# Cloning of a quantum measurement

Alessandro Bisio,* Giacomo Mauro D'Ariano,† and Paolo Perinotti‡
*QUIT Group, Dipartimento di Fisica "A. Volta" and INFN, via Bassi 6, I-27100 Pavia, Italy*§

Michal Sedlák‖
*QUIT Group, Dipartimento di Fisica "A. Volta" via Bassi 6, I-27100 Pavia, Italy and*
*Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, SK-845 11 Bratislava, Slovakia*§

(Received 29 March 2011; published 18 October 2011)

We analyze quantum algorithms for cloning of a quantum measurement. Our aim is to mimic two uses of a device performing an unknown von Neumann measurement with a single use of the device. When the unknown device has to be used before the bipartite state to be measured is available we talk about $1 \rightarrow 2$ learning of the measurement, otherwise the task is called $1 \rightarrow 2$ cloning of a measurement. We perform the optimization for both learning and cloning for arbitrary dimension $d$ of the Hilbert space. For $1 \rightarrow 2$ cloning we also propose a simple quantum network that achieves the optimal fidelity. The optimal fidelity for $1 \rightarrow 2$ learning just slightly outperforms the estimate and prepare strategy in which one first estimates the unknown measurement and depending on the result suitably prepares the duplicate.

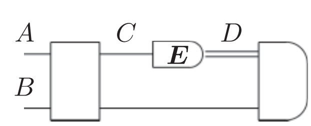
DOI: 10.1103/PhysRevA.84.042330 PACS number(s): 03.67.Ac, 03.65.Ta

## I. INTRODUCTION

Arbitrary processing of classical information can be described by strings of bits, and can be performed by a fixed device, for example, a processor of any personal computer. As a consequence, we do not need to build new devices for different computations, but we just need to copy bit strings carrying the appropriate program. The situation dramatically changes when the systems carrying the information are governed by quantum mechanics. Unknown states of quantum systems cannot be copied perfectly [1] and the no-programming theorem [2] prevents existence of universal quantum processors. This means that quantum programs cannot be copied and that by using registers of qubits (two level quantum systems) one cannot deterministically realize all quantum information processing functions with a fixed processor. So in contrast to classical devices, quantum ones cannot be replicated by just copying the program for them. Copying of quantum states was extensively investigated [3–7]. On the other hand, copying of quantum devices did not receive so much attention even though it is a fundamental and equally important quantum information processing task. Similarly to states, quantum transformations are often used in quantum key distribution schemes [8–11] to encode bits, so analysis of possible attacks by cloning them are needed. Cloning of transformations was analyzed only for the case of unitary transformations [11]. In the present paper we investigate cloning of measurement devices, which can be seen as a cloning of certain measure-and-prepare transformations. More precisely, when a measurement is an intermediate step of a quantum procedure, its outcome can influence the following operations. This feed forward of the classical outcome can be conveniently described using a quantum system into which the outcome is encoded into perfectly distinguishable orthogonal states. In this sense a quantum measurement with only classical outcomes [i.e., a positive operator-valued measure (POVM)] can be seen as a channel, which first measures the input system and based on the outcome prepares a state from a fixed orthogonal set.

The term cloning of observables has been used in Ref. [12] referring to state cloning machines preserving the statistics of a class of observables. In the present paper the objective is to actually mimic two uses of an unknown measurement device, while using it only once.

Let us denote by $A$ and $B$ the systems on which the replicas of the unknown measurement $\boldsymbol{E}$ should act. We denote by $C$ the system on which the measurement $\boldsymbol{E}$ acts and by $D$ the system encoding its outcome. The most general experimental setup for replication of an unknown measurement $\boldsymbol{E}$ can be seen as a sequence of three steps: (i) preprocessing of systems $A,B$ by a quantum channel, (ii) action of the measurement $\boldsymbol{E}$, and (iii) postprocessing phase. In general the preprocessing can produce an auxiliary quantum system that is not affected by the measurement $\boldsymbol{E}$. The postprocessing phase receives both the outcome of $\boldsymbol{E}$ and the auxiliary system and finally produces the outcomes of the replicated measurements. The most general representation of replication strategy is depicted below:




(1)

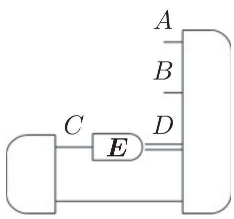
(the rectangular box represents a quantum channel, the round shaped box on the right represents a measurement, and the double wire carries the classical outcome of the measurement $\boldsymbol{E}$). In this case we talk about $1 \rightarrow 2$ *cloning* of a measurement device.

On the other hand, one might ask how well the task can be accomplished when we have to use the measurement before we have access to the state of systems $A,B$. In this case the most

*alessandro.bisio@unipv.it
†dariano@unipv.it
‡paolo.perinotti@unipv.it
§http://www.qubit.it
‖michal.sedlak@unipv.it

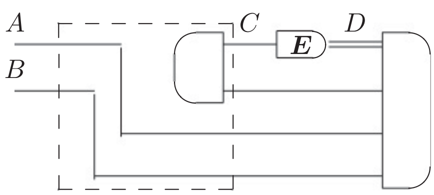
 

general strategy starts with a preparation of a bipartite state of system $C$ and some ancillary system. After the unknown measurement $\boldsymbol{E}$ is applied on $C$ the replication is achieved by a fixed measurement on the ancillary system and systems $A,B,D$. This scenario is depicted below and is called $1 \rightarrow 2$ *learning* of a measurement device:



. (2)

From comparison of Eqs. (1) and (2) one can see that learning is a special case of cloning. Indeed, if the preprocessing of a cloning strategy is restricted to store the input states of $A,B$ into the ancillary system and to prepare a fixed state of system $C$ then the two strategies coincide [see Eq. (3)]:



. (3)

That being so, it is clear that the performance of the optimal learning cannot be better than the performance of the optimal cloning.

In the present paper we will analyze the above two scenarios in which we assume $\boldsymbol{E}$ to be an arbitrary unknown von Neumann measurement. However, one can in principle think of more general versions of the problem, where for example $M$ replicas have to be produced out of $N$ uses of a measurement device. In particular, $N \rightarrow 1$ learning was analyzed in Ref. [13].

The paper is organized as follows. In Sec. II we expose the formulation of the optimal learning and cloning in mathematical terms. In Sec III we review the framework of quantum combs that is used as main tool throughout the paper. In Sec. IV the problem is simplified exploiting all the symmetries that can be useful. Sections V and VI are devoted to derivation of optimal cloning and learning, respectively. The paper is closed by concluding remarks in Sec. VII.

## II. MATHEMATICAL FORMULATION OF THE PROBLEM

Let us now formulate the problem mathematically. First of all, we should be able to evaluate the performance of the chosen replication strategy $\boldsymbol{\mathcal{R}}$. Hence we need a quantity that expresses the closeness of a replicated measurement to a desired bipartite von Neumann measurement. In the following lemma we introduce a function $\mathscr{F}(\boldsymbol{P},\boldsymbol{Q})$ that quantifies the closeness of a POVM $\boldsymbol{Q}$ to a von Neumann POVM $\boldsymbol{P}$ with the same number of outcomes. Throughout the paper we shall use boldface notation for objects that are composed from several elements. For example, $\boldsymbol{P} \equiv \{P_i\}_{i=1}^d$ denotes the POVM with elements $P_i$ and $\boldsymbol{I} \equiv \{I\}$ is the single outcome POVM on the Hilbert space $\mathscr{H}$ of dimension $d$.

*Lemma 1 (Fidelity criterion for POVM).* Let $\boldsymbol{P} \subseteq \mathcal{L}(\mathscr{H})$ and $\boldsymbol{Q} \subseteq \mathcal{L}(\mathscr{H})$ be two POVMs with $d$ possible outcomes, such that one of them is a von Neumann measurement. Consider now the quantity

$$\mathscr{F}(\boldsymbol{P},\boldsymbol{Q}) := \frac{1}{d}\sum_{i=1}^{d} \mathrm{Tr}[P_i Q_i]. \tag{4}$$

Then $\mathscr{F} = 1 \Leftrightarrow P_i = Q_i \; \forall i$ and $\mathscr{F} \leqslant 1$.

