Probabilistic exact universal quantum circuits for transforming unitary operations

Marco Túlio Quintino ⬡, Qingxiuxiong Dong ⬡, Atsushi Shimbo, Akihito Soeda, and Mio Murao ⬡

Department of Physics, Graduate School of Science, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan

This paper addresses the problem of designing universal quantum circuits to transform k uses of a d-dimensional unitary input operation into a unitary output operation in a probabilistic heralded manner. Three classes of protocols are considered, parallel circuits, where the input operations can be performed simultaneously, adaptive circuits, where sequential uses of the input operations are allowed, and general protocols, where the use of the input operations may be performed without a definite causal order. For these three classes, we develop a systematic semidefinite programming approach that finds a circuit which obtains the desired transformation with the maximal success probability. We then analyze in detail three particular transformations: unitary transposition, unitary complex conjugation, and unitary inversion. For unitary transposition and unitary inverse, we prove that for any fixed dimension d, adaptive circuits have an exponential improvement in terms of uses k when compared to parallel ones. For unitary complex conjugation and unitary inversion we prove that if the number of uses k is strictly smaller than $d - 1$, the probability of success is necessarily zero. We also discuss the advantage of indefinite causal order protocols over causal ones and introduce the concept of delayed input-state quantum circuits.

I. INTRODUCTION

In quantum mechanics, deterministic transformations between states are represented by quantum channels and probabilistic transformations by quantum instruments, which consist of quantum channels followed by a quantum measurement. Understanding the properties of quantum channels and quantum instruments is a standard and well established field of research with direct impact for theoretical and applied quantum physics [1,2]. Similarly to states, quantum channels may also be subjected to universal transformation in a paradigm usually referred as *higher order* transformations. Higher order transformations can be formalized by quantum supermaps [3,4] and physically implemented by means of quantum circuits (see Fig. 1). Despite its fundamental value and potential for applications (e.g., quantum circuit designing [3], quantum process tomography [5], testing causal hypothesis [6], channel discrimination [7], aligning reference frames [8,9], and analyzing the role of causal order [10,11]), higher order transformations are still not well understood when compared to quantum channels and quantum instruments.

Reversible operations play an important role in mathematics and in various physical theories such as quantum mechanics and thermodynamics. In quantum mechanics, reversible operations are represented by unitary operators [12,13]. This work considers *universal* transformations between reversible quantum transformations, that is, we seek for quantum circuits which implement the desired transformation for any unitary operation of some fixed dimension without any further specific details of the input unitary operation. From a practical perspective, this universal requirement ensures that the circuit does not require any readjustments or modification when different inputs are considered and the circuit implements the desired transformation even when the description of the

d-dimensional reversible operation is unknown. Note that the universal requirement also imposes strong constraints on transformations which can be physically realized. A well-known example which pinpoints these constraints when considering quantum states is quantum cloning, although it is simple to construct a quantum device that clones qubits which are promised to be in state $|0\rangle$ or $|1\rangle$, it is not possible to design a universal quantum transformation that clones all qubit states [14]. Another interesting example can be found in Ref. [15] where the authors consider universal NOT gates for qubits.

Universal transformations on reversible quantum operations have been studied from several perspectives and motivations such as gate discrimination [16,17], cloning unitary operations [18], preventing quantum systems to evolve [19,20], designing quantum circuits [3], learning the action of a unitary [9,21–24], transforming unitary operations into their complex conjugate [25], understanding the role of causal order in quantum mechanics [10,26], and others [27,28]. Probabilistic exact transformations between multiple uses of reversible operations via quantum circuits have been considered in Ref. [29] where the authors target transforming an arbitrary unitary operation into its inverse and in Ref. [30] where the authors consider the case where the unitary input operation and the unitary output operation are two different representations of the same group element. Also, Refs. [20,31] consider the probabilistic exact circuits which act only in an auxiliary system which interacts to the target one via some random Hamiltonian.

This paper is focused on designing universal quantum circuits which are not exclusively tailored for a particular class of input operations. That is, it should attain the desired transformation for any d-dimensional unitary operation even if its description is not known. In particular, we focus on probabilistic heralded transformations between multiple uses of
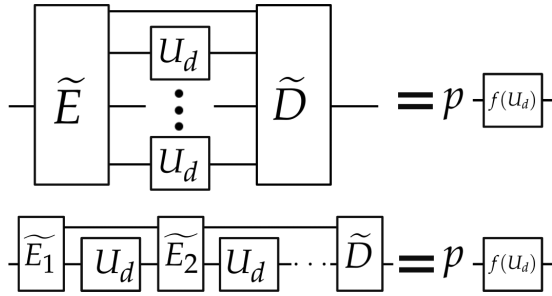
FIG. 1. Pictorial representation of parallel and adaptive quantum circuits that transform k uses of a d-dimensional arbitrary unitary input operation described by U_d into another unitary operation described by $f(U_d)$. The circuit elements $\widetilde{E}$ and $\widetilde{E}_i$ are quantum deterministic operations, i.e., quantum channels, that may be interpreted as encoders and the element $\widetilde{D}$ stands for decoder, a probabilistic quantum operation (quantum instrument), that is, a quantum channel with a quantum measurement that when the "correct" outcome is obtained, the target transformation is obtained perfectly.

reversible operations. In particular, we focus on probabilistic heralded transformations between multiple uses of reversible operations. More precisely, we consider circuits which make use of a quantum measurement with an output associated with success and, when the success outcome is obtained, the transformation is implemented perfectly. We consider three classes of quantum protocols: parallel circuits, where the input operations can be performed simultaneously, adaptive circuits, where the input operations may be used sequentially, and general protocols, which may not be realizable by quantum circuits but are consistent with quantum theory when the use of the input operations may be performed in an indefinite causal order [10,11,32]. We present a systematic approach based on semidefinite programming that allows us to analyze transformations which is linear on quantum operations. We then analyze in details three particular transformations: unitary transposition, unitary complex conjugation, and unitary inversion.

Section II reviews results related to quantum circuits such as quantum supermaps, quantum combs, process matrices, and other important concepts. Section III presents a general semidefinite programming (SDP) approach to design optimal probabilistic exact quantum circuits. Section IV introduces the concept of delayed input-states quantum circuits. Section V analyzes circuits for implementing unitary complex conjugation. Section VI analyzes circuits for unitary transposition. Section VII analyzes circuits for unitary inversion and Sec. VIII concludes and discusses the main results.

II. REVIEW ON HIGHER ORDER QUANTUM OPERATIONS AND SUPERMAPS

In this section we establish our notations and review how to represent and analyze quantum circuits and transformations between quantum operations in terms of supermaps. We refer to transformations between quantum states as lower order operations (i.e., quantum channels and quantum instruments) and transformations between quantum operations (e.g., channels, instruments, quantum circuits) as higher order

operations, which will be named as superchannels and superinstruments.

A. The Pills-Choi-Jamiołkowski isomorphism

We start by reviewing the Choi isomorphism [33–35] (also known as Pills-Choi-Jamiołkowski isomorphism), a useful way to represent linear maps and particularly convenient for completely positive ones. Let $L(\mathcal{H})$ be the set of linear operators mapping a linear (Hilbert) space $\mathcal{H}$ to another space isomorphic to itself. This work only considers finite dimensional quantum systems, hence all linear spaces $\mathcal{H}$ are isomorphic to $\mathbb{C}^d$, d-dimensional complex linear spaces. Any map[1] $\widetilde{\Lambda} : L(\mathcal{H}_{\text{in}}) \to L(\mathcal{H}_{\text{out}})$ has a one to one representation via its Choi operator defined by

$$\mathfrak{C}(\widetilde{\Lambda}) := \sum_{ij} \widetilde{\Lambda}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \in L(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}}), \quad (1)$$

where $\{|i\rangle\}$ is an orthonormal basis. An important theorem regarding the Choi representation is that a map $\widetilde{\Lambda}$ is completely positive (CP) if and only if its Choi operator $\mathfrak{C}(\widetilde{\Lambda})$ is positive semidefinite [35].

When the Choi operator $\mathfrak{C}(\widetilde{\Lambda})$ of some map is given, one can obtain the action of $\widetilde{\Lambda}$ on any operator $\rho_{\text{in}} \in L(\mathcal{H}_{\text{in}})$ via the relation

$$\widetilde{\Lambda}(\rho_{\text{in}}) = \text{Tr}_{\text{in}}\big(\mathfrak{C}(\widetilde{\Lambda})\big[I_{\text{out}} \otimes \rho_{\text{in}}^T\big]\big), \quad (2)$$

where ρ_{in}^T is the transposition of the operator ρ_{in} in terms of the $\{|i\rangle\}$ basis of $\mathcal{H}_{\text{in}}$.

We now present a useful mathematical identity regarding the Choi isomorphism. Let U_d, A, and B be d-dimensional unitary operators.[2] Any unitary quantum operation $\widetilde{U_d}(\rho) := U_d \rho U_d^\dagger$ can be represented by its Choi operator as $\mathfrak{C}(\widetilde{U_d})$ and a straightforward calculation shows that

$$[A \otimes B]\mathfrak{C}(\widetilde{U_d})[A^\dagger \otimes B^\dagger] = \mathfrak{C}(\widetilde{AU_dB^T}). \quad (3)$$

B. Supermaps with single use of the input operations

In quantum mechanics, physical states are represented by positive operators $\rho \in L(\mathcal{H})$, $\rho \geqslant 0$, with unit trace $\text{Tr}(\rho) = 1$. In this language, universal transformations between quantum states are represented by linear maps, to which we refer as just maps, $\widetilde{\Lambda} : L(\mathcal{H}_{\text{in}}) \to L(\mathcal{H}_{\text{out}})$ that are CP [1,2]. Here, by universal we mean that the map $\widetilde{\Lambda}$ is defined for all quantum states $\rho \in L(\mathcal{H}_{\text{in}})$ and the physical transformation can be applied to any of these states. Quantum channels are deterministic quantum operations and are represented by CP maps that preserve the trace of all quantum states $\rho \in L(\mathcal{H}_{\text{in}})$. Probabilistic heralded universal transformations between quantum states are represented by quantum instruments, a set of CP maps $\{\widetilde{\Lambda}_i\}$ that sum to a trace preserving one, i.e., $\widetilde{\Lambda} := \sum_i \widetilde{\Lambda}_i$ is trace preserving (TP). Quantum

---

[1] Symbols with a tilde represent linear maps.

[2] In principle, this identity also holds even when U_d, A, and B are not unitary but general d-dimensional linear operators.

instruments describe measurements in quantum mechanics.[3] When the set of instruments $\{\widetilde{\Lambda}_i\}$ is performed, the outcome i is obtained with probability $\mathrm{Tr}(\widetilde{\Lambda}_i(\rho))$, and the state ρ is transformed to $\frac{\widetilde{\Lambda}_i(\rho)}{\mathrm{Tr}(\widetilde{\Lambda}_i(\rho))}$.

An important realization theorem of quantum channels is given by the Stinespring dilation [36] which states that every quantum channel $\widetilde{\Lambda}$ can be realized by first applying an isometric operation, i.e., a unitary with auxiliary systems and then discarding a part of the system. More precisely, every CPTP map $\widetilde{\Lambda} : L(\mathcal{H}_{\mathrm{in}}) \to L(\mathcal{H}_{\mathrm{out}})$ can be written as

$$\widetilde{\Lambda}(\rho) = \mathrm{Tr}_{\mathcal{A}'}(U[\rho \otimes \sigma_A]U^\dagger), \qquad (4)$$

where $\sigma_A \in L(\mathcal{A})$ is some (constant) auxiliary state, $U : \mathcal{H}_{\mathrm{in}} \otimes \mathcal{A} \to \mathcal{H}_{\mathrm{out}} \otimes \mathcal{A}'$ is a unitary acting on the main and auxiliary system, and $\mathrm{Tr}_{\mathcal{A}'}$ is a partial trace on some subsystem $\mathcal{A}'$ such that $\mathcal{H}_{\mathrm{out}} \otimes \mathcal{A}'$ is isomorphic to $\mathcal{H}_{\mathrm{in}} \otimes \mathcal{A}$.

Quantum instruments also have an important realization theorem that follows from Naimark's dilation [1,37]. Every quantum instrument can be realized by a quantum channel followed by a projective measurement, i.e., a measurement which all its POVM elements are projectors, on some auxiliary system. More precisely, if $\{\widetilde{\Lambda}_i : L(\mathcal{H}_{\mathrm{in}}) \to L(\mathcal{H}_{\mathrm{out}})\}$ represents an instrument, there exist a quantum channel $\widetilde{C} : L(\mathcal{H}_{\mathrm{in}}) \to L(\mathcal{H}_{\mathrm{out}} \otimes \mathcal{A})$ and a projective measurement given by $\{\Pi_i\}$ where $\Pi_i \in L(\mathcal{A})$ which satisfies

$$\widetilde{\Lambda}_i(\rho) = \mathrm{Tr}_{\mathcal{A}'}(\widetilde{C}(\rho)[I_{\mathcal{H}_{\mathrm{out}}} \otimes \Pi_i]). \qquad (5)$$

We now define universal transformations between quantum operations in an analogous way in terms of linear supermaps [4]. Linear supermaps, to which we also refer as just supermaps, are linear transformations between maps. A supermap[4]

$$\widetilde{\widetilde{S}} : [L(H_2) \to L(H_3)] \to [L(H_1) \to L(H_4)] \qquad (6)$$

represents transformations between input maps $\widetilde{\Lambda}_{\mathrm{in}} : L(\mathcal{H}_2) \to L(\mathcal{H}_3)$ to output ones $\widetilde{\Lambda}_{\mathrm{out}} : L(\mathcal{H}_1) \to L(\mathcal{H}_4)$.

For instance, let $\widetilde{U}_d(\rho) := U_d \rho U_d^\dagger$ be the map associated with the d-dimension unitary operation U_d, the supermap that transforms a unitary operation into its inverse is given by $\widetilde{\widetilde{S}}(\widetilde{U}_d) = \widetilde{U}_d^{-1}$.

We say that a supermap $\widetilde{\widetilde{S}}$ is TP preserving (TPP) if it transforms TP maps into TP maps. Similarly, a supermap is CP preserving (CPP) when it transforms CP maps into CP maps, and completely CP preserving (CCPP) if the every trivial extension $\widetilde{\widetilde{S}} \otimes \widetilde{\widetilde{I}}$, of $\widetilde{\widetilde{S}}$ is CPP, where $\widetilde{\widetilde{I}}(\widetilde{\Lambda}) = \widetilde{\Lambda}$, $\forall \widetilde{\Lambda}$. A *superchannel* $\widetilde{\widetilde{C}}$ is a supermap which respects two basic constraints: (1) it transforms valid quantum channels into valid quantum channels (hence CPP and TPP); and (2) when performed to a part of a quantum channel the global channel remains valid (hence CCPP).

---

[3]Every instrument $\{\widetilde{\Lambda}_i\}$ corresponds to a unique positive operator valued measure (POVM) measurement $\{M_i\}$, $M_i \in L(\mathcal{H}_{\mathrm{in}})$, $M_i \geqslant 0$, $\sum_i M_i = I_{\mathcal{H}_{\mathrm{in}}}$, such that $\mathrm{Tr}(\rho M_i) = \mathrm{Tr}(\widetilde{\Lambda}_i(\rho))$ for every state $\rho \in L(\mathcal{H}_{\mathrm{in}})$. The POVM $\{M_i\}$ can be written explicitly as $M_i = \widetilde{\Lambda}_i^\dagger(I_{\mathcal{H}_{\mathrm{out}}})$ where $\widetilde{\Lambda}_i^\dagger$ is the adjoint map of $\widetilde{\Lambda}_i$.

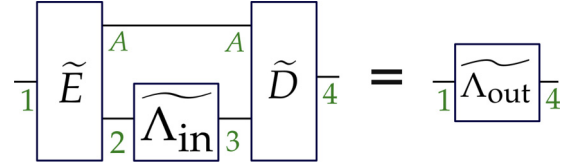[4]Symbols with a double tilde represent linear supermaps.

