcal{S}}_{XYA}\right) = \frac{I_{XYA}}{d_X}$. For the first constraint, consider that

$$\begin{aligned} J^{\mathcal{M}}_{A} &= \mathrm{Tr}_B\left[J^{\mathcal{S}}_{XYAB} \star J^{\vec{\Delta}^m}_{XY}\right] && \text{(F4)}\\ &= \mathrm{Tr}_{XYB}\left[J^{\mathcal{S}}_{XYAB}\left(\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\vec{\Delta}^m}_{XY}\right)\otimes I_{AB}\right)\right] && \text{(F5)}\\ &= \mathrm{Tr}_{XY}\left[\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathcal{S}}_{XYA}\right)\left(J^{\vec{\Delta}^m}_{XY}\otimes I_A\right)\right] && \text{(F6)}\\ &= \mathrm{Tr}_{XY}\left[\frac{I_{XYA}}{d_X}\left(J^{\vec{\Delta}^m}_{XY}\otimes I_A\right)\right] && \text{(F7)}\\ &= I_A, && \text{(F8)} \end{aligned}$$

which is what we set out to show. For the other constraint, note that, in one direction, we have

$$\begin{aligned} \mathrm{Tr}_X\left[\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathcal{S}}_{XYA}\right)\right] &= \mathrm{Tr}_X\left[\frac{I_{XYA}}{d_X}\right] && \text{(F9)}\\ &= I_{YA}. && \text{(F10)} \end{aligned}$$

In another direction, consider that

$$\mathrm{Tr}_X\left[\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathcal{S}}_{XYA}\right)\right] = \left(\bigotimes_{j=1}^{N}\mathrm{Tr}_{X_j}\circ\mathcal{L}^{\mathrm{NS}}_{X_jY_j}\right)\left(J^{\mathcal{S}}_{XYA}\right), \tag{F11}$$

where N is the number of ideal classical channels in the list $\vec{\Delta}^m$. Observe that for any linear operator $P_{X_jY_jR}$ with an arbitrary system R,

$$\begin{aligned} \mathrm{Tr}_{X_j}\circ\mathcal{L}^{\mathrm{NS}}_{X_jY_j}\left(P_{X_jY_jR}\right) &= \mathrm{Tr}_{X_j}\left[{}_{[1-Y_j+X_jY_j]}\left(P_{X_jY_jR}\right)\right] && \text{(F12)}\\ &= \mathrm{Tr}_{X_j}\left[P_{X_jY_jR} - P_{X_jR}\otimes\frac{I_{Y_j}}{d_{Y_j}} + \frac{I_{X_jY_j}}{d_{X_jY_j}}\otimes P_R\right] && \text{(F13)}\\ &= P_{Y_jR}. && \text{(F14)} \end{aligned}$$

As a result, Eq. (F11) gives

$$\mathrm{Tr}_X\left[\mathcal{L}^{\mathrm{NS}}_{XY}\left(J^{\mathcal{S}}_{XYA}\right)\right] = J^{\mathcal{S}}_{YA}, \tag{F15}$$

which, together with Eq. (F10), implies

$$J^{\mathcal{S}}_{YA} = I_{YA}. \tag{F16}$$

Hence, we arrive at

$$\begin{aligned} S^{\mathbf{Free}}_{\epsilon}(\mathcal{C}) \geq \log_2 \min\ & m\\ \text{s.t. } & Y_{AB}\geq 0,\ Y_{AB}\geq J^{\mathcal{M}}_{AB} - J^{\mathcal{C}}_{AB},\ Y_A\leq \epsilon I_A,\ J^{\mathcal{M}}_{AB} = J^{\mathcal{S}}_{XYAB}\star J^{\vec{\Delta}^m}_{XY},\\ & J^{\mathcal{S}}_{XYAB}\geq 0,\ J^{\mathcal{M}}_{A} = I_A,\ J^{\mathcal{S}}_{YA} = I_{YA},\ {}_{[1-A]}J^{\mathcal{S}}_{YAB} = 0. \end{aligned} \tag{F17}$$

Similar to what we did in deriving the SDPs for the capacities, here we use the symmetry of $J^{\vec{\Delta}^m}_{XY}$ to simplify the above optimization program, arriving at

$$\begin{aligned} S^{\mathbf{Free}}_{\epsilon}(\mathcal{C}) \geq \log_2 \min\ & \lceil \mathrm{Tr}\left[F_B\right]\rceil\\ \text{s.t. } & Y_{AB}\geq 0,\ Y_{AB}\geq E_{AB} - J^{\mathcal{C}}_{AB},\ Y_A\leq\epsilon I_A,\\ & 0\leq E_{AB}\leq I_A\otimes F_B,\ E_A = I_A. \end{aligned} \tag{F18}$$

Observe that the right hand side is exactly the SDP for one-shot classical simulation cost assisted by parallel supermaps as given in Ref. [39]. Hence, $S^{\mathbf{Free}}_{\epsilon}(\mathcal{C}) \geq S^{\mathbf{FreePar}}_{\epsilon}(\mathcal{C})$. Therefore, we conclude that $S^{\mathbf{Free}}_{\epsilon}(\mathcal{C}) = S^{\mathbf{FreePar}}_{\epsilon}(\mathcal{C})$. ■