*Proof*. Without loss of generality we can assume that $\boldsymbol{P}$ is a von Neumann measurement and that we have $P_i = |i\rangle\langle i|$ where $|i\rangle$ is an orthonormal basis of $\mathscr{H}$. Then for $Q_i = P_i = |i\rangle\langle i|$ we have

$$\mathscr{F} = \frac{1}{d}\sum_{i=1}^{d}\mathrm{Tr}[P_i Q_i] = \frac{1}{d}\sum_{i=1}^{d}\mathrm{Tr}[|i\rangle\langle i|] = 1. \tag{5}$$

On the other hand, if $\mathscr{F} = 1$ we have

$$\begin{aligned} d &= \sum_{i=1}^{d}\mathrm{Tr}[P_i Q_i] = \sum_{i=1}^{d}\langle i|Q_i|i\rangle = \sum_{i,j=1}^{d}\langle i|Q_j|i\rangle - \sum_{i\neq j}\langle i|Q_j|i\rangle \\ &= \mathrm{Tr}\left[\sum_{j=1}^{d} Q_j\right] - \sum_{i\neq j}\langle i|Q_j|i\rangle = d - \sum_{i\neq j}\langle i|Q_j|i\rangle, \end{aligned} \tag{6}$$

which implies $\sum_{i\neq j}\langle i|Q_j|i\rangle = 0$, because $\sum_{j=1}^{d} Q_j = I$. Since $Q_j \geqslant 0$, we must have $\langle i|Q_j|i\rangle = 0$ for all $i \neq j$, and consequently $Q_j = \alpha_j |j\rangle\langle j|$ with $\alpha_j \geqslant 0$. Finally the condition $\sum_{j=1}^{d}\alpha_j|j\rangle\langle j| = \sum_{j=1}^{d} Q_j = I$ implies $\alpha_j = 1$ and thus $Q_j = P_j$. Proving that $\mathscr{F} \leqslant 1$ is easy. Since $Q_i$ is an element of a POVM we have $\langle i|Q_i|i\rangle \leqslant 1$ and consequently $\mathscr{F} = \frac{1}{d}\sum_{i=1}^{d}\langle i|Q_i|i\rangle \leqslant 1$. ■

Since we assume that the unknown measurement $\boldsymbol{E}$ to be replicated is a von Neumann POVM, we can write it in the following form:

$$E_i = |\phi_i\rangle\langle\phi_i|, \tag{7}$$

where $\{|\phi_i\rangle\}_{i=1}^d$ is an orthonormal basis of the Hilbert space $\mathcal{H}$. All the POVMs of this kind can be generated by rotating a reference POVM $\{|i\rangle\langle i|\}_{i=1}^d$ by elements of the group of unitary transformations $\mathrm{SU}(d)$ as follows:

$$E_i^{(U)} = U|i\rangle\langle i|U^\dagger, \quad U \in \mathrm{SU}(d). \tag{8}$$

Let us denote the bipartite POVM replicated by the strategy $\mathcal{R}$ as $\boldsymbol{G}^{(U)} \equiv \boldsymbol{G}(\boldsymbol{\mathcal{R}},\boldsymbol{E}^{(U)})$. Our task is to find such replicating strategy $\boldsymbol{\mathcal{R}}$ that the elements $G_{ij}^{(U)}$ are as close as possible to $E_i^{(U)} \otimes E_j^{(U)}$. Assuming that the unknown POVM $\boldsymbol{E}^{(U)}$ is randomly drawn according to the Haar distribution, we choose the quantity

$$\begin{aligned} F[\boldsymbol{\mathcal{R}}] &:= \int dU\, \mathscr{F}(\boldsymbol{G}^{(U)},\boldsymbol{E}^{(U)} \otimes \boldsymbol{E}^{(U)}) \\ &= \frac{1}{d^2}\sum_{i,j=1}^{d}\int dU\, \mathrm{Tr}\big[G_{ij}^{(U)}\big(E_i^{(U)} \otimes E_j^{(U)}\big)\big] \end{aligned} \tag{9}$$

as a figure of merit for the replicating strategy. Hence, after choosing one of the two considered scenarios ($1 \rightarrow 2$ cloning or learning) the goal is to find a strategy $\boldsymbol{\mathcal{R}}$, that maximizes $F[\boldsymbol{\mathcal{R}}]$.

### III. PRELIMINARY CONCEPTS

In this section we introduce the necessary notation and review the general theory of *quantum networks*, as developed in [14,15]. The main tool which is necessary in order to develop this framework is the Choi-Jamiołkowsky isomorphism. It is an isomorphism connecting any *quantum operation* (i.e., trace nonincreasing completely positive map) $\mathcal{M} : \mathcal{B}(\mathscr{H}_{\text{in}}) \mapsto \mathcal{B}(\mathscr{H}_{\text{out}})$ to a positive operator $M \in \mathcal{B}(\mathscr{H}_{\text{out}} \otimes \mathscr{H}_{\text{in}})$ defined as follows:

$$M := \mathcal{M} \otimes \mathcal{I}(|\omega\rangle\langle\omega|), \tag{10}$$

where $\mathcal{I}$ is the identical map on $\mathcal{B}(\mathscr{H}_{\text{in}})$, $|\omega\rangle := \sum_n |n\rangle|n\rangle \in \mathscr{H}_{\text{in}} \otimes \mathscr{H}_{\text{in}}$ and we fixed an orthonormal basis $\{|n\rangle\}$ on $\mathscr{H}_{\text{in}}$. The action of $\mathcal{M}$ on a given input state $\rho$ can be expressed in terms of $M$ as

$$\mathcal{M}(\rho) = \text{Tr}_{\text{in}}[M(I \otimes \rho^T)], \tag{11}$$

where $\text{Tr}_{\text{in}}$ denotes the partial trace over $\mathscr{H}_{\text{in}}$ and the superscript $T$ marks the transposition with respect to the basis $\{|n\rangle\}$.

If $\mathcal{M}$ is a *quantum channel* (i.e., a completely positive trace preserving map), its Choi-Jamiołkowsky operator $M$ satisfies

$$\text{Tr}_{\text{out}}[M] = I_{\text{in}}. \tag{12}$$

which expresses the trace preserving condition. On the other hand, when $\mathcal{M}$ is trace nonincreasing $M$ must obey the inequality

$$M \leqslant Z \tag{13}$$

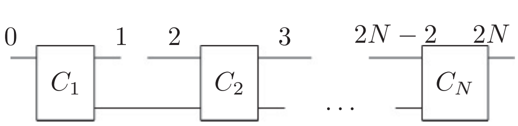
for some $Z$ that satisfies (12).

A collection of quantum operations $\boldsymbol{\mathcal{M}} \equiv \{\mathcal{M}_i\}$ such that $\mathcal{M}_\Omega := \sum_i \mathcal{M}_i$ is a quantum channel is called *quantum instrument*. Physically it represents a quantum device which produces the classical outcome $i$ and the quantum outcome $\mathcal{M}_i(\rho)/\text{Tr}[\mathcal{M}_i(\rho)]$ with probability $\text{Tr}[\mathcal{M}_i(\rho)]$ when the input state is $\rho$. The Choi-Jamiołkowsky isomorphism allows us to represent a quantum instrument as a set $\boldsymbol{M} \equiv \{M_i\}$ of positive operators such that

$$\sum_i M_i = M_\Omega, \tag{14}$$

where $M_\Omega$ is the Choi-Jamiołkowsky operator of a channel.

If we have several quantum devices we can feed the output of some of them into the input of some others, thus building a *quantum network*. Pairs of unconnected inputs and outputs form open slots of the network into which quantum devices can be later inserted. A network with $(N-1)$ open slots has $N$ input and $N$ output systems, that we label by even numbers from 0 to $2N-2$ and by odd numbers from 1 to $2N-1$, respectively. Each network can be visualized as in Eq. (15),




(15)

where the wires represent the connections of output systems to next inputs. This flow of quantum systems induces a *causal order* among the wires, according to which the input system $m$ cannot influence the output system $n$ if $m > n$.

A generalized version of the Choi-Jamiolkowsky isomorphism allows to represent a quantum network $\mathcal{R}$ in terms of a positive operator $R$, called *quantum comb*. $R$ acts on the Hilbert space $\mathscr{H}_{\text{out}} \otimes \mathscr{H}_{\text{in}}$ where $\mathscr{H}_{\text{out}} := \bigotimes_{j=0}^{N-1} \mathscr{H}_{2j+1}$, $\mathscr{H}_{\text{in}} := \bigotimes_{j=0}^{N-1} \mathscr{H}_{2j}$, and $\mathscr{H}_n$ being the Hilbert space of the $n$th system. For a deterministic quantum network (i.e., a network of quantum channels) the causal structure implies the following normalization condition:

$$\text{Tr}_{2k-1}[R^{(k)}] = I_{2k-2} \otimes R^{(k-1)}, \qquad k = 1,\ldots,N, \tag{16}$$

where $R^{(N)} = R$, $R^{(0)} = 1$, $R^{(k)} \in \mathcal{L}(\bigotimes_{n=0}^{2k-1} \mathscr{H}_n)$, $\text{Tr}_{2k-1}$ denotes the partial trace on $\mathscr{H}_{2k-1}$, and $I_{2k-2}$ is an identity operator on $\mathscr{H}_{2k-2}$. Equation (16) can be interpreted as the quantum network analog of Eq. (12).

If we consider probabilistic quantum networks (i.e., networks of quantum operations), their Choi-Jamiołkowsky operators satisfy the following generalized version of Eq. (13):

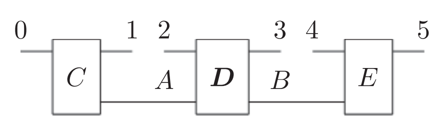
$$0 \leqslant R \leqslant T, \tag{17}$$

where $T$ is the Choi-Jamiołkowsky operator of a deterministic network.