FIG. 2. Every superchannel $\widetilde{\widetilde{C}} : [L(\mathcal{H}_2) \to L(\mathcal{H}_3)] \to [L(\mathcal{H}_1) \to L(\mathcal{H}_4)]$ transforming input channels $\widetilde{\Lambda}_{\mathrm{in}} : L(\mathcal{H}_2) \to L(\mathcal{H}_3)$ into output channels $\widetilde{\Lambda}_{\mathrm{out}} : L(\mathcal{H}_1) \to L(\mathcal{H}_4)$ can be decomposed as $\widetilde{\widetilde{C}}(\widetilde{\Lambda}_{\mathrm{in}}) = \widetilde{D} \circ [\widetilde{\Lambda}_{\mathrm{in}} \otimes \widetilde{I}_A] \circ \widetilde{E}$ where the encoder operation $\widetilde{E} : L(\mathcal{H}_1) \to L(\mathcal{H}_2) \otimes L(\mathcal{H}_A)$ is an isometry operation, $L(\mathcal{H}_A)$ is a space for some possible auxiliary system, and the decoder $\widetilde{D} : L(\mathcal{H}_3 \otimes \mathcal{H}_A) \to L(\mathcal{H}_4)$ is a unitary operation followed by a partial trace on a part of the system.

Any one single use superchannel $\widetilde{\widetilde{C}}$ has a deterministic realization in quantum theory and similarly to the Stinespring dilation theorem, it can be shown that every superchannel admits a decomposition in terms of *encoder* and *decoder* of the form [4]

$$\widetilde{\widetilde{C}}(\widetilde{\Lambda}) = \widetilde{D} \circ [\widetilde{\Lambda} \otimes \widetilde{I}_A] \circ \widetilde{E}, \qquad (7)$$

where $\widetilde{E} : L(\mathcal{H}_1) \to L(\mathcal{H}_2) \otimes L(\mathcal{H}_A)$ is an isometry which maps an input-state $\rho_{\mathrm{in}} \in L(\mathcal{H}_1)$ to the space where the map $\widetilde{\Lambda}$ acts and an auxiliary one $L(\mathcal{H}_A)$, $\widetilde{I}_A$ is the identity map on the auxiliary system [i.e., $\widetilde{I}_A(\sigma_A) = \sigma_A$, $\forall \sigma_A \in L(\mathcal{H}_A)$], $\widetilde{D} : L(\mathcal{H}_3 \otimes \mathcal{H}_A) \to L(\mathcal{H}_4)$ is a unitary operation followed by a partial trace on a part of the system (see Fig. 2).

The Choi representation allows us to describe any supermap $\widetilde{\widetilde{S}} : [L(\mathcal{H}_2) \to L(\mathcal{H}_3)] \to [L(\mathcal{H}_1) \to L(\mathcal{H}_4)]$ as a map $\widetilde{S} : L(\mathcal{H}_3 \otimes \mathcal{H}_2) \to L(\mathcal{H}_4 \otimes \mathcal{H}_1)$ acting on Choi operators. And by exploiting the Choi representation again, we can represent any supermap $\widetilde{\widetilde{S}}$ by a linear operator $S := \mathfrak{C}(\widetilde{\widetilde{S}}) \in L(\mathcal{H}_4 \otimes \mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2)$, which is useful to characterize the set of supermaps with quantum realizations. In Refs. [3,4] the authors show that a $\widetilde{\widetilde{C}}$ is a superchannel if and only if its Choi representation C respects

$$C \geqslant 0,$$
$$\mathrm{Tr}_4 C = \mathrm{Tr}_{43} C \otimes \frac{I_3}{d_3},$$
$$\mathrm{Tr}_{234} C = \mathrm{Tr}_{1234} C \otimes \frac{I_1}{d_1}, \qquad (8)$$
$$\mathrm{Tr}(C) = d_1 d_3,$$

where d_i is the dimension of the linear space $\mathcal{H}_i$. We remark that although we introduce the general formalism where the dimensions d_i may depend on i, we focus our results to the case where $d_i = d$ is independent of i.

Supermaps with a probabilistic heralded quantum realization are given by *superinstruments* and play a similar role of instruments in higher order quantum operations, that is, it formalizes probabilistic transformations on quantum operations. Superinstruments are a set of CCPP supermaps $\{\widetilde{\widetilde{C}}_i\}$ that sums to a superchannel. The probability of obtaining the outcome i when the superinstrument $\{\widetilde{\widetilde{C}}_i\}$ acts on the input-map $\widetilde{\Lambda}$ and input-state ρ is $\mathrm{Tr}([\widetilde{\widetilde{C}}_i(\widetilde{\Lambda})](\rho))$ and the state $\frac{[\widetilde{\widetilde{C}}_i(\widetilde{\Lambda})](\rho)}{\mathrm{Tr}([\widetilde{\widetilde{C}}_i(\widetilde{\Lambda})](\rho))}$ is
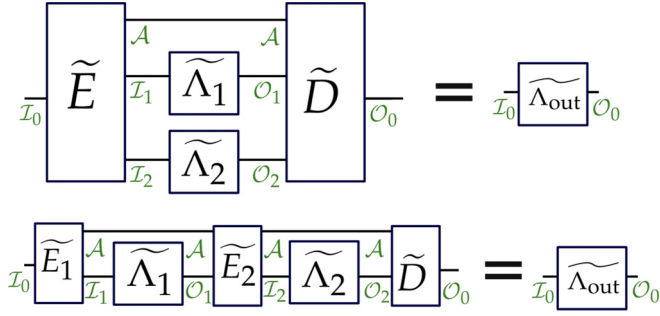
FIG. 3. Illustration of a parallel (upper circuit) and an adaptive (lower circuit) protocol that transform a pair of quantum operations $\widetilde{\Lambda_1}$ and $\widetilde{\Lambda_2}$ into an output one $\widetilde{\Lambda_{\text{out}}}$.

obtained. It follows from Ref. [38] that any superinstrument can be realized by a superchannel followed by a projective measurement, or equivalently,

$$\widetilde{\widetilde{C}}_i(\widetilde{\Lambda}) = \widetilde{D}_i \circ [\widetilde{\Lambda} \otimes \widetilde{I}_A] \circ \widetilde{E}, \tag{9}$$

here $\widetilde{E} : L(\mathcal{H}_1) \rightarrow L(\mathcal{H}_2) \otimes L(\mathcal{H}_A)$ is an isometry which maps an input-state $\rho \in L(\mathcal{H}_1)$ to the space where the map $\widetilde{\Lambda}$ acts and an auxiliary one $L(\mathcal{H}_A)$, $\widetilde{I}_A$ is the identity map on the auxiliary system [i.e., $\widetilde{I}_A(\sigma_A) = \sigma_A$, $\forall \sigma_A \in L(\mathcal{H}_A)$], and the maps $\widetilde{D}_i : L(\mathcal{H}_3 \otimes \mathcal{H}_A) \rightarrow L(\mathcal{H}_4)$ form an instrument corresponding to a projective measurement.

C. Supermaps involving k input operations

In the previous section we have introduced supermaps corresponding to protocols involving a single use of an input operation. We now consider protocols transforming k, potentially different, operations into another. Let $\widetilde{\widetilde{C}}$ be a superchannel which transforms k input channels[5] $\widetilde{\Lambda}_j : L(\mathcal{I}_j) \rightarrow L(\mathcal{O}_j)$ with $j \in \{1, \ldots, k\}$ into an output one $\widetilde{\Lambda}_0 : L(\mathcal{I}_0) \rightarrow L(\mathcal{O}_0)$. We also define the total input-state space as $\mathcal{I} := \bigotimes_{j=1}^{k} \mathcal{I}_j$ and the total output-space state $\mathcal{O} := \bigotimes_{j=1}^{k} \mathcal{O}_j$, hence $\widetilde{\widetilde{C}} : [L(\mathcal{I}) \rightarrow L(\mathcal{O})] \rightarrow [L(\mathcal{I}_0) \rightarrow L(\mathcal{O}_0)]$.

Similarly to the single input-channel case, superchannels transforming k quantum operations are supermaps which: (1) transform k valid quantum channels into a valid quantum channel; and (2) when performed on a part of a quantum channel, the global channel remains valid. Differently from the $k = 1$ case, not all superchannels have a deterministic quantum realization in terms of encoders and decoders in the standard quantum circuit formalism [10]. This impossibility occurs because the definition of quantum realization does not require explicitly that the k channels should be used in a definite causal order and it allows protocols which use the input channels with an indefinite causal order [11].

Protocols that can be implemented in the standard causally ordered circuit formalism are referred to as quantum networks and quantum combs [3] or channels with memory [39]. We

divide these ordered circuits into two classes: (a) parallel ones where k channels can be used simultaneously; and (b) adaptive ones where the k channels are explored in a causal sequential circuit (see Fig. 3).

Parallel protocols transforming k channels are very similar to single-channel superchannels presented in the last subsection. Define $\widetilde{\Lambda} := \bigotimes_{j=1}^{k} \widetilde{\Lambda}_j$, a superchannel $\widetilde{\widetilde{C}}$ represents a parallel protocol if it can be written as $\widetilde{\widetilde{C}}(\widetilde{\Lambda}) = \widetilde{D} \circ [\widetilde{\Lambda} \otimes \widetilde{I}_A] \circ \widetilde{E}$ for some channels $\widetilde{E} : L(\mathcal{I}_0) \rightarrow L(\mathcal{I} \otimes \mathcal{A})$ and $\widetilde{D} : L(\mathcal{O} \otimes \mathcal{A}) \rightarrow \mathcal{O}_0$. It follows from the characterization of Eq. (8) that a Choi operator $C \in L(\mathcal{I}_0 \otimes \bigotimes_{j=1}^{k} \mathcal{I}_j \otimes \bigotimes_{j=1}^{k} \mathcal{O}_j \otimes \mathcal{O}_0)$ represents a parallel protocol if and only if

$$C \geqslant 0,$$

$$\text{Tr}_{\mathcal{O}_0} C = \text{Tr}_{\mathcal{O}\mathcal{O}_0} C \otimes \frac{I_{\mathcal{O}}}{d_{\mathcal{O}}},$$

$$\text{Tr}_{\mathcal{I}\mathcal{O}\mathcal{O}_0} C = \text{Tr}_{\mathcal{I}_0\mathcal{I}\mathcal{O}\mathcal{O}_0} C \otimes \frac{I_{\mathcal{I}_0}}{d_{\mathcal{I}_0}}, \tag{10}$$

$$\text{Tr}(C) = d_{\mathcal{I}_0} d_{\mathcal{O}}.$$

Adaptive circuits can exploit a causal order relation between the channels $\widetilde{\Lambda}_j$ to implement protocols that cannot be done in a parallel way. A simple example is the supermap that concatenates the channels $\widetilde{\Lambda}_1$ and $\widetilde{\Lambda}_2$ to obtain $\widetilde{\Lambda}_2 \circ \widetilde{\Lambda}_1$. This supermap has a trivial implementation in an adaptive circuit (just concatenates the channels) but cannot be implemented in a *deterministic* parallel scheme.

A superchannel $\widetilde{\widetilde{C}} : [L(\mathcal{I}) \rightarrow L(\mathcal{O})] \rightarrow [L(\mathcal{I}_0) \rightarrow L(\mathcal{O}_0)]$ corresponds to an adaptive circuit if it can be written as[6]

$$\widetilde{\widetilde{C}}(\widetilde{\Lambda}) = \widetilde{D} \circ [\widetilde{\Lambda}_k \otimes \widetilde{I}_A] \circ \widetilde{E}_k \circ \cdots \circ [\widetilde{\Lambda}_1 \otimes \widetilde{I}_A] \circ \widetilde{E}_1 \tag{11}$$

for some channels $\widetilde{E}_1 : L(\mathcal{I}_0) \rightarrow L(\mathcal{I}_1 \otimes \mathcal{A})$, $\widetilde{E}_i : L(\mathcal{O}_{i-1} \otimes \mathcal{A}) \rightarrow (\mathcal{I}_i \otimes \mathcal{A})$ with $i \in \{2, \ldots, k\}$, and $\widetilde{D} : L(\mathcal{O}_k \otimes \mathcal{A}) \rightarrow L(\mathcal{O}_0)$. A Choi operator $C \in L(\mathcal{I}_0 \otimes \bigotimes_{j=1}^{k} \mathcal{I}_j \otimes \bigotimes_{j=1}^{k} \mathcal{O}_j \otimes \mathcal{O}_0)$ represents an adaptive superchannel if and only if [3,39]

$$C \geqslant 0,$$

$$\text{Tr}_{\mathcal{O}_0} C = \text{Tr}_{\mathcal{O}_k\mathcal{O}_0} C \otimes \frac{I_{\mathcal{O}_k}}{d_{\mathcal{O}_k}},$$

$$\text{Tr}_{\mathcal{I}_i} C^{(i)} = \text{Tr}_{\mathcal{I}_i\mathcal{O}_i} C^{(i)} \otimes \frac{I_{\mathcal{O}_i}}{d_{\mathcal{O}_i}} \quad \forall i \in \{k, \ldots, 2\}, \tag{12}$$

$$\text{Tr}_{\mathcal{I}_1} C^{(1)} = \text{Tr}_{\mathcal{I}_0\mathcal{I}_1} C^{(1)} \otimes \frac{I_{\mathcal{I}_0}}{d_{\mathcal{I}_0}},$$

$$\text{Tr}(C) = d_{\mathcal{I}_0} d_{\mathcal{O}},$$

where $C^{(i)} := \text{Tr}_{\mathcal{I}_{i+1}\mathcal{O}_{i+1}} C^{(i+1)}$ for $i \in \{1, \ldots, k-1\}$ and $C^{(k)} := \text{Tr}_{\mathcal{O}_k\mathcal{O}_0} C$.

We now consider the most general protocols that transform k quantum channels into a single one. As mentioned before, these superchannels may have an indefinite causal order between the use of these k channels, hence they may not have an implementation in terms of encoders and decoders in the

---

[5]We remark that here the subindex j stands for a label for the channel $\widetilde{\Lambda}_j : L(\mathcal{I}_j) \rightarrow L(\mathcal{O}_j)$, not for some instrument element of an instrument $\{\Lambda_j\}$.

[6]Note that since we do not restrict the dimension of the auxiliary system $\mathcal{A}$, all parallel protocols can be realized by an adaptive circuit.

standard quantum circuit formalism. Even without necessarily having a realization by ordered circuits, it is possible to have a simple characterization of these general superchannels. Before presenting the necessary and sufficient condition for a general (possibly with an indefinite causal order) superchannel, it is convenient to introduce the trace and replace notation introduced in Ref. [26]. Let $A \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ be a general linear operator, we define $_{\mathcal{H}_2}A := \mathrm{Tr}_{\mathcal{H}_2} A \otimes \frac{I_{\mathcal{H}_2}}{d_{\mathcal{H}_2}}$. A Choi operator $C \in L(\mathcal{I}_0 \otimes \mathcal{I}_1 \otimes \mathcal{I}_2 \otimes \mathcal{O}_1 \otimes \mathcal{O}_2 \otimes \mathcal{O}_0)$ represents a general superchannel transforming $k = 2$ channels into a single one if and only if it respects [40]

$$C \geqslant 0,$$
$$_{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_0}C = _{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_2\mathcal{O}_0}C,$$
$$_{\mathcal{I}_2\mathcal{O}_2\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{O}_0}C,$$
$$_{\mathcal{O}_0}C + _{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{O}_0}C + _{\mathcal{O}_2\mathcal{O}_0}C,$$
$$_{\mathcal{I}\mathcal{O}\mathcal{O}_0}C = _{\mathcal{I}_0\mathcal{I}\mathcal{O}\mathcal{O}_0}C,$$
$$\mathrm{Tr}(C) = d_{\mathcal{I}_0}d_{\mathcal{O}_1}d_{\mathcal{O}_2}. \tag{13}$$

We remark that the bipartite process matrices presented in Refs. [11,26] correspond to a particular case of general superchannels with two input channels presented above. This correspondence is made by setting the dimension of the linear spaces $\mathcal{I}_0$ and $\mathcal{O}_0$ as one. This occurs because Refs. [11,26] focus on superchannels that transform pairs of instruments into probabilities, not into quantum operations. Also, the general superchannels presented in Eq. (13) are equivalent to the general process matrices presented in Ref. [40] which uses the terminology common past and common future to denote the spaces $\mathcal{I}_0$ and $\mathcal{O}_0$, respectively.

It is also possible to characterize general superchannels transforming k channels on terms of their Choi operators. For that, one can exploit the methods used in Refs. [26,40] to characterize process matrices (see also Ref. [41]). Using such methods, we have characterized general superchannels which transforms $k = 3$ input channels into a single output one. A Choi operator $C \in L(\mathcal{I}_0 \otimes \mathcal{I}_1 \otimes \mathcal{I}_2 \otimes \mathcal{I}_3 \otimes \mathcal{O}_1 \otimes \mathcal{O}_2 \otimes \mathcal{O}_3 \otimes \mathcal{O}_0)$ represents a general superchannel that transforms $k = 3$ channels into another one if and only if it respects

$$C \geqslant 0,$$
$$_{\mathcal{I}_1\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{O}_0}C = _{\mathcal{I}_1\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}_2\mathcal{O}_2\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}_1\mathcal{O}_1\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_2\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_0}C + _{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_2\mathcal{O}_0}C + _{\mathcal{I}_1\mathcal{O}_1\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}_2\mathcal{O}_2\mathcal{O}_0}C + _{\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{O}_0}C + _{\mathcal{I}_2\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C + _{\mathcal{O}_1\mathcal{I}_2\mathcal{O}_2\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C + _{\mathcal{O}_2\mathcal{I}_3\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{O}_0}C + _{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C = _{\mathcal{O}_1\mathcal{O}_0}C + _{\mathcal{O}_2\mathcal{O}_0}C + _{\mathcal{O}_3\mathcal{O}_0}C$$
$$+ _{\mathcal{O}_1\mathcal{O}_2\mathcal{O}_0}C + _{\mathcal{O}_1\mathcal{O}_3\mathcal{O}_0}C + _{\mathcal{O}_2\mathcal{O}_3\mathcal{O}_0}C,$$
$$_{\mathcal{I}\mathcal{O}\mathcal{O}_0}C = _{\mathcal{I}_0\mathcal{I}\mathcal{O}\mathcal{O}_0}C,$$
$$\mathrm{Tr}(C) = d_{\mathcal{I}_0}d_{\mathcal{O}_1}d_{\mathcal{O}_2}d_{\mathcal{O}_3}. \tag{14}$$

Similarly to the single use case, probabilistic heralded protocols are also represented by superinstruments. Superinstruments also admit a simple representation in terms of their induced Choi operators. A set of parallel, adaptive, and general superinstruments transforming k channels into another is given by a set of positive semidefinite operators $C_i \geqslant 0$ where $C := \sum_i C_i$ is a valid parallel, adaptive, and general superchannel, respectively. The probability of obtaining the outcome i when performing the superinstrument $\{\widetilde{\widetilde{C}}_i\}$ on k input channels represented by $\widetilde{\Lambda} := \bigotimes_{j=1}^k \widetilde{\Lambda}_j$ and the input-state ρ is given by $\mathrm{Tr}([\widetilde{\widetilde{C}}_i(\widetilde{\Lambda})](\rho))$.

III. OPTIMAL UNIVERSAL QUANTUM CIRCUITS VIA SDP

In this section we construct a systematic method to design probabilistic heralded quantum circuits for transforming multiple uses of the same unitary operations. Let $U_d : L(\mathbb{C}^d)$ be a d-dimensional unitary operator and $\widetilde{U}_d$ be a linear map representing the operation associated with U_d, that is, when the operation $\widetilde{U}_d$ is applied into a quantum state $\rho \in L(\mathbb{C}^d)$ the output is given by $\widetilde{U}_d(\rho) = U_d \rho U_d^\dagger$. We consider linear supermaps given by $f : \widetilde{U}_d \mapsto f(\widetilde{U}_d)$ mapping unitary operations into unitary operations. Our goal is to transform k uses of an arbitrary $\widetilde{U}_d$ into $f(\widetilde{U}_d)$ with the highest heralded constant probability p.

From the results of the previous section, this transformation can be implemented via quantum circuits when there exists a superinstrument element, i.e., a CCPP linear supermap, $\widetilde{\widetilde{S}}$ such that $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = p\, f(\widetilde{U}_d)$ for every unitary operation $\widetilde{U}_d$ (see Sec. II). We stress that, even though we have presented an explicit characterization of superinstruments in Sec. II, finding the optimal success probability for this transformation and its associated quantum circuit is, in general, a nontrivial task. First, note that action of the supermap f is only described for unitary channels[7] but the action of a superinstrument element $\widetilde{\widetilde{S}}$ must be defined for any CP linear map. The supermap $\widetilde{\widetilde{S}}$ can then be any CCPP linear supermap that extends the action of f from unitary operations to general CP maps (see Refs. [42,43] for a related lower-order version problem which consists of finding CP extensions of linear maps defined on subspaces). Second, since k uses of the input operation are available, it may be the case that even if f does not have a linear CCPP extension for some number of uses k_0 but it has for $k > k_0$ (see Refs. [15,44] for a lower-order analog of this problem where multiple copies of the input *state* can be used to implement a linear positive non-CP *map*).

Before presenting our general approach we illustrate the subtleties of this extension problem by discussing the universal channel complex conjugation studied in Ref. [25]. Let $\widetilde{\Lambda} : \mathcal{H}_{\mathrm{in}} \to \mathcal{H}_{\mathrm{out}}$ be a quantum channel with the Kraus decomposition given by $\widetilde{\Lambda}(\rho) = \sum_i K_i \rho K_i^\dagger$, we define the complex conjugate of $\widetilde{\Lambda}$ as the map which respects $\widetilde{\Lambda}^*(\rho) = \sum_i K_i^* \rho K_i^{*\dagger}$ for every ρ where the complex conjugation of K_i is made in a fixed orthonormal basis, e.g., the computational basis. One can show that, for any linear spaces $\mathcal{H}_{\mathrm{in}}$ and $\mathcal{H}_{\mathrm{out}}$ with dimension greater than or equal to two, CCPP supermaps

———————

[7]Since we have imposed that f is linear, f is also implicitly defined for linear combination of unitary operations.

respecting $\widetilde{\Lambda}^{\otimes k} \mapsto p\widetilde{\Lambda}^*$ for all channels $\widetilde{\Lambda}$ necessarily have $p = 0$ for any number of uses $k \in \mathbb{N}$ [45]. Hence, it is not possible to design a universal quantum circuit for probabilistic channel adjoint. However, if one relaxes the requirements of general channels and seeks for a quantum circuit that transforms only unitary operations into their adjoints, universal complex conjugation can be implemented deterministically in a parallel circuit with makes $k = d - 1$ uses of the input channel presented in Ref. [25]. In Sec. V we prove that $k = d - 1$ uses are not only sufficient but also necessary. We then see that the notion of CCPP extension and the number of uses play a crucial role in finding superinstruments that implement some desired transformation given by f.

We now present our SDP approach. Let $\{\widetilde{\widetilde{S}}, \widetilde{\widetilde{F}}\}$ be a superinstrument where the outcome of the element $\widetilde{\widetilde{S}}$ indicates success and the outcome of $\widetilde{\widetilde{F}}$ indicates failure. The problem of maximizing the success probability of transforming k uses of an arbitrary d-dimensional unitary input-operation $\widetilde{U_d}$ into $f(\widetilde{U_d})$ can be phrased as

$$\max p$$
$$\text{s.t.} \quad \widetilde{\widetilde{S}}(\widetilde{U_d}^{\otimes k}) = pf(\widetilde{U_d}) \quad \forall U_d; \quad (15)$$
$$\{\widetilde{\widetilde{S}}, \widetilde{\widetilde{F}}\} \text{ is a valid superinstrument,}$$

where the valid superinstrument representing parallel, adaptive, or general protocols. Using the characterization presented in Sec. II, we can rewrite the above maximization problem only in terms of linear and positive semidefinite constraints as

$$\max p$$
$$\text{s.t.} \quad \text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_d^{\otimes k}})^T_{\mathcal{IO}} \otimes I_{\mathcal{O}_0}]\big) = p\mathfrak{C}[f(\widetilde{U_d})] \quad \forall U_d,$$
$$S, F \in L(\mathcal{I}_0 \otimes \mathcal{I} \otimes \mathcal{O} \otimes \mathcal{O}_0), \quad S \geqslant 0, \ F \geqslant 0,$$
$$S + F \text{ is a valid superchannel.}$$
$$(16)$$

Note that the maximization problem presented in Eq. (16) must hold for all unitary operators U_d and has infinitely many constraints. This issue can be bypassed by noting that due to linearity, it is enough to check these constraints only for a set that spans the set spanned by Choi operators of unitary operations. That is, if we can write $\mathfrak{C}(\widetilde{U_d^{\otimes k}}) = \sum_i \alpha_i \mathfrak{C}(\widetilde{U_{d,i}^{\otimes k}})$ and it is true that

$$\text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_{d,i}^{\otimes k}})^T_{\mathcal{IO}} \otimes I_{\mathcal{O}_0}]\big) = p\mathfrak{C}[f(\widetilde{U_{d,i}})] \ \forall i, \quad (17)$$

we have that

$$\text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_d^{\otimes k}})^T_{\mathcal{IO}} \otimes I_{\mathcal{O}_0}]\big)$$
$$= \sum_i \text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \alpha_i\mathfrak{C}(\widetilde{U_{d,i}^{\otimes k}})^T_{\mathcal{IO}} \otimes I_{\mathcal{O}_0}]\big)$$
$$= p\sum_i \alpha_i\mathfrak{C}[f(\widetilde{U_{d,i}})]$$
$$= p\mathfrak{C}[f(\widetilde{U_d})]. \quad (18)$$

Also, one can always find a finite set, in particular, a basis, of unitary operations $\{\widetilde{U_{d,i}}\}$ that spans the set spanned by Choi

operators of d-dimensional unitary operations, i.e.,

$$\text{span}\big(\mathfrak{C}(\widetilde{U_d^{\otimes k}})|U_d \text{ is unitary}\big) = \text{span}\big(\mathfrak{C}(\widetilde{U_{d,i}^{\otimes k}})|U_{d,i} \in \{U_{d,i}\}\big).$$
$$(19)$$

Explicitly obtaining a basis for the subspace $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$ is, in general, not straightforward. For numerical purposes, this problem can be tackled by sampling a large number of unitaries U_d uniformly randomly (according to the Haar measure). If the dimension of this subspace is D, D unitaries sampled uniformly will be linearly independent with unit probability. Since checking linear independence can be done in an efficient way, we can construct a basis for this set by sampling unitaries randomly until we cannot find more linearly independent ones.

Also note that the dimension D of the subspace $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$ may grow very fast with k and d, this will increase the number of constraints in the SDP we have presented. Since having a large number of constraints may make the SDP intractable for practical purposes [it may take a very long time to run the code or to consume a very large amount of random-access memory (RAM)], it is worth noticing that if one runs the SDP (16) with a set of operators $\{U_{d,i}\}$ that do not form a basis for $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$, the solution p of the SDP is not the maximal success probability but an upper bound on the maximal success probability (it is the same SDP with fewer constraints). We also point out that since the methods to solve an SDP also provide the instrument element S that attains the maximal success probability p, even if the set $\{U_{d,i}\}$ does not form a basis for $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$, it may still be the case that the solution obtained is also the global optimal value.[8] In order to check this hypothesis we can extract the superinstrument element S of the SDP in which the operators $\{U_{d,i}\}$ that do not form a basis for $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$. Then, we generate a basis $\{U'_{d,j}\}$ for $\text{span}(\mathfrak{C}(\widetilde{U_d^{\otimes k}}) | U_d \text{ is unitary})$ and verify that

$$\text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_{d,j}'^{\otimes k}})^T_{\mathcal{IO}} \otimes I_{\mathcal{O}_0}]\big) = p\mathfrak{C}[f(\widetilde{U_{d,j}'})] \quad (20)$$

for every[9] j.

We have implemented our code using MATLAB [46] with the interpreter CVX [47] and tested with the solvers MOSEK, SeDuMi, and SDPT3 [48–50]. In Table I we apply this method

---

[8]We thank Alastair Abbott for pointing out this fact to us.

[9]When $d = 3$, $k = 2$, we have applied this technique to tackle the unitary transposition and inversion problem. In this case, we have run our numerical SDPs only for a subset of the space of unitary channels generated by $\text{span}(\mathfrak{C}(\widetilde{U_3^{\otimes 2}}) | U_3 \text{ is unitary})$. Numerically, we can see that the linear space spanned by $(\mathfrak{C}(\widetilde{U_3^{\otimes 2}}) | U_3 \text{ is unitary})$ has 994 linearly independent unitary channels but we have only considered a random subset containing 200 linear independent elements of the form $\mathfrak{C}(\widetilde{U_3^{\otimes 2}})$ in our calculations. After obtaining an upper bound to the problem, we have verified that the superinstrument element S transforms the full basis with 994 linearly independent unitary channels into their inverses, ensuring that the previous upper bound is tight.

TABLE I. Table with optimal success probability we have obtained for heralded protocols transforming k uses of $\widetilde{U_d}$ into a single use of its transpose $\widetilde{U_d^T}$. The values in bold were proved analytically and the rest were obtained via numerical SDP optimization.

$d = 2$	Parallel	Adaptive	Indefinite causal order
$k = 1$	$\frac{1}{4} = \textbf{0.25}$	$\frac{1}{4} = \textbf{0.25}$	$\frac{1}{4} = \textbf{0.25}$
$k = 2$	$\frac{2}{5} = \textbf{0.4}$	$0.4286 \approx \frac{3}{7}$	$0.4444 \approx \frac{4}{9}$
$k = 3$	$\frac{1}{2} = \textbf{0.5}$	$0.7500 \approx \frac{3}{4}$	0.9416
$d = 3$	Parallel	Adaptive	Indefinite causal order
$k = 1$	$\frac{1}{9} \approx \textbf{0.1111}$	$\frac{1}{9} \approx \textbf{0.1111}$	$\frac{1}{9} \approx \textbf{0.1111}$
$k = 2$	$\frac{2}{10} = \textbf{0.2}$	$0.2222 \approx \frac{2}{9}$	$0.2500 \approx \frac{2}{8}$

to obtain the maximal success probability to transform k uses of a d-dimensional unitary operation, i.e., $f(\widetilde{U_d}) = \widetilde{U_d^T}$ under different constraints. In Sec. VII C we reproduce Table I of Ref. [29] which contains results for the maximal probability for unitary inversion, i.e., $f(\widetilde{U_d}) = \widetilde{U_d^{-1}}$. All our code are available at Ref. [51] and can be freely used, edited, and distributed under the MIT license [52] and make extensive use of the toolbox QETLAB [53].