We are now ready to introduce the quantum network generalization of the concept of quantum instrument. We call *generalized instrument* a set of probabilistic quantum networks $\boldsymbol{\mathcal{R}} := \{\mathcal{R}_i\}$ such that the set $\boldsymbol{R} := \{R_i\}$ of the corresponding Choi operators satisfies

$$\sum_i R_i = R_\Omega, \tag{18}$$

where $R_\Omega$ is the Choi operator of a deterministic network. We can say that a generalized instrument is the mathematical representation of a network of quantum devices that produces both the classical outcome $i$ and the quantum outcome $\mathcal{R}_i(\rho) \in \mathcal{L}(\mathscr{H}_{\text{out}})$ with probability $\text{Tr}[\mathcal{R}_i(\rho)]$ when the state $\rho \in \mathcal{L}(\mathscr{H}_{\text{in}})$ is fed into the free inputs of the network. A typical example of a generalized instrument is a quantum network in which one of the devices is a quantum instrument:




(19)

In Eq. (19) we have two channels $C$ and $E$ connected through wires $A$ and $B$ to the quantum instrument $\boldsymbol{D} \equiv \{D_i\}$.

Two quantum networks $\mathcal{R}_1$ and $\mathcal{R}_2$ can be connected by linking some outputs of $\mathcal{R}_1$ ($\mathcal{R}_2$) with inputs of $\mathcal{R}_2$ ($\mathcal{R}_1$), thus forming a new network $\mathcal{R}_3 := \mathcal{R}_1 * \mathcal{R}_2$. We adopt the convention that the wires to be connected are identified by the same label. The connection of the two quantum networks is mathematically represented by the *link product* of the corresponding Choi operators $R_1$ and $R_2$, which is defined as

$$R_1 * R_2 = \text{Tr}_{\mathscr{K}}\big[R_1^{\theta_{\mathscr{K}}} R_2\big], \tag{20}$$

$\theta_{\mathscr{K}}$ denoting partial transposition over the Hilbert space $\mathscr{K}$ of the connected systems (recall that we identify the Hilbert spaces of connected systems with the same label).

As an example the generalized instrument $\boldsymbol{R}$ from Eq. (19) is described by operators $R_i = C * D_i * E$.

As we pointed out in the Introduction, the classical outcome of the inserted measurement can influence the next operation of the network. In order to take the feed forward of the classical outcome into account it is convenient to describe the measurement device to be replicated as a measure and prepare quantum channel

$$\mathcal{E}^{(U)}(\rho) = \sum_{i=1}^{d} \mathrm{Tr}\big[E_i^{(U)}\rho\big]|i\rangle\langle i|, \tag{21}$$

which measures the POVM $\boldsymbol{E}^{(U)}$ on the input state and in the case of outcome $i$ prepares the state $|i\rangle$ from a fixed orthonormal basis on the output of the channel. Within this framework the classical outcome is encoded into a quantum system by preparing it into a state from a set of orthogonal states. The Choi-Jamiołkowski representation of the channel $\mathcal{E}^{(U)}$ is the following:

$$E^{(U)} = \sum_{i=1}^{d} |i\rangle\langle i| \otimes E_i^{(U)^T} = \sum_{i=1}^{d} |i\rangle\langle i| \otimes U^*|i\rangle\langle i|U^T, \tag{22}$$

where $X^T$ denotes the transpose of $X$ in the basis $\{|i\rangle\}_{i=1}^{d}$.

Since we want the replicating network $\mathcal{R}$ to behave as two copies of the POVM $\boldsymbol{E}^{(U)}$ upon insertion of a single use of $\mathcal{E}^{(U)}$, we have that $\mathcal{R}$ is actually a generalized instrument $\boldsymbol{R} \equiv \{R_{ij}\}_{i,j=1}^{d} \in \mathcal{L}(\mathscr{H}_A \otimes \mathscr{H}_B \otimes \mathscr{H}_C \otimes \mathscr{H}_D)\}$ where $i,j$ is the couple of outcomes of the two replicated measurements and the labeling of the Hilbert spaces follows from Eqs. (1) and (2).

Specializing Eq. (18) using Eq. (16), the normalization of the generalized instrument $\boldsymbol{R}$ has to obey the following equations:

$1 \to 2$ *cloning*

$$\sum_{i,j} R_{ij} = R_\Omega = I_D \otimes S_{ABC} \quad \mathrm{Tr}_C[S] = I_{AB}, \tag{23}$$

$1 \to 2$ *learning*

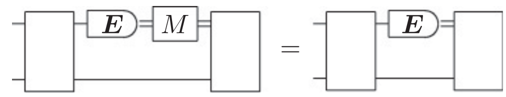
$$\sum_{i,j} R_{ij} = R_\Omega = I_{ABD} \otimes \rho_C \quad \mathrm{Tr}[\rho] = 1. \tag{24}$$

The replicated POVM is then equal to

$$\begin{aligned} G_{ij}^{(U)} &= \big[R_{ij} * E_{CD}^{(U)}\big]^T \\ &= \left[\sum_k \langle k|_C \langle k|_D (U^\dagger \otimes I) R_{ij} (U \otimes I)|k\rangle_C |k\rangle_D\right]^T. \end{aligned} \tag{25}$$

## IV. SYMMETRIES OF THE REPLICATING NETWORK

In this section we utilize the symmetries of the figure of merit (9) to simplify the optimization problem. These considerations apply both to cloning and learning of a measurement device. The first simplification relies on the fact that some wires of the network carry only classical information, representing the outcome of the measurement. The classical information encoded in the choice of a state from basis $\{|i\rangle\}$ can be read without disturbance by the measure and prepare channel $\mathcal{M}$ with Choi-Jamiolkowski operator $M \equiv E^{(I)}$, where $E^{(I)}$ is defined in Eq. (22) by choosing $U = I$. Thus, inserting channel $\mathcal{M}$ between the use of a measurement device $E^{(U)}$ and the network $\mathcal{R}$ will not change the operation of the scheme, i.e.,




(26)

As a consequence we have the following lemma.

*Lemma 2 (Restriction to diagonal network).* The optimal generalized instrument $\mathcal{R}$, $\sum_{i,j} R_{ij} = R_\Omega$ maximizing Eq. (9) can be chosen to satisfy:

$$R_{ij} = \sum_k R'_{ij,k} \otimes |k\rangle\langle k|_D, \tag{27}$$

where $0 \leqslant R'_{ij,k} \in \mathcal{L}(\mathscr{H}_A \otimes \mathscr{H}_B \otimes \mathscr{H}_C)$.

*Proof.* Let $S_{ij}$ be the Choi representation of a generalized instrument corresponding to a quantum network $\mathcal{S}$. Let us define network $\mathcal{R}$ as

$$R_{ij} := \sum_k \langle k|S_{ij}|k\rangle \otimes |k\rangle\langle k|, \tag{28}$$

which can be seen as $R_{ij} = S_{ij} * M = S_{ij} * E^{(I)}$ [see Eq. (22)] with the link performed on system $D$ carrying the classical information. We can easily prove that $\mathcal{R}$ is a generalized instrument. Indeed we have

$$\sum_{i,j} R_{ij} = \sum_{i,j} S_{ij} * E^{(I)} = S_\Omega * E^{(I)}, \tag{29}$$

where the link is performed only on the space $\mathscr{H}_D$. The operator in Eq. (29) is the Choi-Jamiołkowski operator of a deterministic quantum network satisfying the same normalization conditions as $S_\Omega$. Since $M * E^{(U)} = E^{(U)}$ we show that $\mathcal{S}$ and $\mathcal{R}$ produce the same replicated POVM $G_{ij}^{(U)}$ when linked with the single use of $E^{(U)}$, as follows:

$$\begin{aligned} \big(G_{ij}^{(U)}\big)^T &= S_{ij} * E_{CD}^{(U)} = S_{ij} * M * E_{CD}^{(U)} = R_{ij} * E_{CD}^{(U)} \\ &= \sum_k (\langle k|_D \langle k|_C U^\dagger) S_{ij} (|k\rangle_D U |k\rangle_C), \end{aligned} \tag{30}$$

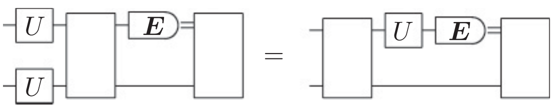
where the explicit form of the star product will be used later. The thesis then holds with $R'_{ij,k} := \langle k|S_{ij}|k\rangle$. ■

The restriction to diagonal networks allows us to simplify the figure of merit [Eq. (9)] as follows:

$$\begin{aligned} F[\mathcal{R}] &:= \int dU \mathscr{F}(\boldsymbol{G}^{(U)}, \boldsymbol{E}^{(U)} \otimes \boldsymbol{E}^{(U)}) \\ &= \frac{1}{d^2} \int dU \sum_{i,j,k} \mathrm{Tr}\big[R'_{ij,k} \\ &\quad \times U^{*\otimes 2} \otimes U|ijk\rangle\langle ijk|U^{T^{\otimes 2}} \otimes U^\dagger\big], \end{aligned} \tag{31}$$

where $|ijk\rangle \equiv |i\rangle_A |j\rangle_B |k\rangle_C$ and we applied Eq. (27).

Since the performance of the scheme is evaluated as an average over all possible "orientations" of the replicated measurement device, there exists a symmetrization procedure that can make any strategy covariant [i.e., having property from Eq (32)], without affecting the figure of merit:




(32)

This translates into mathematical terms as follows.

*Lemma 3 (Restriction to covariant networks)*. The operators $R'_{ij,k}$ that maximize Eq. (31) can be chosen to satisfy the commutation relation

$$[R'_{ij,k}, U_A^* \otimes U_B^* \otimes U_C] = 0. \tag{33}$$

*Proof.* Suppose that the generalized instrument corresponding to $S'_{ij,k}$ is optimal. Then one can easily check that also the instrument $R'_{ij,k}$ defined as follows:

$$R'_{ij,k} := \int dU (U^{*\otimes 2} \otimes U) S'_{ij,k} (U^{T\otimes 2} \otimes U^\dagger) \tag{34}$$

is suitably normalized and satisfies $[R'_{ij,k}, U^* \otimes U^* \otimes U] = 0$. Generalized instrument $\mathcal{R}$ corresponds to a strategy where random unitary $U^\dagger$, $U^\dagger$, $U$ is applied before and after the original strategy $\mathcal{S}$ to systems $A$, $B$, $C$, respectively. From the integration in Eq. (31) it is obvious that the value of $F$ for the above choice of $R'_{ij,k}$ is the same as for $S'_{ij,k}$. ■

The commutation relation (33) allows us to rewrite the figure of merit as

$$F[\mathcal{R}] = \frac{1}{d^2} \sum_{i,j,k} \langle ijk | R'_{ij,k} | ijk \rangle_{ABC}. \tag{35}$$

Another symmetry we can utilize is related to a simultaneous relabeling of the outcomes of the inserted and produced measurements. We shall denote by $\sigma$ the element of $\mathbb{S}_d$, the group of permutations of $d$ elements, and by $T_\sigma$ the linear operator that permutes the elements of basis $\{|i\rangle\}$ according to this permutation, in formula $T_\sigma |i\rangle = |\sigma(i)\rangle$. Let us note that the complex conjugation and transposition are defined with respect to the basis $\{|i\rangle\}$, so $T_\sigma = T_\sigma^*$.

*Lemma 4 (Relabeling symmetry)*. Without loss of generality we can assume that the operators $R'_{ij,k}$ that maximize Eq. (31) satisfy the relation

$$R'_{ij,k} = R'_{\sigma(ij,k)}. \tag{36}$$

where we shortened $\sigma(ij,k) := (\sigma(i)\,\sigma(j), \sigma(k))$.

*Proof.* Suppose that network $\mathcal{S}$ characterized by operators $S_{ij}$ is optimal and satisfies both conditions (27) and (33). Let us then define

$$\begin{aligned} R'_{ij,k} &:= \frac{1}{d!} \sum_{\sigma \in \mathbb{S}_d} \left(T_\sigma^{\dagger\otimes 2} \otimes T_\sigma{}^\dagger\right) S'_{\sigma(ij,k)} \left(T_\sigma{}^{\otimes 2} \otimes T_\sigma\right) \\ &= \frac{1}{d!} \sum_{\sigma \in \mathbb{S}_d} S'_{\sigma(ij,k)}, \end{aligned} \tag{37}$$

where the last identity in (37) follows from the commutation relation (33) with $U = T_\sigma$. The operators $R'_{ij,k}$ correspond to a valid quantum network $\mathcal{R}$, because $\mathcal{R}$ is a convex combination of networks $\mathcal{Z}^\sigma$ defined by Eq. (27) with $Z'^{\sigma}_{ij,k} = S'_{\sigma(ij,k)}$. Quantum network $\mathcal{R}$ operationally corresponds to relabeling of the outcomes of the inserted and replicated measurements by permutation $\sigma$. The figure of merit for $\mathcal{R}$ is

$$\begin{aligned} F[\mathcal{R}] &= \frac{1}{d^2} \sum_{i,j,k} \langle ijk | R'_{ij,k} | ijk \rangle \\ &= \frac{1}{d^2 d!} \sum_{\sigma \in \mathbb{S}_d} \sum_{i,j,k} \langle \sigma(ijk) | S'_{\sigma(ij,k)} | \sigma(ijk) \rangle = F[\mathcal{S}]. \end{aligned} \tag{38}$$

It is easy to prove that $R'_{ij,k}$ satisfies Eq. (36). ■

*Remark 1.* The properties (27), (33), and (36) induce the following structure of the replicated POVMs:

$$G^{(U)}_{\sigma(ij)} = (UT_\sigma)^{\otimes 2} G^{(I)}_{ij} (T_\sigma^\dagger U^\dagger)^{\otimes 2}. \tag{39}$$

The advantage of using the relabeling symmetry is the reduction of the number of independent parameters of the quantum generalized instrument. Let us define the equivalence relation between strings $ijk$ and $i'j'k'$ as

$$ijk \sim i'j'k' \quad \Leftrightarrow \quad ijk = \sigma(i'j'k') \tag{40}$$

for some permutation $\sigma$. Thanks to Eq. (36) there are only as many independent $R'_{ij,k}$ as there are equivalence classes among sequences $(ij,k)$. There are four or five equivalence classes depending on the dimension $d$ being two or greater than two, respectively. We denote the set of these equivalence classes by $\mathsf{L} := \{xxx, xxy, xyx, xyy, xyz\}$.

Based on lemma (4) we can write the optimal generalized instrument as follows:

$$R_{ab,c} := R'_{ij,k} = R'_{\sigma(ij,k)}, \tag{41}$$

where $(ab,c)$ is a string of indices that represents one equivalence class from $\mathsf{L}$.

The figure of merit can finally be written as follows:

$$F[\mathcal{R}] = \frac{1}{d^2} \sum_{(ab,c)\in \mathsf{L}} n(ab,c) \langle R_{ab,c} \rangle, \tag{42}$$

where $n(ab,c)$ is the cardinality of the equivalence class denoted by $(ab,c)$, and $\langle R_{ab,c} \rangle = \langle ijk | R'_{ij,k} | ijk \rangle$ for any string $ijk$ in the equivalence class denoted by $(ab,c)$. As a consequence of Schur's lemmas, Eq. (33) implies the following structure for the operators $R_{ab,c}$:

$$R_{ab,c} = \bigoplus_\nu P^\nu \otimes r^\nu_{ab,c}, \tag{43}$$

where $\nu$ labels the irreducible representations in the Clebsch-Gordan series of $U_A^* \otimes U_B^* \otimes U_C$, and $P^\nu$ acts as the identity on the invariant subspaces of the representations $\nu$, while $r^\nu_{ab,c}$ acts on the multiplicity space of the same representation.

Depending on the dimension $d = 2$ or $d > 2$ we have two different decompositions. In the former case, we have

$$R_{ab,c} = P^\alpha \otimes r^\alpha_{ab,c} + P^\beta r^\beta_{ab,c}, \tag{44}$$

where $r^\alpha_{ab,c}$ is a positive $2 \times 2$ matrix, while $r^\beta_{ab,c}$ is a nonnegative real number. The projections $P^\xi$ on the invariant spaces of the representation $U^* \otimes U^* \otimes U$ are the following:

$$\begin{aligned} P^\alpha \otimes |i\rangle\langle j| &= \sum_{m=1}^d \left|\Psi^i_m\right\rangle\!\left\langle\Psi^j_m\right|, \quad i,j \in \{+,-\}, \\ P^\beta &= I \otimes P^+ - P^\alpha \otimes |+\rangle\langle +|, \end{aligned} \tag{45}$$

where $|\Psi_m^{\pm}\rangle = (|\omega\rangle|m\rangle \pm |m\rangle|\omega\rangle)/[2(d\pm 1)]^{\frac{1}{2}}$, and $P^+$, $P^-$, are the projections onto the symmetric and antisymmetric subspace, respectively. When $d>2$, on the other hand, we have

$$R_{ab,c} = P^\alpha \otimes r^\alpha_{ab,c} + P^\beta r^\beta_{ab,c} + P^\gamma r^\gamma_{ab,c}, \tag{46}$$

where $r^\alpha_{ab,c}$ is a positive $2\times 2$ matrix, while $r^\beta_{ab,c}$ and $r^\gamma_{ab,c}$ are non-negative real numbers. The projections $P^\xi$ on the invariant subspaces are the following:

$$\begin{aligned} P^\alpha \otimes |a\rangle\langle b| &= \sum_{m=1}^{d} \big|\Psi_m^a\big\rangle\big\langle\Psi_m^b\big|, \quad a,b\in\{+,-\rangle,\\ P^\beta &= I\otimes P^+ - P^\alpha \otimes |+\rangle\langle +|,\\ P^\gamma &= I\otimes P^- - P^\alpha \otimes |-\rangle\langle -|. \end{aligned} \tag{47}$$

The last symmetry we are going to introduce relies on the possibility of exchanging the inputs (Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$) of the two replicated measurements with simultaneously exchanging their measurement outcomes, while the figure of merit is left unchanged.