IV. DELAYED INPUT-STATE PROTOCOLS

In this section we define a particular subclass of quantum circuits in which we refer to delayed input-state protocols. This class consists of circuits where the input state is provided after the input operation which will be transformed (see Fig. 4). The concept of delayed input state generalizes the class of supermaps considered in the context of unitary learning and unitary store-and-retrieve problems [9,21–24]. As we will show next, parallel quantum circuits used for unitary transposition and unitary inversion can be assumed to be in the delayed input-state form without loss of generality
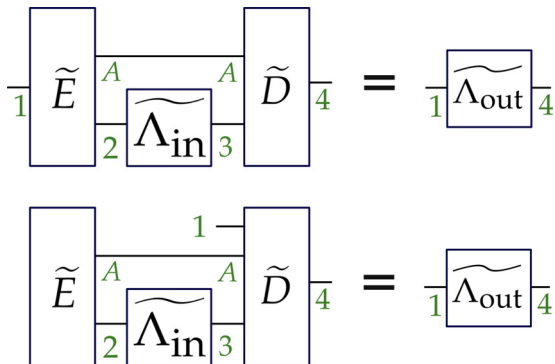

FIG. 4. Comparison between a standard quantum circuit (upper circuit) and a delayed input-state protocol (lower circuit) that transforms general operations. In a delayed input-state protocol, the input state labeled by the space 1 is not used by the encoder operation $\widetilde{E}$. The encoder only prepares a (potentially entangled) state which partially goes to the input-channel $\widetilde{\Lambda}_{\text{in}}$, and then to the decoder channel $\widetilde{D}$, which can perform a joint operation between the input state and the auxiliary system.

and the definition of delayed input-state protocols is useful to prove various theorems presented in this paper.

Consider a scenario where Alice has k uses of a general unitary operation $\widetilde{U_d}$ until some time t_1. In a later time t_2, where $\widetilde{U_d}$ cannot be accessed anymore, she would like to implement $f(\widetilde{U_d})$ on some arbitrary quantum state chosen at time t_2. This scenario can be seen as a particular case of the general unitary transformation problem where the input state is only provided after the operation $\widetilde{U_d}$. Let us start with the $k = 1$ case where only a single use of the general input-operation $\widetilde{\Lambda}_{\text{in}}$ is allowed (see Fig. 4).

In this single use case, every superchannel admits a realization in terms of a quantum circuit with an encoder and a decoder [4]. Let $\widetilde{\widetilde{C}}$ be a superchannel transforming an input-operation $\widetilde{\Lambda}_{\text{in}} : L(\mathcal{H}_2) \to L(\mathcal{H}_3)$ into $\widetilde{\widetilde{C}}(\widetilde{\Lambda}_{\text{in}}) = \widetilde{\Lambda}_{\text{out}} : L(\mathcal{H}_1) \to L(\mathcal{H}_4)$ and $\rho_{\text{in}} \in L(\mathcal{H}_1)$ be the input state on which she would like to apply $\widetilde{\Lambda}_{\text{out}}$. A protocol to implement the superchannel $\widetilde{\widetilde{C}}$ can be realized as follows:

(1) Alice performs an encoder channel $\widetilde{E} : L(\mathcal{H}_1) \to L(\mathcal{H}_2 \otimes \mathcal{H}_A)$ on the input-state $\rho_{\text{in}} \in L(\mathcal{H}_1)$.

(2) The input-operation $\widetilde{\Lambda}_{\text{in}} : L(\mathcal{H}_2) \to L(\mathcal{H}_3)$ is performed on a part of the state $\widetilde{E}(\rho_{\text{in}}) \in L(\mathcal{H}_2 \otimes \mathcal{H}_A)$.

(3) The decoder $\widetilde{D} : L(\mathcal{H}_3 \otimes \mathcal{H}_A) \to L(\mathcal{H}_4)$ is applied to the state $[\widetilde{\Lambda}_{\text{in}} \otimes \widetilde{I_A}](\widetilde{E}(\rho_{\text{in}}))$ to obtain the final output state

$$[\widetilde{\widetilde{C}}(\widetilde{\Lambda}_{\text{in}})](\rho_{\text{in}}) = \widetilde{\Lambda}_{\text{out}}(\rho_{\text{in}}). \tag{21}$$

In a delayed input-state protocol, the encoder channel $\widetilde{E}$ does not have access to the input-state ρ_{in}, since this state is only provided after the use of the operation $\widetilde{\Lambda}_{\text{in}}$. Instead of having an encoder channel, Alice must then prepare a fixed state $\phi_E \in L(\mathcal{H}_2 \otimes \mathcal{H}_A)$ that is independent of ρ_{in}. More precisely, a superchannel $\widetilde{\widetilde{C}}_D$ represents a $k = 1$ delayed input-state protocol if it can be realized by the following protocol:

(1) Alice prepares a state $\phi_E \in L(\mathcal{H}_2 \otimes \mathcal{H}_A)$.

(2) The input-operation $\widetilde{\Lambda}_{\text{in}} : L(\mathcal{H}_2) \to L(\mathcal{H}_3)$ is performed on a part of the state $\phi_E \in L(\mathcal{H}_2 \otimes \mathcal{H}_A)$ prepared by Alice.

(3) The decoder $\widetilde{D}_D : L(\mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \mathcal{H}_A) \to L(\mathcal{H}_4)$ is applied to the state $\rho_{\text{in}} \otimes [(\widetilde{\Lambda}_{\text{in}} \otimes \widetilde{I_A})(\phi_E)]$ to obtain the final output-state

$$[\widetilde{\widetilde{C}}_D(\widetilde{\Lambda}_{\text{in}})](\rho_{\text{in}}) = \widetilde{\Lambda}_{\text{out}}(\rho_{\text{in}}). \tag{22}$$

We now consider parallel delayed input-state protocols with $k > 1$ uses of the input-channel $\widetilde{\Lambda}_{\text{in}}$. By definition, a parallel superchannel $\widetilde{\widetilde{C}} : [L(\mathcal{I}) \to L(\mathcal{O})] \to [L(\mathcal{I}_0) \to L(\mathcal{O}_0)]$ that transforms k identical input operations into another can be represented by an encoder channel $\widetilde{E} : L(\mathcal{I}_0) \to L(\mathcal{I} \otimes \mathcal{A})$ and a decoder channel $\widetilde{D} : L(\mathcal{O} \otimes \mathcal{A}) \to L(\mathcal{O}_0)$ such that

$$\widetilde{\widetilde{C}}(\widetilde{\Lambda^{\otimes k}}) = \widetilde{D} \circ [\widetilde{\Lambda^{\otimes k}} \otimes \widetilde{I^{\otimes k}}] \circ \widetilde{E}. \tag{23}$$

That is, in order to perform the output operation $\widetilde{\Lambda}_{\text{out}} = \widetilde{\widetilde{C}}(\widetilde{\Lambda^{\otimes k}})$ on an arbitrary input-state $\rho_{\text{in}} \in L(\mathcal{I}_0)$, we first perform the encoder operation on ρ_{in}, then the k uses of $\widetilde{\Lambda}$ on a part of the output of the encoder, and then the decoder $\widetilde{D}$:

$$[\widetilde{\widetilde{C}}(\widetilde{\Lambda^{\otimes k}})](\rho_{\text{in}}) = \widetilde{D}\{[\widetilde{\Lambda^{\otimes k}} \otimes \widetilde{I_A}](\widetilde{E}(\rho_{\text{in}}))\}. \tag{24}$$

In a delayed input-state protocol the encoder cannot not make use of the input-state $\rho_{\rm in}$. Instead of an encoder channel $\widetilde{E}$ we now consider some fixed (potentially entangled) quantum state $\phi_E \in L(\mathcal{I} \otimes \mathcal{A})$. On a delayed input-state protocol, the decoder $\widetilde{D}_D : L(\mathcal{I}_0 \otimes \mathcal{O} \otimes \mathcal{A}) \to L(\mathcal{O}_0)$ acts directly on input-state $\rho_{\rm in}$. We then say that a parallel superchannel $\widetilde{\widetilde{C}}_D$ represents a delayed input-state parallel protocol if can be written as

$$[\widetilde{\widetilde{C}}_D(\widetilde{\Lambda^{\otimes k}})](\rho_{\rm in}) = \widetilde{D}_D\{[(\widetilde{\Lambda^{\otimes k}} \otimes \widetilde{I}_\mathcal{A})(\phi_E)] \otimes \rho_{\rm in}\}, \quad (25)$$

for some decoder channel $\widetilde{D}_D : L(\mathcal{I}_0 \otimes \mathcal{O} \otimes \mathcal{A}) \to L(\mathcal{O}_0)$ and some state $\phi_E \in L(\mathcal{I} \otimes \mathcal{A})$. If we define a $\psi_{\widetilde{\Lambda^{\otimes k}}} := [\widetilde{\Lambda^{\otimes k}} \otimes \widetilde{I}_\mathcal{A}](\phi_E)$, we can rewrite Eq. (25) as

$$[\widetilde{\widetilde{C}}_D(\widetilde{\Lambda^{\otimes k}})](\rho_{\rm in}) = \widetilde{D}(\psi_{\widetilde{\Lambda^{\otimes k}}} \otimes \rho). \quad (26)$$

Parallel delayed input-state superchannels $\widetilde{\widetilde{C}}_D$ also have a simple characterization in terms of its Choi operator $C_D \in L(\mathcal{I}_0 \otimes \mathcal{I} \otimes \mathcal{O} \otimes \mathcal{O}_0)$. Since the encoder acts trivially on the space $L(\mathcal{I}_o)$, it follows from the same tools used to characterize standard ordered circuits [3] that C_D represents a parallel delayed input-state protocol if and only

$$C_D \geqslant 0,$$
$$\mathrm{Tr}_{\mathcal{O}_0} C_D = \frac{I_{\mathcal{I}_0}}{d_{\mathcal{I}_0}} \otimes \mathrm{Tr}_{\mathcal{I}_0 \mathcal{O} \mathcal{O}_0} C_D \otimes \frac{I_\mathcal{O}}{d_\mathcal{O}}, \quad (27)$$
$$\mathrm{Tr}(C_D) = d_{\mathcal{I}_0} d_\mathcal{O}.$$

Or equivalently, C_D respects the standard parallel supermap restrictions of Eq. (10) and also

$$\mathrm{Tr}_{\mathcal{O} \mathcal{O}_0} C_D = \mathrm{Tr}_{\mathcal{I}_0 \mathcal{O} \mathcal{O}_0} C_D \otimes \frac{I_{\mathcal{I}_0}}{d_{\mathcal{I}_0}}. \quad (28)$$

The formal definition and a simple Choi characterization of adaptive delayed input-state protocols follow straightforwardly from the discussions of the parallel case presented here.[10] The case of superchannels with indefinite causal order is more subtle. Since they have no encoder and decoder ordered quantum circuit implementation their physical interpretation is not evident. We let the precise definition and the characterization of noncausally ordered delayed input-state protocols for future research.

Probabilistic heralded parallel (adaptive) delayed input-state protocols are given by superinstruments whose elements add to a superchannel representing a parallel (adaptive) delayed input-state protocol. It follows from the circuit realization of quantum instruments [38] that every parallel (adaptive) delayed input-state protocol can be realized by an encoder $(k - 1$ encoders) where the input state is not required and a decoder, which makes use of the input state, followed by a projective measurement.

We will now show that any probabilistic supermap can be implemented via a parallel probabilistic delayed input-state protocol with a smaller, but nonzero success probability. That is, if a supermap $\widetilde{\widetilde{S}}$ represents a superinstrument element of some higher order transformation, there exists a delayed input-state parallel superinstrument which, when successful, implements the action of $\widetilde{\widetilde{S}}$ in a probabilistic heralded way. This theorem holds true even if the supermap $\widetilde{\widetilde{S}}$ corresponds to an indefinite causal order protocol. Intuitively, one can understand this theorem in terms of state teleportation and probabilistic heralded gate teleportation (see Sec. VI A for a review of gate teleportation). In order to "parallelize" any superinstrument one can use the gate teleportation to rearrange the position of all input operations in parallel. Also, one can always delay the use of the input state by exploiting the state teleportation protocol [54]. Although the teleportation and gate teleportation protocol may fail, the success probability is strictly positive for any fixed dimension, ensuring that the success probability of the parallel circuit is nonzero.

Lemma 1. Let $\widetilde{\widetilde{S}} : [L(\bigotimes_{i=1}^k \mathcal{I}_i) \to L(\bigotimes_{o=1}^k \mathcal{O}_o)] \to [L(\mathcal{I}_0) \to L(\mathcal{O}_0)]$ be a supermap representing a general (possibly with indefinite causal order) probabilistic protocol that makes k uses of a unitary operation $\widetilde{U}_d$ and transforms to some other unitary operation $f(\widetilde{U}_d)$ with probability p_U, i.e., $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = p_U f(\widetilde{U}_d)$. There exists a parallel delayed input-state protocol implementing the supermap $\widetilde{\widetilde{S}}$ with a probability greater than or equal to $\frac{p_U}{d_{\rm total}}$, where $d_{\rm total}$ is the product of all linear space dimensions, i.e., $d_{\rm total} = \prod_{i=0}^k d_{\mathcal{I}_i} \prod_{i=0}^k d_{\mathcal{O}_i}$.

Proof. By assumption, $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = p_U f(\widetilde{U}_d)$ for all U_d with probability p_U. Since $\widetilde{\widetilde{S}}$ must be a superinstrument element, the corresponding Choi operator S is positive and respects $\mathrm{Tr}(S) \leqslant d_\mathcal{O} d_{\mathcal{I}_0}$, hence $0 \leqslant \frac{1}{d_{\rm total}} S \leqslant \frac{1}{d_\mathcal{I} d_\mathcal{O}} I$, where $I \in L(\mathcal{I}_0 \otimes \mathcal{I} \otimes \mathcal{O} \otimes \mathcal{O}_0)$ is the identity operator. Note that the Choi operator $C_P := \frac{1}{d_\mathcal{I} d_{\mathcal{O}_0}} I$ represents a valid parallel delayed input-state superchannel, i.e., it satisfies the parallel delayed input-state superchannel conditions of Eq. (27). We thus define the new superinstrument via the Choi of its elements as $S_P := \frac{1}{d_{\rm total}} S$ and $F_P := \frac{1}{d_\mathcal{I} d_{\mathcal{O}_0}} I - \frac{1}{d_{\rm total}} S$. It follows that $F_P \geqslant 0$ and $S_P + F_P = C_P = \frac{1}{d_\mathcal{I} d_{\mathcal{O}_0}} I$ is a valid parallel delayed input-state superchannel, hence the operators S_P and F_P form a valid delayed-input state parallel superinstrument. By linearity, we can verify that $\widetilde{\widetilde{S}}_P(\widetilde{U}_d^{\otimes k}) = \frac{1}{d_{\rm total}} \widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = \frac{p_U}{d_{\rm total}} f(\widetilde{U}_d)$, ensuring that when the output associated with S_P is obtained, the probabilistic parallel delayed input-state protocol represented by the superinstrument elements S_P and F_P performs the transformation of the supermap $\widetilde{\widetilde{S}}$ with probability $\frac{p_U}{d_{\rm total}}$. ∎

V. UNIVERSAL UNITARY COMPLEX CONJUGATION

In this section we consider the problem of transforming k uses of an arbitrary d-dimensional unitary $\widetilde{U}_d$ into its complex conjugate $\widetilde{U}_d^*$ for some fixed basis. We prove that when $k < d - 1$ uses are accessible, any exact unitary complex conjugation quantum protocol, including protocols with indefinite causal order, necessarily have zero success

probability. In Ref. [25] the authors present a deterministic parallel quantum circuit that transforms $k = d - 1$ uses of a d-dimensional unitary operation $\widetilde{U}_d$ into its complex conjugate[11] $\widetilde{U}_d^*$. Hence, when combined with Ref. [25], our result reveals a characteristic threshold property for exact unitary complex conjugation: if $k < d - 1$, universal exact unitary complex conjugation is impossible (zero success probability), if $k = d - 1$ exact unitary complex conjugation is possible with probability one with a parallel circuit implementation.

Theorem 1 (Unitary complex conjugation: no-go). Any universal probabilistic heralded quantum protocol (including protocols without definite causal order) transforming $k < d - 1$ uses of a d-dimensional unitary operation $\widetilde{U}_d$ into its complex conjugate $\widetilde{U}_d^*$ with probability p that does not depend on $\widetilde{U}_d$ necessarily has $p = 0$, i.e., null success probability.