*Lemma 5*. The operators $R_{ab,c}$ in Eq. (43) can be chosen to satisfy

$$R_{ab,c} = \mathsf{S} R_{ba,c}\mathsf{S} \quad \forall (ab,c)\in \mathsf{L}, \tag{48}$$

where $\mathsf{S}$ is the swap operator $\mathsf{S}|k\rangle_A|j\rangle_B = |j\rangle_A|k\rangle_B$.

*Proof*. The proof can be done by the following averaging argument. Let us define $\overline{R}_{ij,k} := \frac{1}{2}(R'_{ij,k} + \mathsf{S}R'_{ji,k}\mathsf{S})$. It is easy to prove that $\{\overline{R}_{ij,k}\}$ satisfies the corresponding normalization [Eq. (23) for cloning or Eq. (24) for learning] and that it gives the same value of $F$ as $R'_{ij,k}$. ■

Equation (48) together with the decomposition (46) gives for $\forall (ab,c)\in\mathsf{L}$,

$$\sigma_z r^\alpha_{ab,c}\sigma_z = r^\alpha_{ba,c}, \quad r^\beta_{ab,c} = r^\beta_{ba,c}, \quad r^\gamma_{ab,c} = r^\gamma_{ba,c}, \tag{49}$$

where $\sigma_z = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$.

As a consequence of Eq. (43) , the figure of merit in Eq. (42) can be written as

$$\begin{aligned} F &= \frac{1}{d^2}\sum_{(ab,c)\in\mathsf{L}} n(ab,c)\,\mathrm{Tr}\Big[|ijk\rangle\langle ijk|\sum_\nu P^\nu\otimes r^\nu_{ab,c}\Big]\\ &= \sum_\nu \frac{1}{d}\sum_{(ab,c)\in\mathsf{L}} \mathrm{Tr}\big[\Delta^\nu_{ab,c}s^\nu_{ab,c}\big] = F_\alpha + F_\beta + F_\gamma, \end{aligned} \tag{50}$$

where

$$\Delta^\nu_{ab,c} := \mathrm{Tr}_{\mathscr{H}_\nu}[|ijk\rangle\langle ijk|], \tag{51}$$

$$s^\nu_{ab,c} := \frac{n(ab,c)}{d} r^\nu_{ab,c}, \tag{52}$$

and $ij,k$ is any triple of indices in the class denoted by $ab,c$. Notice that $n(xx,x)=d$, $n(xx,y)=n(xy,x)=n(xy,y)=d(d-1)$, $n(xy,z)=d(d-1)(d-2)$ and in the case $d=2$ $F_\gamma=0$ (i.e., does not appear).

In particular, by direct calculation we have

$$\begin{aligned} &\Delta^\alpha_{xx,x} = \begin{pmatrix}\frac{2}{d+1} & 0\\ 0 & 0\end{pmatrix}, \quad \Delta^\alpha_{xx,y} = \frac{1}{2}\begin{pmatrix}\frac{1}{d+1} & \frac{1}{\sqrt{d^2-1}}\\ \frac{1}{\sqrt{d^2-1}} & \frac{1}{d-1}\end{pmatrix},\\ &\Delta^\alpha_{xy,y} = \Delta^\alpha_{xy,z} = 0, \quad \Delta^\alpha_{xy,x} = \sigma_z\Delta^\alpha_{xx,y}\sigma_z,\\ &\Delta^\beta_{xx,x} = \frac{d-1}{d+1}, \quad \Delta^\beta_{xx,y} = \Delta^\beta_{xy,x} = \frac{d}{2(d+1)},\\ &\Delta^\beta_{xy,y} = 1, \quad \Delta^\beta_{xy,z} = \frac{1}{2}, \quad \Delta^\gamma_{xx,x} = \Delta^\gamma_{xy,y} = 0,\\ &\Delta^\gamma_{xx,y} = \Delta^\gamma_{xy,x} = \frac{d-2}{2(d-1)}, \quad \Delta^\gamma_{xy,z} = \frac{1}{2}. \end{aligned} \tag{53}$$

## V. OPTIMAL CLONING

In this section we turn our attention to the cloning scenario. Cloning is less restrictive than learning, since we allow the two states to be measured to be available at the same time as the single use of the measurement device. The normalization condition for the $1\to 2$ cloning reads

$$\sum_{ij,k} |k\rangle\langle k|_D \otimes R'_{ij,k} = I_D\otimes S_{ABC}, \qquad \mathrm{Tr}_C[S] = I_{AB} \tag{54}$$

which implies the following:

$$\begin{aligned} I_{AB} &= \mathrm{Tr}_C[R_{xx,x}] + (d-1)(d-2)\mathrm{Tr}_C[R_{xy,z}]\\ &\quad + (d-1)\mathrm{Tr}_C[R_{xx,y} + R_{xy,x} + R_{xy,y}]. \end{aligned} \tag{55}$$

From the commutation $[R_{ab,c}, U_A^*\otimes U_B^*\otimes U_C]$ it follows that $[\mathrm{Tr}_C[R_{ab,c}], U_A^*\otimes U_B^*]=0$ and taking the decomposition $R_{ab,c} = \sum_\nu P^\nu\otimes r^\nu_{ab,c}$ along with definition (52), the normalization constraint (55) becomes

$$P^\pm = P^\pm \sum_\nu \sum_{(ab,c)\in\mathsf{L}} \mathrm{Tr}_C\big[P^\nu\otimes s^\nu_{ab,c}\big]P^\pm. \tag{56}$$

We take the trace of the previous equation to obtain the following equivalent formulation of the normalization constraints:

$$d_+ = d_\alpha \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,+,+}_{ab,c} + d_\beta \sum_{(ab,c)\in\mathsf{L}} s^\beta_{ab,c}, \tag{57}$$

$$d_- = d_\alpha \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,-,-}_{ab,c} + d_\gamma \sum_{(ab,c)\in\mathsf{L}} s^\gamma_{ab,c}, \tag{58}$$

where $d_\pm \equiv \mathrm{Tr}[P^\pm]$, $d_\nu\equiv \mathrm{Tr}[P^\nu]$. If we introduce the notation

$$\begin{aligned} &s^\beta_{a,bc} := \begin{pmatrix} s^\beta_{a,bc} & 0\\ 0 & 0\end{pmatrix}, \quad s^\gamma_{a,bc} := \begin{pmatrix} 0 & 0\\ 0 & s^\gamma_{a,bc}\end{pmatrix},\\ &\Pi^+ = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \quad \Pi^- = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \end{aligned} \tag{59}$$

the normalization constraints (57) and (58) can be rewritten as

$$\begin{aligned} &\Pi^+\Bigg(\sum_{\nu,(a,bc)\in\mathsf{L}} d_\nu s^\nu_{a,bc}\Bigg)\Pi^+ = \begin{pmatrix} d_+ & 0\\ 0 & 0\end{pmatrix},\\ &\Pi^-\Bigg(\sum_{\nu,(a,bc)\in\mathsf{L}} d_\nu s^\nu_{(a,bc)}\Bigg)\Pi^- = \begin{pmatrix} 0 & 0\\ 0 & d_-\end{pmatrix}. \end{aligned} \tag{60}$$

In order to solve the optimization problem we have to find the set $\mathsf{s} := \{s^\nu_\ell, \ell\in\mathsf{L}, \nu\in\{\alpha,\beta\gamma\}\}$, $s^\nu_\ell\in\mathcal{L}(\mathbb{C}^2)$, $s^\nu_\ell\geqslant 0$ subjected to the constraint (60) that maximizes the figure of merit (50); we will denote as $\mathsf{M}$ the set of all the $\mathsf{s}$ satisfying Eq. (60). Since the figure of merit (50) is linear and the set $\mathsf{M}$ is convex, a trivial result of convex analysis states that the maximum of a convex function over a convex set is achieved at an extremal point of the convex set. We now give two necessary conditions for a given $\mathsf{S}$ to be an extremal point of $\mathsf{M}$. Let us start with the following.

*Definition 1 (Perturbation)*. Let $\mathsf{S}$ be an element of $\mathsf{M}$. A set of Hermitian operators $\mathsf{Z} := \{z_\ell^\nu\}$ is a *perturbation* of $\mathsf{S}$ if there exists $\epsilon \geqslant 0$ such that

$$\mathsf{S} + h\mathsf{Z} \in \mathsf{M} \quad \forall h \in [-\epsilon,\epsilon], \tag{61}$$

where we defined $\mathsf{S} + h\mathsf{Z} := \{s_\ell^\nu + hz_\ell^\nu | h \in [-\epsilon,\epsilon]\}$.