Proof. From Lemma 1 we see that if there exists a superinstrument that transforms k uses of $\widetilde{U}_d$ into its complex conjugate $\widetilde{U}_d^*$ with some possibly smaller but still positive probability, there also exists a parallel superinstrument $\widetilde{\widetilde{S}}$ that transforms k uses of $\widetilde{U}_d$ into its complex conjugate $\widetilde{U}_d^*$ with some positive probability p, i.e., $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = p\widetilde{U}_d^*$. From the realization theorem of the superinstruments (see Eq. (9) and Ref. [9]), there exist an isometry $\widetilde{E} : L(\mathcal{I}_0) \rightarrow L(\mathcal{I}) \otimes L(\mathcal{A})$ and an instrument element corresponding to success $\widetilde{D}_S$: $L(\mathcal{I} \otimes \mathcal{A}) \rightarrow L(\mathcal{O}_0)$ such that

$$\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = \widetilde{D}_S \circ \left[\widetilde{U}_d^{\otimes k} \otimes \widetilde{I}_A \right] \circ \widetilde{E}. \tag{29}$$

Let $\rho_{\mathcal{IA}}$ be the Naimark dilation, the instrument element $\widetilde{D}_S$ is given by

$$\widetilde{D}_S(\rho_{\mathcal{IA}}) = \text{Tr}_{\mathcal{A}}(D\rho_{\mathcal{IA}}D^\dagger) \tag{30}$$

for some operator $D \in L(\mathcal{I} \otimes \mathcal{A})$. Set $\{|a\rangle\}$ as a basis for the auxiliary system $\mathcal{A}$. The previous equation becomes

$$\widetilde{D}_S(\rho_{\mathcal{IA}}) = \sum_a \langle a|D\rho_{\mathcal{IA}}D^\dagger|a\rangle. \tag{31}$$

The operators $D_a := \langle a|D$ form a possible set of operators realizing the instrument $\widetilde{D}_S$.

Since we assume $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k}) = p\widetilde{U}_d'$, $\widetilde{\widetilde{S}}(\widetilde{U}_d^{\otimes k})$ must return a pure state whenever the input state is a pure state. The instrument element $\widetilde{E}$ is an isometry, hence its output is always a pure state if the input state is pure. This forces $\widetilde{D}_S$ to preserve the purity of pure input states, which in turn implies that $\langle a|D\rho_{\mathcal{IA}}D^\dagger|a\rangle$ must be the same for all a up to a proportionality constant. Let $\widetilde{D}_a$ denote the map given by $\widetilde{D}_a(\rho_{\mathcal{IA}}) := D_a\rho_{\mathcal{IA}}D_a^\dagger = \langle a|D\rho_{\mathcal{IA}}D^\dagger|a\rangle$. The above argument shows that $\widetilde{D}_a \circ [\widetilde{U}_d^{\otimes k} \otimes \widetilde{I}_A] \circ \widetilde{E}$ is also a valid universal conjugation supermap.

Without loss of generality we assume that $k = d - 2$, since we may always opt to *not* use any of the input operations for the remaining cases of $k < d - 2$. The imaginary unit $\sqrt{-1}$ throughout this section will be denoted by the Italic font j.

---

[11] Reference [45] also proves that when $d > 2$, $k > 1$ uses of the input-unitary operation $\widetilde{U}_d$ are required for any non-null probabilistic heralded implementation.

By hypothesis, every pure state $|\psi\rangle \in \mathcal{I}_0 \cong \mathbb{C}^d$ and unitary operator $U_d \in L(\mathbb{C})$ must respect

$$D_a \left[U_d^{\otimes d-2} \otimes I \right] E|\psi\rangle = e^{j\phi_{\psi,U_d}} \sqrt{p} U_d^*|\psi\rangle, \tag{32}$$

where ϕ_{ψ,U_d} is a global phase that may depend on $|\psi\rangle$ and U_d. We see, however, that ϕ_{ψ,U_d} must be independent of the input state. Set $\{|i\rangle\}_{i=0}^{d-1}$ as the computational basis for $\mathbb{C}^d$ and the phase ϕ_{i,U_d} for when the input-state $|\psi\rangle$ is equal to $|i\rangle$. Take a maximally entangled state $|\phi_d^+\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle|i\rangle$ in $\mathcal{I}_0 \otimes \mathcal{I}_R$, where $\mathcal{I}_R$ is a "copy" of $\mathcal{I}_0$, i.e., another d-dimensional quantum system left untouched by $\widetilde{\widetilde{S}}$. We denote the corresponding phase by $\phi_{\phi_d^+,U_d}$. Let $M_U := D_a[U_d^{\otimes d-2} \otimes I]E$. Then, by linearity of M_U and Eq. (32), we conclude that $\phi_{i,U_d} = \phi_{\phi_d^+,U_d}$, hence no dependence on i. The subscript of ϕ_{ψ,U_d} for the input state shall be omitted as ϕ_{U_d}.

We now parametrize the operators E and D_a via their action on this basis as

$$E|i\rangle_{\mathcal{I}_0} = \sum_{\vec{i},i,a} \alpha_{\vec{i},i,a}|i_1, \ldots, i_{d-2}\rangle_{\mathcal{I}} \otimes |a\rangle_{\mathcal{A}},$$
$$_{\mathcal{O}_0}\langle i|D_a = \sum_{\vec{i},i,a} \beta_{\vec{i},i,a}|i_1, \ldots, i_{d-2}\rangle_{\mathcal{I}} \otimes |a\rangle_{\mathcal{A}}, \tag{33}$$

where $\vec{i} = [i_1, \ldots, i_{d-2}]$ is a vector such that $i_\lambda \in \{0, \ldots, d-1\}$ for any $\lambda = 1, \ldots, d - 2$. Hereafter, we restrict to unitary operators U_d that are diagonal in the computational basis such that $U_d = \sum_i e^{j\theta_i}|i\rangle\langle i|$ where θ_i is any real number. For such diagonal U_d its complex conjugate can be written as $U_d^* = \sum_i e^{-j\theta_i}|i\rangle\langle i|$. By Eq. (32),

$$\langle i'|D_a\left[U_d^{\otimes d-2} \otimes I\right]E|i'\rangle = e^{j\phi_{U_d}}\sqrt{p}\langle i'|U_d^*|i'\rangle. \tag{34}$$

Substituting the definition (33), we obtain

$$\sum_{\vec{i},i,a} \alpha_{\vec{i},i,a}\beta_{\vec{i},i,a} \, e^{j[\sum_{\lambda=1}^{d-2} \theta_{i_\lambda}]} = e^{j\phi_{U_d}}\sqrt{p}e^{-j\theta_{i'}} \tag{35}$$

or, equivalently,

$$\sum_{\vec{i},i,a} \alpha_{\vec{i},i,a}\beta_{\vec{i},i,a} \, e^{j[\theta_{i'}+\sum_{\lambda=1}^{d-2} \theta_{i_\lambda}]} = e^{j\phi_{U_d}}\sqrt{p}, \tag{36}$$

for all $i, i' \in \{0, \ldots, d-1\}$ and the diagonal U_d. Note that each U_d corresponds to some choice of real numbers $\vec{\theta}_\lambda = [\theta_0, \theta_1, \ldots, \theta_{d-1}]$ and vice versa. Moreover, the left-hand side of Eq. (36) depends on i', but the right-hand side does not.

In combinatorics, a *weak composition* of an integer n is a sequence of non-negative integers that sum to n. The weak compositions that appear in this proof are that of $d - 1$ with d elements. The set of all such weak compositions will be denoted by Γ and its elements (i.e., the individual weak decomposition) by $\vec{\gamma} = [\gamma_0, \ldots, \gamma_{d-1}]$, where the subscripts denote the elements of $\vec{\gamma}$.

In Eq. (36) the summation on i ranges between 0 and $d-1$ and $\vec{i}$ over all possible combinations of $\vec{i} = [i_1, \ldots, i_{d-2}]$ where each i_n ranges between 0 and $d-1$. Let ν_l denote the number of times an integer l between 0 and $d - 1$ appears in $\vec{i}$ and i. Recall that $\vec{i}$ consists of $d - 2$ variables, thus $\vec{i}$

and i in total are $d-1$ variables. We see that $\sum_{l=0}^{d-1} \nu_l = d-1$. Clearly the sequence $[\nu_0, \ldots, \nu_{d-1}]$ belongs to Γ. With slight abuse of notation, let us set $[\vec{i}, i] = [i_1, \ldots, i_{d-2}, i]$. Each $[\vec{i}, i]$ corresponds to $\vec{\gamma} \in \Gamma$. Each $[\vec{i}, i]$ with a given $\vec{\gamma}$ can be differentiated by an additional parameter, say κ. More specifically, let $K(\vec{\gamma})$ denote the set of all sequences $[\vec{i}, i]$ with the weak decomposition $\vec{\gamma}$. This extra parameter κ is then a natural number that enumerates the sequences in $K(\vec{\gamma})$ (e.g., via lexicographic ordering). Thus the summation $\sum_{\vec{i}, i, a}$ in Eq. (36) can be relabeled as $\sum_{\vec{\gamma}, \kappa, a}$. Introducing $\alpha'_{\vec{\gamma}} := \sum_{\kappa, a} \alpha_{\vec{\gamma}, \kappa, a} \beta_{\vec{\gamma}, \kappa, a}$, we have

$$\sum_{\vec{\gamma}} \alpha'_{\vec{\gamma}} \, e^{j[\sum_{l=0}^{d-1} \gamma_l \theta_l]} = e^{j\phi_{U_d}} \sqrt{p}. \tag{37}$$

Observe that for different $\vec{\gamma}$, the functions $e^{j[\sum_{l=0}^{d-1} \gamma_l \theta_l]}$ are linearly independent since θ_{i_λ} may take any value in the set of the real numbers. Each $\vec{\gamma}$ contains d elements and must sum up to $d-1$. One of the elements, say $\gamma_{l'}$, must be zero because the elements are non-negative. Set $i' = l'$ in Eq. (35) and use Eq. (37) to replace $e^{j\phi_{U_d}} \sqrt{p}$ on the left-hand side of Eq. (35). Then all the terms that appear on the right-hand side contain an exponent with a non zero coefficient in front of $\theta_{l'}$, while the coefficients of $\theta_{l'}$ are zero on the left-hand side. This equation can only be satisfied by setting $\alpha'_{\vec{\gamma}} = 0$, because $\exp(jk\theta)$ and $\exp(jk'\theta)$ are linearly independent functions of θ, for any pair of distinct integers k and k'. Thus, $p = 0$. ∎

VI. UNIVERSAL UNITARY TRANSPOSITION

This section addresses the problem of universal unitary transposition. We consider probabilistic heralded exact universal quantum protocols transforming k uses of a general d-dimensional unitary operation $\widetilde{U_d}$ into its transpose $\widetilde{U_d}$ in terms of a fixed basis. When only parallel protocols are considered, we show that the maximal success probability is exactly $p_s = 1 - \frac{d^2-1}{k+d^2-1}$. Also, by exploiting ideas of the port-based teleportation [55], one can design a delayed input-state parallel circuit that attains this maximal probability. When adaptive quantum circuits are considered, we present an explicit protocol that attains a success probability of $p_s = 1 - (1 - \frac{1}{d^2})^{\lceil \frac{k}{d} \rceil}$, which, for any constant dimension d, has an exponential improvement over any parallel protocol. We then analyze quantum protocols with indefinite causal order via the SDP approach presented in Sec. III and show that indefinite causal order protocols do have an advantage over causally ordered ones.

A. Gate teleportation and single-use unitary transposition

Quantum teleportation is a universal protocol that can be used to send an arbitrary d-dimensional quantum state via classical communication assisted by quantum entanglement. We are going to describe the protocol for pure states, as the extension to general mixed states follows from linearity. Suppose Alice holds the qudit state $|\psi\rangle \in \mathbb{C}^d$ and shares with Bob a d-dimensional maximally entangled state $|\phi_d^+\rangle := \sum_{i=0}^{d-1} \frac{1}{\sqrt{d}} |ii\rangle$. In order to "teleport" her state to Bob, Alice performs a general Bell measurement on $|\psi\rangle$ and her share
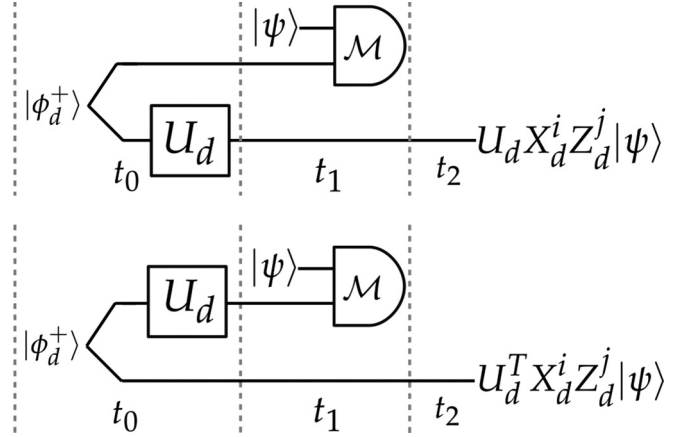

FIG. 5. Illustration of gate teleportation (upper circuit) and unitary transposition protocol (lower circuit).

of the entangled state and then sends the outcome of her measurement to Bob. The generalized Bell measurements have POVM elements given by

$$\mathcal{M} := \left\{ \left[(X_d^i Z_d^j)^\dagger \otimes I_d \right] |\phi_d^+\rangle\langle\phi_d^+| \left[(X_d^i Z_d^j) \otimes I_d \right] \right\}_{i,j=0}^{i,j=d-1}, \tag{38}$$

where

$$\begin{aligned} X_d^i &:= \sum_{l=0}^{d-1} |l \oplus i\rangle\langle l|, \\ Z_d^j &:= \sum_{l=0}^{d-1} \omega^{jl} |l\rangle\langle l|, \end{aligned} \tag{39}$$

$\omega := e^{\frac{2\pi\sqrt{-1}}{d}}$, and $l \oplus i$ denotes $l + i$ modulo d. The operators X_d^i and Z_d^j are known as the shift and clock operators, respectively, and can be seen as a generalization of the qubit Pauli operators. Straightforward calculation shows that, after Alice's measurement, the state held by Bob is given by $X_d^i Z_d^j |\psi\rangle$.

After the measurement process is complete, Alice sends the measurement outcomes i and j of her joint measurement to Bob. Bob can then apply the unitary operation $(Z_d^j)^{-1}(X_d^i)^{-1}$ on his state to recover the state $|\psi\rangle$. Remark that, with probability $p = \frac{1}{d^2}$ Alice obtains the outcomes $i = j = 0$ and Bob does not need to perform any correction.

The standard teleportation protocol can be adapted to teleport the use of a unitary operation in a process known as gate teleportation [56]. The idea here is that if Bob performs a unitary operation U_d on his half of the maximally entangled state before Alice performs the joint Bell measurement, the final state is given by $U_d X_d^i Z_d^j |\psi\rangle$, see Fig. 5. In this protocol, the operation U_d performed by Bob acts on the state $|\psi\rangle$ held by Alice when the outcomes are $i = j = 0$, which happens with probability $p = \frac{1}{d^2}$. Gate teleportation can be represented as a quantum circuit (see Fig. 5) and has applications in fault tolerant quantum computation [56].

Our method to transform a single use of a general d-dimensional unitary operation $\widetilde{U}_d$ into its transpose[12] $\widetilde{U}_d^T$ is based on the circuit interpretation of gate teleportation. The maximally entangled state respects the property $I \otimes A|\phi_d^+\rangle = A^T \otimes I|\phi_d^+\rangle$ for any linear operator $A \in L(\mathbb{C}^d)$. If Alice performs a general unitary U_d on her half of the maximally entangled state, the state held by Bob after the protocol is $U_d^T X_d^i Z_d^j |\psi\rangle$. With probability $p = \frac{1}{d^2}$, the outcome $i = j = 0$ is obtained and $U_d^T X_d^i Z_d^j |\psi\rangle$ is equal to $U_d^T |\psi\rangle$, see Fig. 5.