By the definition of perturbation it is easy to prove that an element $\mathsf{S}$ of $\mathsf{M}$ is extremal if and only if it admits only the trivial perturbation $z_\ell^\nu = 0\,\forall \ell,\nu$. We now exploit this definition to prove two necessary conditions for extremality.

*Lemma 6*. Let $\mathsf{S}$ be an extremal element of $\mathsf{M}$. Then $s_\ell^\nu$ has to be rank one for all $\ell,\nu$.

*Proof.* Suppose that there is a $s_{\ell'}^{\nu'} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathsf{S}$ which is not rank one; then there exist $\epsilon$ such that $\mathsf{Z} := \{0,\ldots,0,z_{\ell'}^{\nu'},0,\ldots,0\}$, $z_{\ell'}^{\nu'} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is an admissible perturbation. ■

The above lemma tells us that without lost of generality we can assume the optimal $\mathsf{S}$ to be a set of rank one matrices. Let us now consider a set $\mathsf{S}$ such that $s_\ell^\nu$ is rank one for all $\ell,\nu$; any admissible perturbation $\mathsf{Z}$ of $\mathsf{S}$ must satisfy

$$z_\ell^\nu = c_\ell^\nu s_\ell^\nu, \quad c_\ell^\nu \in \mathbb{R}, \tag{62}$$

$$\Pi^+ \left( \sum_{\nu,\ell} d_\nu c_\ell^\nu s_\ell^\nu \right) \Pi^+ = \Pi^- \left( \sum_{\nu} d_\nu c_\ell^\nu s_\ell^\nu \right) \Pi^- = 0. \tag{63}$$

where the constraint (62) is required in order to have $s_\ell^\nu + hz_\ell^\nu \geqslant 0$, while Eq. (63) tells us that $\mathsf{S} + h\mathsf{Z}$ satisfies the normalization (60). Let us now consider the map

$$f : \mathcal{L}(\mathbb{C}^2) \longrightarrow \mathbb{C}^2, \quad f(A) := \begin{pmatrix} \mathrm{Tr}[\Pi^+ A] \\ \mathrm{Tr}[\Pi^- A] \end{pmatrix},$$
$$f \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ d \end{pmatrix},$$

exploiting this definition Eq. (63) becomes

$$\sum_{\nu,\ell} c_\ell^\nu f\left(s_\ell^\nu\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{64}$$

Suppose now that the set $\bar{\mathsf{S}}$ has $N \geqslant 3$ nonzero elements; then $\{f(\bar{s}_\ell^\nu)\}$ is a set of $N \geqslant 3$ vectors of $\mathbb{C}^2$ that cannot be linearly independent. That being so, there exists a set of coefficients $\{c_\ell^\nu\}$ such that $\sum_{\nu,\ell} c_\ell^\nu f(s_\ell^\nu) = 0$ and then $z_\ell^\nu = c_\ell^\nu \bar{s}_\ell^\nu$ is a perturbation of $\bar{\mathsf{S}}$. We have then proved the following lemma.

*Lemma 7*. Let $\mathsf{S}$ be an extremal element of $\mathsf{M}$. Then $\mathsf{S}$ cannot have more than two nonzero elements.

Lemmas 6 and 7 provide two necessary conditions for extremality that allow us to restrict the search of the optimal $\mathsf{S}$ among the ones that satisfy

$$\begin{gathered} \mathsf{S} = \left\{s_{\ell'}^{\nu'}, s_{\ell''}^{\nu''}\right\}, \quad \mathrm{Rnk}\left(s_{\ell'}^{\nu'}\right) = \mathrm{Rnk}\left(s_{\ell''}^{\nu''}\right) = 1, \\ \Pi^i \left( \sum_{\nu,\ell} d_\nu s_\ell^\nu \right) \Pi^i = d_i, \quad i = +,-\,. \end{gathered} \tag{65}$$

The set of the above $\mathsf{S}$ is small enough to allow us to compute the value of $F$ for all the possible cases. It turns out that there are two choices achieving the highest value of fidelity

$$F = \frac{4}{3d}. \tag{66}$$

They are defined by $\mathsf{S} = \{s_{xx,x}^\alpha, s_{xy,x}^\alpha\}$ and $\mathsf{S} = \{s_{xx,x}^\alpha, s_{xy,y}^\alpha\}$, where

$$\begin{gathered} s_{xx,x}^\alpha = \begin{pmatrix} \frac{9d_+-1}{9d} & 0 \\ 0 & 0 \end{pmatrix} \equiv A, \quad B \equiv \begin{pmatrix} \frac{1}{9d} & \frac{\sqrt{d_-}}{3d} \\ \frac{\sqrt{d_-}}{3d} & \frac{d_-}{d} \end{pmatrix}, \\ s_{xy,x}^\alpha = B, \quad s_{xy,y}^\alpha = \sigma_z B \sigma_z. \end{gathered} \tag{67}$$

From the linearity of the link product and our figure of merit it follows that also any convex combination of the above two strategies will give the optimal performance. In the rest of the paper we consider the equal convex combination of the above two strategies:

$$s_{xx,x}^\alpha = A, \quad s_{xy,x}^\alpha = \tfrac{1}{2}B, \quad s_{xy,y}^\alpha = \tfrac{1}{2}\sigma_z B \sigma_z, \tag{68}$$

because it treats the two clones in the same way. Using Eq. (30) one can derive the form of the replicated POVM corresponding to the above choice of the optimal generalized instrument:

$$\begin{aligned} G_{ii} &= \left[1 - \frac{2}{9d(d+1)}\right] P^+\left(E_i^{(U)} \otimes I_B\right) P^+, \\ G_{ij} &= \frac{1}{d-1}\left[Q^+\left(E_i^{(U)} \otimes I_B\right)Q^+ + Q^-\left(E_j^{(U)} \otimes I_B\right)Q^-\right], \end{aligned}$$

where $Q^\pm = 1/\sqrt{9d(d+1)}\; P^+ \pm 1/\sqrt{2}\; P^-$.

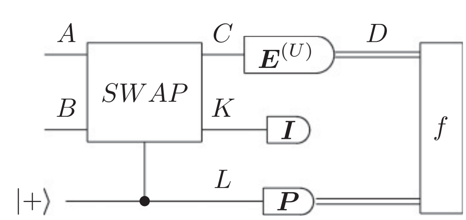
### A. Realization scheme for the optimal cloning network

In this section we describe the inner structure of the optimal cloning network. First we notice that the choice from Eq. (68) corresponds to the generalized instrument

$$\begin{aligned} R_{ii} &= |i\rangle\langle i| \otimes \frac{9d_+ - 1}{9d} \sum_k |\Psi_k^+\rangle\langle\Psi_k^+|, \\ R_{ij} &= |i\rangle\langle i| \otimes \frac{1}{2(d-1)} \sum_k |\phi_k\rangle\langle\phi_k| \\ &\quad + |j\rangle\langle j| \otimes \frac{1}{2(d-1)} \sum_k \tilde{\sigma}_z |\phi_k\rangle\langle\phi_k| \tilde{\sigma}_z, \end{aligned} \tag{69}$$

$$\begin{gathered} |\phi_k\rangle = \sqrt{\frac{1}{9d}}|\Psi_k^+\rangle + \sqrt{\frac{d_-}{d}}|\Psi_k^-\rangle \\ \tilde{\sigma}_z|\Psi_k^\pm\rangle = \pm|\Psi_k^\pm\rangle. \end{gathered}$$

The generalized instrument $\boldsymbol{R}$ can be realized by the following network:




(70)

The first step consists of a controlled-SWAP gate, which is described by the unitary

$$\begin{aligned}\text{CSWAP} = {} & T_{A\to C} \otimes T_{B\to K} \otimes |0\rangle\langle 0|_L \\ & + T_{A\to K} \otimes T_{B\to C} \otimes |1\rangle\langle 1|_L\end{aligned}$$

with the control qubit prepared in the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. We defined $T_{X\to Y} = \sum_i |i\rangle_Y |i\rangle_X$ and named $\mathscr{H}_L$ the two-dimensonal Hilbert space of the control qubit with $\{|0\rangle,|1\rangle\}$ being an orthonormal basis on $\mathscr{H}_L$.

In the second step we have three commuting actions:

(1) the single use of the measurement device $\boldsymbol{E}^{(U)}$ is applied on system $C$ and its outcome is recorded on a classical memory $D$,

(2) system $K$ is discarded, and

(3) system $L$ undergoes a three-outcome measurement described by the POVM $\boldsymbol{P}$ defined as follows:

$$\begin{aligned} P_1 &= \frac{[9d(d+1)-2]}{9d(d+1)}|+\rangle\langle+|, \\ P_2 &= |\psi\rangle\langle\psi|, \quad P_3 = \sigma_z|\psi\rangle\langle\psi|\sigma_z, \\ |\psi\rangle &= \sqrt{\frac{1}{9d(d+1)}}|+\rangle + \sqrt{\frac{1}{2}}|-\rangle. \end{aligned} \tag{71}$$

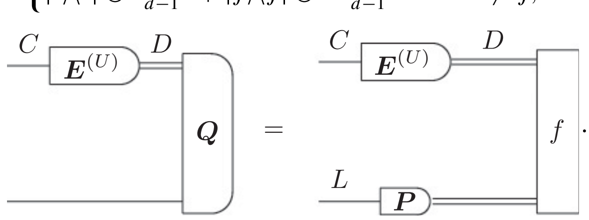
The last step is just a classical processing $f$ of the outcome $k$ of the measurement $\boldsymbol{E}^{(U)}$ and of the outcome $n$ of POVM $\boldsymbol{P}$. The function $f$ that produces the outcome $(i,j) = f(k,n)$ of the network is defined as follows:

$$f(k,n) = \begin{cases} (k,k) & \text{if } n=1 \\ (k,j) \quad j\neq k & \text{if } n=2 \\ (j,k) \quad j \neq k & \text{if } n=3 \end{cases}, \tag{72}$$

where the outcome $j$ in the second and third case is randomly generated with flat distribution.

In order to prove that this network is described by the generalized instrument in Eq. (69) we first realize that the action of the POVM $\boldsymbol{P}$ and of the processing $f$ can be represented by the bipartite POVM $\boldsymbol{Q}$ on systems $D$ and $L$ defined as

$$Q_{i,j} = \begin{cases} |i\rangle\langle i| \otimes \frac{[9d(d+1)-2]}{9d(d+1)}|+\rangle\langle+| & \text{if } i=j \\ |i\rangle\langle i| \otimes \frac{|\psi\rangle\langle\psi|}{d-1} + |j\rangle\langle j| \otimes \frac{\sigma_z|\psi\rangle\langle\psi|\sigma_z}{d-1} & \text{if } i\neq j, \end{cases}$$




(73)

Finally, one can check the identity

$$R_{ij} = |+\rangle\langle+| * M_{\text{CSWAP}} * (Q_{i,j} \otimes I_K), \tag{74}$$

where $M_{\text{CSWAP}}$ is the Choi-Jamiolkowski operator of the control SWAP unitary channel.

It is worth noticing that the optimal cloning of the measurement device has some features in common with the optimal cloning of unitaries. Both in the cloning of unitaries and in the cloning of von Neumann measurements the first step is to perform a control-SWAP of the two input systems with the control qubit prepared in the superposition $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. We could give an intuitive explanation of this feature in terms of quantum parallelism: for a bipartite input $|\chi\rangle_0|\xi\rangle_1$ the unknown measurement acts on both input states via a superposition $\sqrt{\frac{1}{2}}(|\chi\rangle_2|\xi\rangle_A + |\xi\rangle_2|\chi\rangle_A)$.

## VI. OPTIMAL LEARNING

Our goal in this scenario is to create two replicas of the measurement after it was used once. Let us consider the normalization constraint for the generalized instrument $R_{ij}$. Since $\sum_{i,j} R_{ij}$ has to be a deterministic network, we have

$$\sum_{ijk} |k\rangle\langle k|_D \otimes R'_{ij,k} = I_{ABD} \otimes \rho_C, \quad \text{Tr}[\rho] = 1, \tag{75}$$

where $\rho$ has to be positive operator. The commutation relation (33) implies $[\rho,U] = 0$ and so we have $\rho = \frac{1}{d}I_C$. Writing $I_{ABCD}$ as $\sum_k |k\rangle\langle k|_D \otimes (I_{m_\alpha} \otimes P^\alpha + P^\beta + P^\gamma)$ we can rewrite the normalization conditions as follows:

$$\begin{aligned} \sum_{(ab,c)\in\mathsf{L}} s^{\nu}_{ab,c} &= \frac{1}{d}, \quad \nu = \beta,\gamma, \\ \sum_{(ab,c)\in\mathsf{L}} s^{\alpha}_{ab,c} &= \frac{1}{d}I_{m_\alpha}. \end{aligned} \tag{76}$$

Let us now maximize the figure of merit under these constraints. The maximization of $F_\beta$ and $F_\gamma$ is simple and yields

$$F_\beta = \frac{1}{d^2}, \quad F_\gamma = \frac{1}{2d^2}, \tag{77}$$

$$\begin{aligned} s^{\beta}_{xx,x} = s^{\beta}_{xy,x} = s^{\beta}_{xy,y} = s^{\beta}_{xy,z} = 0, \\ s^{\gamma}_{xx,x} = s^{\gamma}_{xx,y} = s^{\gamma}_{xy,x} = s^{\gamma}_{xy,y} = 0, \end{aligned} \tag{78}$$

$$s^{\beta}_{xx,y} = s^{\gamma}_{xy,z} = \frac{1}{d}.$$

Let us now consider the maximization of $F_\alpha$; the normalization constraint for the $\alpha$ subspace gives

$$\begin{aligned} \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,+,+}_{ab,c} = \frac{1}{d}, \quad \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,+,-}_{ab,c} = 0, \\ \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,-,-}_{ab,c} = \frac{1}{d}, \quad \sum_{(ab,c)\in\mathsf{L}} s^{\alpha,-,+}_{ab,c} = 0. \end{aligned} \tag{79}$$

Inserting the explicit expression of the $\Delta^{\alpha}_{ab,c}$ into Eq. (50) and taking into account Eq. (49) we have

$$\begin{aligned} dF_\alpha = {} & \text{Tr}\left[\begin{pmatrix} s^{\alpha,+,+}_{xx,x} & s^{\alpha,+,-}_{xx,x} \\ s^{\alpha,-,+}_{xx,x} & s^{\alpha,-,-}_{xx,x} \end{pmatrix}\begin{pmatrix} \frac{2}{d+1} & 0 \\ 0 & 0 \end{pmatrix}\right] \\ & +\text{Tr}\left[\begin{pmatrix} s^{\alpha,+,+}_{xy,x} & s^{\alpha,+,-}_{xy,x} \\ s^{\alpha,-,+}_{xy,x} & s^{\alpha,-,-}_{xy,x} \end{pmatrix}\begin{pmatrix} \frac{1}{d+1} & \frac{1}{\sqrt{d^2-1}} \\ \frac{1}{\sqrt{d^2-1}} & \frac{1}{d-1} \end{pmatrix}\right] \\ = {} & \frac{2s^{\alpha,+,+}_{xx,x}}{d+1} + \frac{s^{\alpha,+,+}_{xy,x}}{d+1} + \frac{s^{\alpha,-,-}_{xy,x}}{d-1} + \frac{2s^{\alpha,+,-}_{xy,x}}{\sqrt{d^2-1}} \\ \leqslant {} & \frac{5d-3}{2d(d^2-1)} - \frac{3s^{\alpha,+,+}_{xy,x}}{d+1} + 2\sqrt{\frac{s^{\alpha,+,+}_{xy,x}}{2d(d^2-1)}}, \end{aligned} \tag{80}$$

where in the derivation of the bound (80) we used the positivity of $s^{\alpha}_{xy,x}$ and the constraints (79). The upper bound (80) can be achieved by taking

$$s^{\alpha}_{xx,x} = \begin{pmatrix} \frac{1}{d} - 2a & 0 \\ 0 & 0 \end{pmatrix}, \quad s^{\alpha}_{xy,x} = \begin{pmatrix} a & \sqrt{\frac{1}{2d}a} \\ \sqrt{\frac{1}{2d}a} & \frac{1}{2d} \end{pmatrix},$$
$$s^{\alpha}_{xy,z} = s^{\alpha}_{xx,y} = 0, \tag{81}$$

where we defined $a := s^{\alpha,+,+}_{xy,x}$. Equation (80) gives the value of $F_\alpha$ as a function of $a$; the maximization of $F_\alpha(a)$ with the constraint $0 \leqslant a \leqslant \frac{1}{2d}$ is easy and gives

$$F_\alpha = \frac{4(2d-1)}{3d^2(d^2-1)} \quad \text{for } a = \frac{d+1}{18d(d-1)}. \tag{82}$$

and then for $d \geqslant 3$ we have

$$F = F_\alpha + F_\beta + F_\gamma = \frac{9d^2 + 16d - 17}{6d^2(d^2-1)} \sim \frac{3}{2d^2}. \tag{83}$$

For $d = 2$ the invariant subspace $\mathscr{H}_\gamma$ does not appear and the fidelity becomes $F = F_\alpha + F_\beta = \frac{7}{12}$.