B. Port-based teleportation and parallel unitary transposition

Port-based teleportation [55] has the same main goal as the standard state teleportation protocol. Alice wants to "teleport" an arbitrary d-dimensional state $|\psi\rangle$ to Bob with classical communication assisted by shared entanglement. The original motivation of port-based teleportation is to perform a teleportation protocol that does not require a correction made via Pauli operators, but it can be made simply by selecting some particular "port." For that, it allows a more general initial resource state and more general joint measurements. The three main differences of port-based teleportation when compared to the standard teleportation protocol presented in the previous section can be summarized by:

(1) In port-based teleportation, instead of sharing a d-dimensional maximally entangled state, Alice and Bob may share a general d^k-dimensional entangled state $|\phi\rangle \in (\mathbb{C}^d \otimes \mathbb{C}^d)^{\otimes k}$. This general entangled state $|\phi\rangle$ can be seen as k pairs of qudits, referred to as ports.

(2) Instead of performing a generalized Bell measurement, Alice can perform a general joint measurement on $|\psi\rangle$ and her half of the k entangled states shared with Bob.

(3) Instead of performing the Pauli correction, Bob chooses a particular port based on Alice's message and discards the rest of the ports of his system.

We note that since no Pauli correction is made, port-based teleportation can only perform the teleportation task approximately or probabilistically. In this paper we only consider the probabilistic exact port-based teleportation where Alice performs a $k + 1$ outcome measurement, where k outcomes are associated with the k ports she shares with Bob and another outcome corresponding to failure. If Alice obtains the outcome of failure, she sends the failure flag to Bob and the protocol is aborted. If she obtains an outcome corresponding to some port l, she communicates this corresponding outcome to Bob and the state $|\psi\rangle$ is teleported to Bob's port labeled by l.

The optimal probabilistic single port ($k = 1$) case is essentially the standard state teleportation. Consider the case where Alice and Bob share the d-dimensional maximally entangled state $|\phi_d^+\rangle$. If we set the measurement performed by Alice as $M_1 = |\phi_d\rangle\langle\phi_d|$ and $M_{\text{fail}} = I - |\phi_d\rangle\langle\phi_d|$, with probability $p = \frac{1}{d^2}$, the state $|\psi\rangle$ is obtained in the single port 1, and with probability $p_F = 1 - \frac{1}{d^2}$ the protocol fails.

Reference [57] shows that the optimal probabilistic port-based gate teleportation protocol for any dimension d and number of states k with success probability $p = 1 - \frac{d^2-1}{k+d^2-1}$. Reference [57] also characterizes the optimal d^k-dimensional shared entangled state and the optimal joint measurement Alice must perform. The optimal state resource state is described by exploiting the Schur-Weyl duality

$$\mathbb{C}^{d \otimes k} \cong \bigoplus_{\mu \in \text{irrep}(U^{\otimes k})} \mathbb{C}_\mu^{\dim(\mu)} \otimes \mathbb{C}^{m_\mu}, \tag{40}$$

where $\text{irrep}(U^{\otimes k})$ is the set of all irreducible representations μ of the group of special unitary $\text{SU}(U_d)$ contained in the decomposition $U^{\otimes k}$ and m_μ is the multiplicity of the representation μ. The optimal resource state used for port-based teleportation can be written as

$$|\phi_{\text{PBT}}\rangle := \bigoplus_{\mu \in \text{irrep}(U^{\otimes k})} \sqrt{p_\mu}|\phi^+(\mu)\rangle \otimes |\psi_{m_\mu}\rangle, \tag{41}$$

where

$$|\phi^+(\mu)\rangle := \frac{1}{\sqrt{\dim(\mu)}} \sum_i |i_\mu i_\mu\rangle \in \mathbb{C}_\mu^{\dim(\mu)} \otimes \mathbb{C}_\mu^{\dim(\mu)} \tag{42}$$

is the maximally entangled state on the linear space of the irreducible representation μ, $\{p_\mu\}$ is a probability distribution, and $|\psi_{m_\mu}\rangle \in \mathbb{C}^{m(\mu)} \otimes \mathbb{C}^{m(\mu)}$ is a pure quantum state.

In Sec. VI A we have exploited the standard state gate teleportation to construct a protocol that can be used to transform a general unitary U_d into its transpose U_d^T. We now exploit port-based gate teleportation to construct a parallel protocol that transforms k uses of U_d to obtain its transpose.

The first important observation is that the state $|\phi_{\text{PBT}}\rangle$ [Eq. (41)] respects

$$U_d^{\otimes k} \otimes I|\phi_{\text{PBT}}\rangle = I \otimes U_d^{T \otimes k}|\phi_{\text{PBT}}\rangle. \tag{43}$$

This identity holds true because every tensor product of k unitaries U_d can be decomposed as[13]

$$U_d^{\otimes k} \cong \bigoplus_{\mu \in \text{irrep}(U^{\otimes k})} U(\mu) \otimes I_{m_\mu} \tag{44}$$

for some unitaries $U(\mu)$ acting on the irreducible representation space $\mathbb{C}_j^{\dim(j)}$ [57]. Hence, similarly to the case of the single use unitary transposition, we can adapt port-based gate teleportation to obtain a general protocol to transform k uses of a general unitary operation $\widetilde{U}_d$ into its transpose $\widetilde{U}_d^T$. It is enough to perform the operation $\widetilde{U}_d$ on each of her half of entangled qudit states (see Fig. 6). We will show in Sec. VI D that this protocol is also optimal in terms of success probability.

C. Review on probabilistic exact unitary learning

We make a brief summary of a problem known as unitary learning (also known as storage and retrieval of unitary operations) [9,21–24]. As we will show in Sec. VI D, the problem

---

[12] The transposition is taken in the computational basis $\{|i\rangle\}_{i=0}^{d-1}$ in which the maximally entangled state $|\phi_d^+\rangle = \sum_i \frac{1}{\sqrt{d}}|ii\rangle$ is defined.

[13] Here the symbol $\cong$ is used to emphasize that Eq. (44) is true up to an isometry.
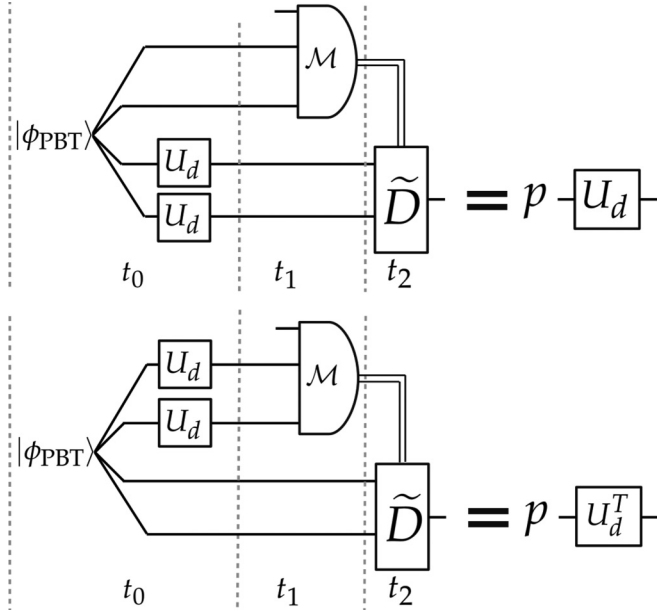
FIG. 6. Illustration of the modified port-based teleportation protocol that makes $k = 2$ uses of an arbitrary d-dimensional unitary operation $\widetilde{U}_d$ where the state $|\psi_{\text{PBT}}\rangle$ is described in Eq. (41) and the decoder $\widetilde{D}$ simply selects a particular port accordingly to the outcome of the joint measurement $\mathcal{M}$. The upper circuit exploits port-based gate teleportation to store $k = 2$ uses of a input-operation $\widetilde{U}_d$ and returns a single use of it with probability p. The lower circuit exploits port-based gate teleportation to transform k uses of $\widetilde{U}_d$ into a single use of its transpose $\widetilde{U}_d^T$. The upper and lower circuits are successful with probability $p = 1 - \frac{d^2-1}{k+d^2-1}$.

of probabilistic unitary learning is closely connected to the problem of parallel unitary transposition and results related to unitary learning will be useful to prove the optimality of our parallel unitary transposition protocol. Suppose that, until some time t_1, Alice has access to k uses of some general d-dimensional unitary operation U_d of which the description is not provided. In a later moment t_2, where Alice cannot access U_d any more, she wants to implement the action of this unitary on some general quantum state ρ chosen at time t_2. A parallel strategy[14] to succeed in this task is to perform the k uses of U_d on parts of an entangled quantum state ϕ_E before t_1 to obtain a quantum state $\psi_M := [U_d^{\otimes k} \otimes I]\phi_E[U_d^{\dagger \otimes k} \otimes I]$. Alice then saves this state ψ_M until a later time t_2 where she performs a global decoder operation $\widetilde{D}$ on the state ψ_M together with the target state ρ, which is desired to

---

[14]In principle, one may also consider adaptive protocols to perform better in the unitary learning problem. In an adaptive protocol, one can perform different encoder operations in between the use of the unitary to create more general protocols. One may also consider protocols where the unitaries U_d are used without a definite causal order. References [9,24] show that, for the unitary learning problem, the protocol with highest success probability (exact implementation) and highest expected fidelity (deterministic implementation) can always be parallelized.

satisfy[15] $\widetilde{D}(\psi_M \otimes \rho) = U_d \rho U_d^{\dagger}$. References [9,21–23] consider deterministic nonexact unitary learning protocols and analyze strategies that simulate the action of $\widetilde{U}_d$ with the maximal average fidelity, while Ref. [24] considers probabilistic heralded protocols that can be used to retrieve (a single use of) $\widetilde{U}_d$ exactly but may fail with some probability.

The unitary learning problem described above can be rephrased as the problem of finding delayed input-state protocols that transform k uses of a general unitary operation $\widetilde{U}_d$ into itself. In Sec. VI D we present a one-to-one connection between probabilistic unitary learning protocols and delayed input-state parallel protocols transforming k uses of a general unitary operation $\widetilde{U}_d$ into its transpose $\widetilde{U}_d^T$. Essentially, we show that any probabilistic unitary learning with success probability p can be translated into a parallel unitary transposition protocol with success probability p. This one-to-one connection is related to the fact that the optimal resource state used for unitary learning and the optimal resource state used for parallel delayed input-state unitary transposition can be both chosen as a state $|\phi\rangle$ which respects the property

$$U_d^{\otimes k} \otimes I|\phi\rangle = I \otimes U_d^{T \otimes k}|\phi\rangle, \tag{45}$$

as shown in the next subsection.

D. Optimal parallel unitary transposition protocols

We show how any parallel protocol that can be used to transform k copies of a general unitary operation $\widetilde{U}_d$ into its transpose $\widetilde{U}_d^T$ can be adapted into a delayed input-state protocol keeping the same success probability.

Lemma 2. Any parallel probabilistic heralded protocol transforming k copies of a general unitary $\widetilde{U}_d$ into $\widetilde{U}_d^T$ with a constant probability p can be converted to a delayed input-state parallel protocol with the same probability p.

Proof. Let S be the Choi operator of the superinstrument element associated with success and F be the Choi operator of the superinstrument element associated with failure. Superinstrument element S transforms k copies of $\widetilde{U}_d$ into $\widetilde{U}_d^T$ with probability p, i.e.,

$$\text{Tr}_{\mathcal{IO}}\big(S[I_{\mathcal{I}_0} \otimes \mathfrak{C}\big(\widetilde{U}_d^{\otimes k}\big)^T \otimes I_{\mathcal{O}_0}]\big) = p\mathfrak{C}\big(\widetilde{U}_d^T\big) \quad \forall U_d, \tag{46}$$

and $S + F$ is a valid parallel superchannel.

Since S transforms every unitary operator into its transpose, we can make the change of variable $U_d \mapsto B U_d A^T$ where A and B are arbitrary d-dimensional unitary operators. With that, unitary transposition can be seen as $(B U_d A^T)^{\otimes k} \mapsto p(B U_d A^T)^T = p A U_d^T B^T$. Our goal now is to show that if S respects Eq. (46), any operator S' respecting

$$S' = \big[A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{*\otimes k} \otimes A_{\mathcal{O}}^{*\otimes k} \otimes B_{\mathcal{O}_0}\big] S$$
$$\times \big[A_{\mathcal{I}_0}^{\dagger} \otimes B_{\mathcal{I}}^{T\otimes k} \otimes A_{\mathcal{O}}^{T\otimes k} \otimes B_{\mathcal{O}_0}^{\dagger}\big] \tag{47}$$

---

[15]We note that although the main goal is to obtain a decoder channel $\widetilde{D}$ and entangled state ϕ_E such that $\widetilde{D}(\psi_M \otimes \rho) = U_d \rho U_d^{\dagger}$ where $\psi_M := [U_d^{\otimes k} \otimes I]\phi_E[U_d^{\dagger \otimes k} \otimes I]$, the unitary learning task cannot be realized in a deterministic and exact way for a general unitary U_d.

satisfies

$$\text{Tr}_{\mathcal{IO}}\big(S'\big[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_d^{\otimes k}})^T \otimes I_{\mathcal{O}_0}\big]\big) = p\mathfrak{C}(\widetilde{U_d^T}) \quad \forall U_d. \quad (48)$$

To prove this fact, first note that the identity presented in Eq. (3) implies that $\mathfrak{C}(\widetilde{AU_d^T B^T}) = [A \otimes B]\,\mathfrak{C}(\widetilde{U_d^T})[A^\dagger \otimes B^\dagger]$, and

$$\mathfrak{C}\big(\big[\widetilde{BU_d A^T}\big]^{\otimes k}\big) = [B^{\otimes k} \otimes A^{\otimes k}]\mathfrak{C}(\widetilde{U_d^{\otimes k}})[B^{\dagger \otimes k} \otimes A^{\dagger \otimes k}], \quad (49)$$

which implies

$$\mathfrak{C}\big(\big[\widetilde{BU_d A^T}\big]^{\otimes k}\big)^T = [B^{*\otimes k} \otimes A^{*\otimes k}]\mathfrak{C}(\widetilde{U_d^{\otimes k}})^T [B^{T\otimes k} \otimes A^{T\otimes k}]. \quad (50)$$

Substituting Eq. (50) and $\mathfrak{C}(\widetilde{AU_d^T B^T}) = [A \otimes B]\,\mathfrak{C}(\widetilde{U_d^T})$ $[A^\dagger \otimes B^\dagger]$ in Eq. (46) we obtain

$$\text{Tr}_{\mathcal{IO}}\big(S\big\{I_{\mathcal{I}_0} \otimes [B^{*\otimes k} \otimes A^{*\otimes k}]\mathfrak{C}(\widetilde{U_d^{\otimes k}})^T[B^{T\otimes k} \otimes A^{T\otimes k}] \otimes I_{\mathcal{O}_0}\big\}\big)$$
$$= \big[A_{\mathcal{I}_0} \otimes B_{\mathcal{O}_0}\big]\mathfrak{C}(\widetilde{U_d^T})\big[A_{\mathcal{I}_0}^\dagger \otimes B_{\mathcal{O}_0}^\dagger\big]. \quad (51)$$

If we apply the operator $A_{\mathcal{I}_0}^\dagger \otimes B_{\mathcal{O}_0}^\dagger$ on the left side and the operator $A_{\mathcal{I}_0} \otimes B_{\mathcal{O}_0}$ on the right side of Eq. (51) and use the cyclic property of the trace, we find that S can be substituted by

$$S'' := \big[A_{\mathcal{I}_0}^\dagger \otimes B_{\mathcal{I}}^{T\otimes k} \otimes A_{\mathcal{O}}^{T\otimes k} \otimes B_{\mathcal{O}_0}^\dagger\big] S$$
$$\times \big[A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{*\otimes k} \otimes A_{\mathcal{O}}^{*\otimes k} \otimes B_{\mathcal{O}_0}\big]. \quad (52)$$

Since A and B are arbitrary unitary operators, we can take the invertible transformations $A^\dagger \mapsto A$ and $B^\dagger \mapsto B$ to obtain the symmetry of Eq. (47).

The symmetry presented in Eq. (47) motivates the definition of a Haar measure "twirled" map

$$\widetilde{\tau}(S) := \int_{\text{Haar}} \big[A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{*\otimes k} \otimes A_{\mathcal{O}}^{*\otimes k} \otimes B_{\mathcal{O}_0}\big]S$$
$$\times \big[A_{\mathcal{I}_0}^\dagger \otimes B_{\mathcal{I}}^{T\otimes k} \otimes A_{\mathcal{O}}^{T\otimes k} \otimes B_{\mathcal{O}_0}^\dagger\big]dAdB. \quad (53)$$

We now define a twirled version of the superinstrument as $S_\tau := \widetilde{\tau}(S)$ and $F_\tau := \widetilde{\tau}(F)$, which respects the conditions of valid superinstruments and S_τ also transforms k uses of any $\widetilde{U_d}$ into $\widetilde{U_d^T}$ with probability p. We now notice that both S_τ and F_τ respect

$$\text{Tr}_{\mathcal{O}\mathcal{O}_0}S_\tau = \int_{\text{Haar}} \big[A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{*\otimes k}\big]\text{Tr}_{\mathcal{O}\mathcal{O}_0}(S)\big[A_{\mathcal{I}_0}^\dagger \otimes B_{\mathcal{I}}^{T\otimes k}\big]dAdB$$
$$\propto I_{\mathcal{I}_0} \otimes \text{Tr}_{\mathcal{I}_0\mathcal{O}\mathcal{O}_0}S_\tau \quad (54)$$

since the identity is the only operator that commutes with all unitary operations (Schur's lemma). It follows then that the superchannel $C_\tau := S_\tau + F_\tau$ respects the conditions of a parallel delayed input-state protocol.

Lemma 3. For every delayed input-state parallel protocol transforming k uses of a general unitary operation $\widetilde{U_d}$ into its transpose $\widetilde{U_d^T}$ with success probability p that is independent of $\widetilde{U_d}$, there exists a probabilistic unitary learning protocol with a success probability p.

Conversely, for every probabilistic unitary learning protocol with a success with probability p that is independent of $\widetilde{U_d}$, there exists a delayed input-state parallel protocol

transforming k uses of a general unitary operation $\widetilde{U_d}$ into its transpose $\widetilde{U_d^T}$ with a constant success probability p.

Proof. We start by showing how one can adapt a parallel protocol transforming k uses of a general unitary operation $\widetilde{U_d}$ into its transpose $\widetilde{U_d^T}$ into a unitary learning one with the same success probability.

Let S be the Choi operator of the superinstrument element associated with success and F be the Choi operator of the superinstrument element associated with failure. Superinstrument element S transforms k copies of $\widetilde{U_d}$ into $\widetilde{U_d^T}$ with probability p, i.e.,

$$\text{Tr}_{\mathcal{IO}}\big(S\big[I_{\mathcal{I}_0} \otimes \mathfrak{C}(\widetilde{U_d^{\otimes k}})^T \otimes I_{\mathcal{O}_0}\big]\big) = p\mathfrak{C}(\widetilde{U_d^T}) \quad \forall U_d, \quad (55)$$

Lemma 2 states that this protocol can be converted to have a delayed input state and without lost of generality, the superchannel $C = S + F$ respects the commutation relation

$$\big[C, A_{\mathcal{I}_0}^* \otimes B_{\mathcal{I}}^{\otimes k} \otimes A_{\mathcal{O}}^{\otimes k} \otimes B_{\mathcal{O}_0}^*\big] = 0 \quad (56)$$

for every pair of unitary operations $A, B \in SU(d)$.

When a Choi operator C represents a delayed input-state protocol, the operator $C_{\mathcal{I}} := \text{Tr}_{\mathcal{I}_0\mathcal{O}\mathcal{O}_0}C$ is proportional to the reduced state $\text{Tr}_{\mathcal{A}}(\phi_E)$ of the state $\phi_E \in L(\mathcal{I} \otimes \mathcal{A})$ prepared by Alice before the use of the input operations.[16] From the commutation relation in Eq. (56), we see that $C_{\mathcal{I}}$ respects

$$\big[C_{\mathcal{I}}, B_{\mathcal{I}}^{\otimes k}\big] = 0. \quad (57)$$

The Schur-Weyl duality states that k identical d-dimensional unitaries B can be decomposed as (see Sec. VI B)

$$B^{\otimes k} \cong \bigoplus_{\mu \in \text{irrep}(U_d^{\otimes k})} B(\mu) \otimes I_{m(\mu)}, \quad (58)$$

where $B(\mu) \in L(\mathbb{C}_\mu^{\dim(\mu)})$ is a unitary operator, and $I_{m(\mu)}$ is the identity on the multiplicity space $\mathbb{C}^{m(\mu)}$. Since the reduced state $\text{Tr}_{\mathcal{A}}(\phi_E)$ respects the relation $[\text{Tr}_{\mathcal{A}}(\phi_E), B^{\otimes k}] = 0$, Schur's lemma ensures that the reduced encoder state has the form of

$$\text{Tr}_{\mathcal{A}}(\phi_E) \propto \bigoplus_\mu I_\mu \otimes \rho_{m_\mu}, \quad (59)$$

where I_μ is the identity on the the linear space $\mathbb{C}_\mu^{\dim(\mu)}$ and ρ_{m_μ} is some state on the multiplicity space of μ. Without loss of generality, we can assume that $\phi_E = |\phi_E\rangle\langle\phi_E|$ is a pure state with a reduced state that respects Eq. (59). It follows then that $|\phi_E\rangle$ can be written as

$$|\phi_E\rangle := \bigoplus_{\mu \in \text{irrep}(U^{\otimes k})} \sqrt{p_\mu}|\phi^+(\mu)\rangle \otimes |\psi_{m_\mu}\rangle, \quad (60)$$

where

$$|\phi^+(\mu)\rangle := \frac{1}{\sqrt{\dim(\mu)}}\sum_i |i_\mu i_\mu\rangle \in \mathbb{C}_\mu^{\dim(\mu)} \otimes \mathbb{C}_\mu^{\dim(\mu)} \quad (61)$$

---

[16]See Fig. 4 for a pictorial illustration for the case $k = 1$. Let $\phi_E \in L(\mathcal{H}_2 \otimes \mathcal{H}_A)$ be the state created by the encoder of the delayed input-state protocol of Fig. 4. In this case, $C_2 := \text{Tr}_{134}C$ is proportional to the reduced state $\text{Tr}_{\mathcal{A}}\phi_E$.

is the maximally entangled state on the linear space of the irreducible representation μ, $\{p_\mu\}$ is a probability distribution, and $|\psi_{m_\mu}\rangle \in \mathbb{C}^{m(\mu)} \otimes \mathbb{C}^{m(\mu)}$ are some purifications of ρ_{m_μ}.

We now make an important observation. Although the state $|\phi_E\rangle$ is not the maximally entangled state, it respects

$$U_d^{\otimes k} \otimes I|\phi_E\rangle = I \otimes U_d^{T\otimes k}|\phi_E\rangle. \tag{62}$$

This identity holds true because any tensor product of k identical unitaries U_d can be decomposed as

$$U_d^{\otimes k} \cong \bigoplus_{\mu \in \mathrm{irrep}(U^{\otimes k})} U(\mu) \otimes I_{m_\mu} \tag{63}$$

for some unitaries $U(\mu)$ acting on the invariant representation space $\mathbb{C}_j^{\dim(j)}$. Any delayed input-state protocol that can be used for unitary transposition can be used for unitary learning, since it is enough to perform the unitaries $U_d^{\otimes k}$ on the "other" half of the entangled state $|\phi_E\rangle$ on which the joint operation is not performed.

We now show how to transform probabilistic unitary learning protocols to heralded unitary transposition protocols. In Ref. [24], the authors have shown that, without loss of generality, any probabilistic unitary learning protocol can be made parallel and, moreover, with the entangled state $|\phi_E\rangle \in L(\mathcal{I} \otimes \mathcal{A})$ which respects the property

$$U_d^{\otimes k} \otimes I|\phi_E\rangle = I \otimes U_d^{T\otimes k}|\phi_E\rangle. \tag{64}$$

Hence, if we perform the unitary operations $U_d^{\otimes k}$ into the half of the entangled state $|\psi\rangle$ on which the joint measurement is performed, the unitary recovered after the learning protocol will be U_d^T instead of U_d.

We are now in position to prove that the protocol based on port-based gate teleportation presented in Sec. VI B is optimal.

Theorem 2 (Optimal parallel unitary transposition). The modified port-based gate teleportation protocol can be used to transform k uses of an arbitrary d-dimensional unitary operation $\widetilde{U}_d$ into its transpose $\widetilde{U}_d^T$ with success probability $p = 1 - \frac{d^2-1}{k+d^2-1}$ in a parallel delayed input-state protocol. Moreover, this protocol attains the optimal success probability among all parallel protocols with probability p that does not depend on $\widetilde{U}_d$.

Proof. As shown above, the identity $U^{\otimes k} \otimes I|\phi_{\mathrm{PBT}}\rangle = I \otimes U^{T\otimes k}|\phi_{\mathrm{PBT}}\rangle$ ensures that port-based gate teleportation can be used to construct a delayed input-state parallel protocol that obtains U_d^T with k uses of U_d with probability $p = 1 - \frac{d^2-1}{k+d^2-1}$.

Lemma 3 shows that any protocol transforming k uses of $\widetilde{U}_d$ into its transpose $\widetilde{U}_d^T$ in a parallel protocol with probability p can be used to successfully "learn" the input-operation $\widetilde{U}_d$ with probability p and k uses. Reference [24] shows that the optimal protocol for unitary learning a unitary U_d with k uses cannot have constant probability greater than $p = 1 - \frac{d^2-1}{k+d^2-1}$, which bounds our maximal probability of success and finishes the proof.
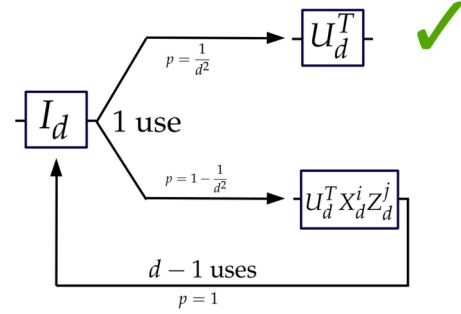

FIG. 7. A flowchart illustrating the adaptive unitary transpose protocol.

E. Adaptive unitary transposition protocols

In this subsection we present an adaptive circuit that transforms k uses of an arbitrary d-dimensional unitary operation $\widetilde{U}_d$ into a single use of its transpose $\widetilde{U}_d^T$ with heralded probability $p = 1 - (1 - \frac{1}{d^2})^{\lceil \frac{k}{d} \rceil}$ (see Fig. 7).

(1) We start by making a single use of the input-operation $\widetilde{U}_d$ to implement the probabilistic heralded transposition protocol based on gate teleportation described in Sec. VI D. When the generalized Bell measurement returns the outcomes i and j, the operator U_d is transformed into $V_1 = U_d^T X_d^i Z_d^j$, where X_d^i and Z_d^j are the clock and shift operators, respectively [see Eq. (39)].

(2) If both outcomes i and j correspond to the identity operator, i.e., $i = j = 0$, we have $V_1 = U_d^T$ and we stop the protocol with success. If some other outcome is obtained, we make $d - 1$ uses of U_d to implement the unitary complex conjugate protocol [25] to obtain U_d^*. We then apply $X_d^{i^{-1}} Z_d^{j^{-1}} U_d^*$ into V_1 to "cancel" the transformation of step 1 to obtain identity operator $[X_d^{i^{-1}} Z_d^{j^{-1}} U_d^* U_d^T Z_d^j X_d^i = I_d]$.

(3) Go to step 1.

We see that step 1 fails returning $\widetilde{U}_d^T$ with probability $(1 - \frac{1}{d^2})$ and we need in total d uses of the input-operation $\widetilde{U}_d$ to complete steps 1 and 2. These steps may be repeated up to $\lceil \frac{k}{d} \rceil$ times, hence they lead to a success probability of $p = 1 - (1 - \frac{1}{d^2})^{\lceil \frac{k}{d} \rceil}$.

F. Optimal protocols via SDP formulation and indefinite causal order advantage

We apply the SDP methods obtained in Sec. III to the case of unitary transposition and present the optimal success probability in Table I. By checking Table I one observes that the adaptive circuit we have presented in Sec. VI E is not optimal. One possible intuitive understanding is that the adaptive protocol we have presented in Sec. VI E "wastes" $d - 1$ uses of the input-operation $\widetilde{U}_d$ to recover the input state. We also notice that indefinite causal order protocols provide a strictly large success probability when compared to causally ordered ones. It is interesting to observe that although indefinite causal order protocols have been reported useful in tasks such as nonsignaling channel discrimination [58], quantum computation [32], and quantum channel capacity activation [27,28], in these previous examples, the advantage of

indefinite causal order is obtained by exploiting the *quantum switch* [10], a process which is not useful in our task of unitary channel transformation, since the quantum switch would transform k uses of the any unitary operation $\widetilde{U_d}$ into simple k concatenations of $\widetilde{U_d}$, or equivalently, a single use of $\widetilde{U_d^k}$. Our results for indefinite causal order then reveals the existence of a different class of indefinite causal order protocols, similarly to the one reported for unitary inverse in Ref. [29].

VII. UNIVERSAL UNITARY INVERSION PROTOCOLS

We now address the problem of transforming k uses of a general d-dimensional unitary operation $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ with probabilistic heralded quantum circuits. We have presented our adaptive circuit in Ref. [29] and here we present a parallel implementation and provide more details on the adaptive circuit.

Before presenting our protocols we prove that, similarly to the complex conjugation case, any protocol performing exact universal unitary inversion with $k < d - 1$ uses of the unitary input-operation $\widetilde{U_d}$ necessarily has null success probability. Also, this no-go result also holds even when protocols with indefinite causal order are considered.

Theorem 3 (Unitary inversion: no-go). Any universal probabilistic heralded quantum protocol (including protocols without definite causal order) transforming $k < d - 1$ uses of a d-dimension unitary operation $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ with success probability p that does not depend on U_d necessarily has $p = 0$, i.e., null success probability.

Proof. Assume that there exists a quantum protocol transforming k uses of a general d-dimensional unitary operation $\widetilde{U_d}$ into its inverse $\widetilde{U_d^{-1}}$ with a nonzero success probability p. We can then exploit the single-use unitary transposition protocol presented in Sec. VI A to obtain $\widetilde{U_d^*}$ with success probability $p/d^2 > 0$, which contradicts Lemma 1. $\blacksquare$

A. Parallel unitary inversion protocols

We start by showing that, similarly to universal parallel transposition, any universal parallel unitary inversion protocol can be made in a delayed input-state way.

Lemma 4. Any parallel probabilistic heralded parallel protocol k uses of a general unitary $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ with constant probability p can be conversed to a delayed input-state parallel protocol with the same probability p.

Proof. The proof follows the same steps as the one in Theorem 2. The only difference is that for unitary transposition, the superinstrument element S can be chosen as an operator that commutes with unitaries of the form $A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{*\otimes k} \otimes A_{\mathcal{O}}^{*\otimes k} \otimes B_{\mathcal{O}_0}$ and for unitary inversion S can be chosen as an operator which commutes with all unitaries of the form of $A_{\mathcal{I}_0} \otimes B_{\mathcal{I}}^{\otimes k} \otimes A_{\mathcal{O}}^{\otimes k} \otimes B_{\mathcal{O}_0}$.