Using Eq. (30) it is possible to derive the form of the replicated POVM corresponding to the optimal generalized instrument:

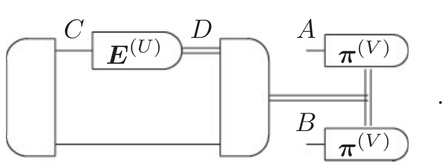
$$\begin{aligned} G_{ii} &= \frac{16d-2}{9d(d^2-1)} P^+\big(E^{(U)}_i \otimes I_B\big)P^+ + \frac{d^2-3}{d(d^2-1)}P^+, \\ G_{ij} &= \frac{1}{d-1}\big[Q'^+\big(E^{(U)}_i \otimes I_B\big)Q'^+ + Q'^-\big(E^{(U)}_j \otimes I_B\big)Q'^-\big] \\ &\quad + \frac{2}{d(d-1)^2(d-2)}P^-\big(E^{(U)}_i \otimes I_B + E^{(U)}_j \otimes I_B\big)P^- \\ &\quad + \frac{d-3}{(d-1)^2(d-2)}P^-, \end{aligned}$$

where $Q'^{\pm} = 1/\sqrt{9d(d-1)}\,(P^+ \pm 3\,P^-)$.

One can now compare the performance of the optimal $1 \to 2$ cloning and learning. The optimal values of $F$ depending on the dimension $d$ are plotted in Fig. 1. As expected, the optimal cloning strategy largely outperforms the optimal learning strategy with a fidelity, which is a factor $d$ larger, as one can see from Eqs. (66) and (83). Similar distinction arises also for comparison of cloning and learning of unitary channels (for details see Ref. [16]).

### A. Estimate and prepare strategy

One can achieve both cloning and learning of a measurement by first estimating the unknown measurement and then constructing the measurement devices according to the results of the estimation [see Eq (84)]:




(84)

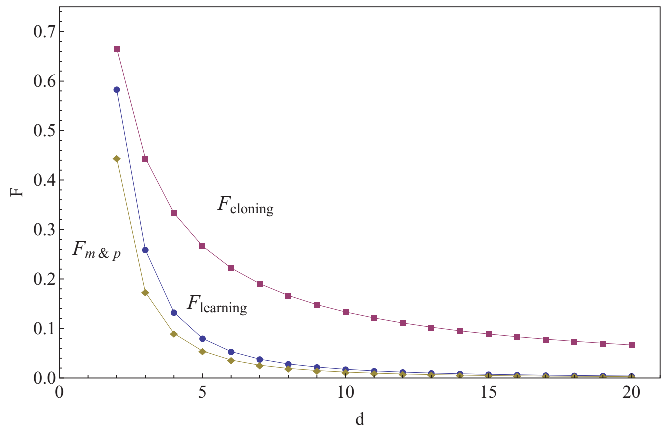



FIG. 1. (Color online) Optimal $1 \to 2$ cloning and learning of a measurement device: we present the values of $F$ for different values of the dimension $d$. The squared dots represent the optimal $1 \to 2$ cloning, the round dots represent the optimal $1 \to 2$ learning and lowest curve corresponds to a strategy in which one performs the optimal estimation followed by the preparation of the estimated measurement.

In the above equation $T_V$ is the quantum network describing the estimation and $\boldsymbol{\pi}^{(V)} \otimes \boldsymbol{\pi}^{(V)}$ are the two replicas of the original POVM $\boldsymbol{E}^{(U)}$ that are prepared depending on the outcome $V$ of the estimation. Since the estimation is done independently on the availability of states $A, B$ estimate and prepare strategy is a special case of learning, which is in turn a special case of cloning of a measurement.

Since the unknown Von Neuman measurements, which are parametrized by unitary operators, are picked randomly according to a Haar measure, we assume covariant estimation procedure, that is, the shift in the parameters of the measurement to be estimated induces the same shift in the probability distribution for the estimate $V \in U(d)$. Mathematically estimation from Fig. 1 is described by a special type of generalized quantum instrument. Its covariance and normalization are expressed by the following equations: $T_V = V T_I V^\dagger$, $\int dV\, T_V = \rho_0 \otimes I_1$, where $T_V \in \mathcal{B}(\mathscr{H}_1 \otimes \mathscr{H}_0)$. We also assume covariance $\pi^{(V)}_i = V \pi^{(I)}_i V^\dagger$ of the POVM $\boldsymbol{\pi}^{(V)}$ that is prepared based on the outcome $V$. These two assumptions allow us to optimize the operator seeds $T_I$ and $\boldsymbol{\pi}^{(I)}$ for the $1 \mapsto 1$ learning of a measurement using the same techniques we used earlier in the paper. As one would expect the optimal seed of the replicated POVM corresponds to a Von Neuman measurement $\pi^{(I)}_i = |i\rangle\langle i|$ and the estimation stage is determined by $T_I = \sum_k |kk\rangle\langle kk|$. For the $1 \mapsto 2$ learning or cloning of the measurement we can decide to use the same estimation stage, but prepare two measurements based on outcome $V$. This strategy is described by a generalized quantum instrument

$$\begin{aligned} R_{ij} &= \int dV \big(\pi^{(V)}_i \otimes \pi^{(V)}_j\big)^T \otimes T_V \\ &= \sum_k \int dV (V^*)^{\otimes 2} \otimes V |ijk\rangle\langle ijk| (V^T)^{\otimes 2} \otimes V^\dagger \otimes |k\rangle\langle k|. \end{aligned} \tag{85}$$

We can now insert the above expression into Eq. (31) and as a consequence of Schur's lemmas for representation $V^* \otimes V^* \otimes V$ we can write the figure of merit as

$$F[\mathcal{R}] = \frac{1}{d^2} \sum_{(ab,c)\in\mathsf{L}} n(ab,c) \sum_{\nu} \frac{1}{d_\nu} \mathrm{Tr}\left[\left(\Delta^{\nu}_{ab,c}\right)^2\right], \quad (86)$$

where we used $\Delta^{\nu}_{ab,c}$ defined in Eq. (53) and the notation introduced after remark 1. The direct evaluation of Eq. (86) gives the following value of fidelity:

$$F_{\mathrm{e\&p}} = \frac{d^2 + 6d + 10}{d^2(d+1)(d+2)} \quad (87)$$

for the estimate and prepare strategy.

## VII. CONCLUSIONS

In the present paper we focused on $1 \to 2$ cloning and $1 \to 2$ learning of von Neumann measurements. Even though both problems can be easily formulated in the usual language of quantum mechanics, the necessity to handle the measurement outcome in the remaining part of the scheme makes the optimization complicated and requires suitable mathematical tools. We represented the unknown measurement to be replicated as a measure and prepare channel and we employed framework of quantum combs to perform the network optimization. Thanks to symmetries of the figure of merit the problem was simplified and solved for arbitrary dimension of the measurement's Hilbert space $d$. For $1 \to 2$ cloning of a measurement we found that the optimal fidelity is $\frac{4}{3d}$, while the optimal fidelity for $1 \to 2$ learning scales as $\frac{3}{2d^2}$ and outperforms just slightly the estimate and prepare strategy in which one first estimates the unknown measurement and prepares the duplicate based on that result. As Fig. 1 suggests, in higher dimensional Hilbert spaces the single use of the unknown measurement device provides just a small piece of information needed for its accurate replication and hence efficient cloning and learning of measurement become impossible. In Sec. V A we proposed a realization of optimal $1 \to 2$ cloning of measurements. The proposed scheme has some similarities to optimal cloning of unitary transformations, since they both begin by the control-SWAP operation, which reflects the presence of quantum parallelism. In this paper we exploit the measure and prepare representation of von Neumann measurement that allowed us to deal with feed forward of classical information in quantum networks. This tool could be in principle used to tackle other quantum information processing tasks in which classical information is involved, for example, estimation and cloning of quantum instruments.


[1] W. K Wootters and W. H Zurek, Nature (London) **299**, 802 (1982).

[2] M. A. Nielsen and I. L. Chuang, Phys. Rev. Lett. **79**, 321 (1997).

[3] V. Bužek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).

[4] R. F. Werner, Phys. Rev. A **58**, 1827 (1998).

[5] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, Phys. Rev. Lett. **76**, 2818 (1996).

[6] G. M. D'Ariano, C. Macchiavello, and P. Perinotti, Phys. Rev. Lett. **95**, 060503 (2005).

[7] V. Scarani, S. Iblisdir, N. Gisin, and A. Acin, Rev. Mod. Phys. **77**, 1225 (2005).

[8] S. Pirandola, S. Mancini, S. Lloyd, and S. L. Braunstein, Nat. Phys. **4**, 726 (2008).

[9] K. Bostroem and T. Felbinger, Phys. Rev. Lett. **89**, 187902 (2002).

[10] M. Lucamarini and S. Mancini, Phys. Rev. Lett. **94**, 140501 (2005).

[11] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 180504 (2008).

[12] A. Ferraro, M. Galbiati, M. G. A. Paris, J. Phys. A **39**, L219 (2006).

[13] A. Bisio, G. M. D'Ariano, P. Perinotti, and M. Sedlak, Phys. Lett. A **375**, 3425 (2011.

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

[15] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **80**, 022339 (2009).

[16] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Rev. A **81**, 32324 (2010).