We are now in condition to present a universal circuit for parallel unitary inversion and also to obtain an upper bound on the maximal success probability. Our protocol makes use of the unitary complex conjugation and unitary transposition and it is proven to be optimal for qubits.

Theorem 4 (Universal unitary inverse). There exists a parallel delayed input-state probabilistic quantum circuit that transforms k uses of an arbitrary d-dimensional unitary operation $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ with success probability $p_S = 1 - \frac{d^2-1}{k'+d^2-1}$ where $k' := \lfloor \frac{k}{d-1} \rfloor$ is the greatest integer that is less than or equal to $\frac{k}{d-1}$.

The maximal success probability transforming k uses of an arbitrary d-dimensional unitary operation $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ in a parallel quantum circuit is upper bounded by $p_{\max} \leqslant 1 - \frac{d^2-1}{k(d-1)+d^2-1}$.

Proof. We construct our protocol by concatenating the protocol for unitary complex conjugation of Ref. [25] with the unitary transposition one presented in Sec. VI B. First we divide the k uses of the input-operation $\widetilde{U_d}$ into $k' = \lfloor \frac{k}{d-1} \rfloor$ groups containing $d - 1$ uses of $\widetilde{U_d}$ and discard possible extra uses. We then exploit the unitary conjugation protocol to obtain k' uses of U_d^*. After, we implement the unitary transposition protocol of Sec. VI B on k' uses of $\widetilde{U_d^*}$ to obtain a single use of $\widetilde{U_d^{-1}}$ with probability of success given by $p = 1 - \frac{d^2-1}{k'+d^2-1}$.

Next, we prove the upper bound. Let $p_{\text{inv}}(d, k)$ be the success probability of transforming k uses of an arbitrary unitary input-operation $\widetilde{U_d}$ into a single use of its inverse $\widetilde{U_d^{-1}}$ with a parallel circuit. Suppose one has access to $l = k(d - 1)$ uses of an input-operation $\widetilde{U_d}$. One possible protocol to transform these l uses of $\widetilde{U_d}$ into its transpose with a parallel circuit is the following, first we perform the deterministic parallel complex conjugation protocol on l uses to obtain k uses of $\widetilde{U_d^*}$. We then perform the parallel unitary inversion on k uses of $\widetilde{U_d^*}$ to obtain $\widetilde{U_d^{*-1}} = \widetilde{U_d^T}$ with probability $p_{\text{inv}}(d, k)$. This parallel unitary transposition protocol has then success probability of $q_{\text{T}}(d, l) = p_{\text{inv}}(d, k)$. Theorem 2 states that any parallel circuit that transforms l uses of an arbitrary unitary into its transpose respects $q_{\text{T}}(d, l) \leqslant 1 - \frac{d^2-1}{l+d^2-1}$, which implies

$$p_{\text{inv}}(d, k) \leqslant 1 - \frac{d^2 - 1}{l + d^2 - 1} = 1 - \frac{d^2 - 1}{k(d-1) + d^2 - 1},$$
(65)

which completes the proof. $\blacksquare$

B. Adaptive unitary inversion circuit

For completeness we now summarize the protocol for adaptive unitary inversion presented in Ref. [29]. Our protocol to obtain $\widetilde{U_d^{-1}}$ follows similar steps of the protocol to implement $\widetilde{U_d^{-1}}$ presented in the previous section goes as follow (see Fig. 8).

(1) We start by making a $d - 1$ uses of the unitary U_d to implement the probabilistic heralded transposition protocol for unitary inverse used in Theorem 4 in the main text. When the generalized Bell measurements return the outcomes i and j, the operation U_d is transformed into $V_1 = U_d^{-1} X_d^i Z_d^j$, where X_d^i and Z_d^j are the clock and shift operators, respectively [see Eq. (39)].

(2) If both outcomes i and j correspond to the identity operator, i.e., $i = j = 0$, we have $V_1 = U_d^{-1}$ and we stop the
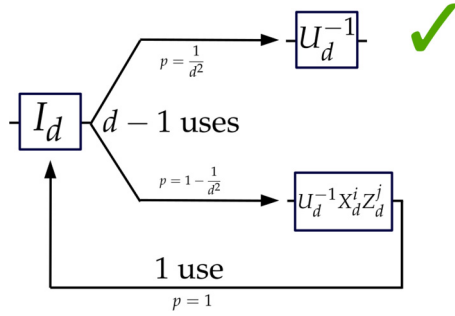
FIG. 8. Flowchart illustrating the adaptive unitary inverse protocol.

protocol with success. If some other outcome is obtained, we make a single use of $\widetilde{U_d}$ to apply $X_d^{i-1} Z_d^{j-1} U_d$ into V_1 and invert the transformation of step 1 to obtain identity operator $[X_d^{i-1} Z_d^{j-1} U_d U_d^{-1} Z_d^j X_d^i = I_d]$.

(3) Go to step 1.

We see that step 1 requires $d-1$ uses and returns $\widetilde{U_d}^{-1}$ with probability $(1 - \frac{1}{d^2})$. We need in total d uses of $\widetilde{U_d}$ to complete steps 1 and 2. Iteration of this protocol leads to a success probability of $p = 1 - (1 - \frac{1}{d^2})^{\lfloor \frac{k+1}{d} \rfloor}$.

C. Optimal protocols via SDP formulation and indefinite causal order advantage

We now apply the SDP methods of Sec. III to the case of unitary inversion and reproduce Table I of Ref. [29] in Table II. For qubits ($d = 2$) we note that with the Pauli qubit unitary operator Y, $Y U_2 Y = U_2^*$ for every U_2 with determinant one. Hence, any protocol for transforming a single use of a qubit unitary operation into a single use of its transposition can be converted into a qubit unitary inversion protocol and vice versa. Hence, the results and conclusions for qubits are equivalent to the ones presented in Sec. VI F.

For qutrits ($d = 3$), Theorem 3 ensures that circuits with a single use necessarily have success probability equal zero. For the case $d = 3$ and $k = 2$, parallel, adaptive, and indefinite causal order protocols have attained the same success

TABLE II. Maximum success probabilities for universally inverting k uses of U_d, by parallel quantum circuit, adaptive quantum circuit, and protocols with indefinite causal orders. Values in bold are analytical and the rest via numerical SDP optimization. This table is extracted from Table I of Ref. [29].

$d = 2$	Parallel	Adaptive	Indefinite causal order
$k = 1$	$\frac{1}{4} = \mathbf{0.25}$	$\frac{1}{4} = \mathbf{0.25}$	$\frac{1}{4} = \mathbf{0.25}$
$k = 2$	$\frac{2}{5} = \mathbf{0.4}$	$0.4286 \approx \frac{3}{7}$	$0.4444 \approx \frac{4}{9}$
$k = 3$	$\frac{1}{2} = \mathbf{0.5}$	$0.7500 \approx \frac{3}{4}$	0.9416
$d = 3$	Parallel	Adaptive	Indefinite causal order
$k = 1$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$k = 2$	$0.1111 \approx \frac{1}{9}$	$0.1111 \approx \frac{1}{9}$	$0.1111 \approx \frac{1}{9}$

probability, suggesting that our parallel unitary inversion protocol may be optimal when $d = k - 1$.

VIII. CONCLUSIONS

We have addressed the problem of designing probabilistic heralded universal quantum protocols that transform k uses of an arbitrary (possibly unknown) d-dimensional unitary quantum operation $\widetilde{U_d}$ to exactly implement a single use of some other operation given by $f(\widetilde{U_d})$. For the cases where f is a linear supermap, we have provided an SDP algorithm that can be used to analyze parallel, adaptive, and indefinite causal order protocols. For the parallel and adaptive cases, our algorithm finds a quantum circuit that universally implements the desired transformation with the optimal probability of success for any k and d. For the indefinite causal order case, the algorithm finds a quantum process that obtains the desired transformation with the optimal probability of success for any k and d.

For the particular case of unitary complex conjugation, i.e., $f(\widetilde{U_d}) = \widetilde{U_d^*}$, we have proved that when $k < d - 1$ the success probability is necessarily zero, even when indefinite causal order protocols are considered. Since a deterministic parallel quantum circuit to transform $k = d - 1$ uses of a general unitary operation $\widetilde{U_d}$ into a single use of its complex conjugation was presented in Ref. [25], we can argue that the theoretical possibility of implementing universal exact unitary complex conjugation is completely solved.

For the particular case of unitary transposition, i.e., $f(\widetilde{U_d}) = \widetilde{U_d^T}$, we have shown that the optimal success probability with parallel circuits is exactly $p = 1 - \frac{d^2-1}{k+d^2-1}$. When adaptive circuits are considered, we have presented an explicit protocol that has success probability $p = 1 - (1 - \frac{1}{d^2})^{\lceil \frac{k}{d} \rceil}$, which has an exponential improvement over any parallel protocol. We have also shown that indefinite causal order protocols outperforms causally ordered ones by tackling the cases $d = 2, k \leqslant 3$ and $d = 3, k = 2$ numerically.

For the particular case of unitary inversion, i.e., $f(\widetilde{U_d}) = \widetilde{U_d^{-1}}$, we have proved that when $k < d - 1$ the success probability is necessarily zero, even when indefinite causal order protocols are considered. When $k \geqslant d - 1$ we have presented parallel and adaptive circuits to succeed in this task and proved that the success probability of our adaptive protocol presented in Ref. [29] has probability of success given by $p = 1 - (1 - \frac{1}{d^2})^{\lfloor \frac{k+1}{d} \rfloor}$ and we prove it to be exponentially higher than any success probability obtained by parallel circuits.

[1] K. Kraus, in *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics, Vol. 190, edited by A. Böhm, J. D. Dollard, and W. H. Wootters (Springer-Verlag, Berlin, Heidelberg, 1983), 1st ed.

[2] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, 2000).

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

[4] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, Europhys. Lett. **83**, 30004 (2008).

[5] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Optimal Quantum Tomography of States, Measurements, and Transformations, Phys. Rev. Lett. **102**, 010404 (2009).

[6] G. Chiribella and D. Ebler, Quantum speedup in the identification of cause-effect relations, Nat. Commun. **10**, 1472 (2019).

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

[8] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, Quantum communication using a bounded-size quantum reference frame, New J. Phys. **11**, 063013 (2009).

[9] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Optimal quantum learning of a unitary transformation, Phys. Rev. A **81**, 032324 (2010).

[10] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).

[11] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).

[12] M. Wolf, Quantum Channels and Operations Guided Tour (Corolarry 6.2), http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf.

[13] D. Cariello, An Elementary Description of the Positive Maps with Positive Inverse, *CNMAC12*, http://sbmac.org.br/eventos/cnmac/xxxiv_cnmac/pdf/541.pdf.

[14] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature (London) **299**, 802 (1982).

[15] V. Bužek, M. Hillery, and R. F. Werner, Optimal manipulations with qubits: Universal-NOT gate, Phys. Rev. A **60**, R2626 (1999).

[16] A. Acín, Statistical Distinguishability between Unitary Operations, Phys. Rev. Lett. **87**, 177901 (2001).

[17] G. Mauro D'Ariano, P. L. Presti, and M. G. A. Paris, Improved discrimination of unitary transformations by entangled probes, J. Opt. B **4**, S273 (2002).

[18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Optimal Cloning of Unitary Transformation, Phys. Rev. Lett. **101**, 180504 (2008).

[19] I. S. B. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, Universal refocusing of systematic quantum noise, arXiv:1602.07963.

[20] M. Navascués, Resetting Uncontrolled Quantum Systems, Phys. Rev. X **8**, 031008 (2018).

[21] G. Vidal, L. Masanes, and J. I. Cirac, Storing Quantum Dynamics in Quantum States: A Stochastic Programmable Gate, Phys. Rev. Lett. **88**, 047905 (2002).

[22] M. Sasaki and A. Carlini, Quantum learning and universal quantum matching machine, Phys. Rev. A **66**, 022303 (2002).

[23] S. Gammelmark and K. Mølmer, Quantum learning by measurement and feedback, New J. Phys. **11**, 033017 (2009).

[24] M. Sedlák, A. Bisio, and M. Ziman, Optimal Probabilistic Storage and Retrieval of Unitary Channels, Phys. Rev. Lett. **122**, 170502 (2019).

[25] J. Miyazaki, A. Soeda, and M. Murao, Complex conjugation supermap of unitary quantum maps and its universal implementation protocol, Phys. Rev. Res. **1**, 013007 (2019).

[26] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New J. Phys. **17**, 102001 (2015).

[27] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, Phys. Rev. Lett. **120**, 120502 (2018).

[28] S. Salek, D. Ebler, and G. Chiribella, Quantum communication in a superposition of causal orders, arXiv:1809.06655.

[29] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Reversing Unknown Quantum Transformations: A Universal Quantum Circuit for Inverting General Unitary Operations, Phys. Rev. Lett. **123**, 210502 (2019).

[30] A. Bisio, G. M. D'Ariano, P. Perinotti, and M. Sedlák, Optimal processing of reversible quantum channels, Phys. Lett. A **378**, 1797 (2014).

[31] D. Trillo, B. Dive, and M. Navascues, Remote time manipulation, arXiv:1903.10568.

[32] M. Araújo, F. Costa, and Č. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, Phys. Rev. Lett. **113**, 250402 (2014).

[33] J. de Pillis, Linear transformations which preserve Hermitian and positive semidefinite operators, Pacific J. Math. **23**, 129 (1967).

[34] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Phys. **3**, 275 (1972).

[35] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).

[36] W. F. Stinespring, Positive functions on C*-algebras, Proc. Am. Math. Soc. **6**, 211 (1955).

[37] M. Neumark, Spectral functions of a symmetric operator, Izv. Akad. Nauk SSSR Ser. Mat. **4**, 277 (1940).

[38] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum networks: General theory and applications, Acta Phys. Slovaca **61**, 273 (2011).

[39] D. Kretschmann and R. F. Werner, Quantum channels with memory, Phys. Rev. A **72**, 062323 (2005).

[40] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, A purification postulate for quantum mechanics with indefinite causal order, Quantum **1**, 10 (2017).

[41] G. Chiribella and D. Ebler, Optimal quantum networks and one-shot entropies, New J. Phys. **18**, 093053 (2016).

[42] A. Chefles, R. Jozsa, and A. Winter, On the existence of physical transformations between sets of quantum states, Int. J. Quantum. Inform. **2**, 11 (2004).

[43] T. Heinosaari, M. A. Jivulescu, D. Reeb, and M. M. Wolf, Extending quantum operations, J. Math. Phys. **53**, 102208 (2012).

[44] Q. Dong, M. T. Quintino, A. Soeda, and M. Murao, Implementing positive maps with multiple copies of an input state, Phys. Rev. A **99**, 052352 (2019).

[45] J. Miyazaki, On the difference between positivity and complete positivity of maps in quantum theory, Ph.D. thesis, The University of Tokyo, 2017, https://irdb.nii.ac.jp/en/00926/0004117630.

[46] MATLAB. The MathWorks Inc., Natick, Massachusetts.

[47] http://cvxr.com/cvx.

[48] https://www.mosek.com/.

[49] http://sedumi.ie.lehigh.edu/.

[50] http://www.math.nus.edu.sg/~mattohkc/sdpt3.html.

[51] https://github.com/mtcq/unitary_inverse.

[52] https://opensource.org/licenses/MIT.

[53] N. Johnston, QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9, http://qetlab.com, Jan., 2016.

[54] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[55] S. Ishizaka and T. Hiroshima, Asymptotic Teleportation Scheme as a Universal Programmable Quantum Processor, Phys. Rev. Lett. **101**, 240501 (2008).

[56] D. Gottesman and I. L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, Nature (London) **402**, 390 (1999).

[57] M. Studziński, S. Strelchuk, M. Mozrzymas, and M. Horodecki, Port-based teleportation in arbitrary dimension, Sci. Rep. **7**, 10871 (2017).

[58] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301 (2012).