Optimal processing of reversible quantum channels

Alessandro Bisio [a],[*], Giacomo Mauro D'Ariano [a], Paolo Perinotti [a], Michal Sedlák [b],[c],[*]

[a] *QUIT Group, Dipartimento di Fisica, INFN Sezione di Pavia, via Bassi 6, 27100 Pavia, Italy*
[b] *Department of Optics, Palacký University, 17. Listopadu 1192/12, CZ-771 46 Olomouc, Czech Republic*
[c] *Institute of Physics, Slovak Academy of Sciences, Dúbravská Cesta 9, 845 11 Bratislava, Slovakia*

ARTICLE INFO

ABSTRACT

We consider the general problem of the optimal transformation of N uses of (possibly different) unitary channels to a single use of another unitary channel in any finite dimension. We show how the optimal transformation can be fully parallelized, consisting in a preprocessing channel followed by a parallel action of all the N unitaries and a final postprocessing channel. Our techniques allow to achieve an exponential reduction in the number of the free parameters of the optimization problem making it amenable to an efficient numerical treatment. Finally, we apply our general results to find the analytical solution for special cases of interest like the cloning of qubit phase gates.

1. Introduction

In the past decades, considerable progress has been made in the understanding of the mathematical structure of quantum theory. Recently the view of Quantum Theory as an operational probabilistic theory [1] was revitalized by the success of quantum information theory, which helped framing the operational axiomatization program into an information theoretic context [2–5]. This approach has been a fruitful line of investigation [6,7] and remarkably lead to a derivation of Quantum theory from operational and informational principles [8].

The founding pillar of this view is the basic notion of test, that includes as a special case that of preparation and observation. The second ingredient defining an operational probabilistic theory is provided by the rules for calculating the probabilities of the experimental outcomes. In this perspective, transformations of quantum states can be characterized by the minimal axioms that ensure preservation of the probabilistic structure of quantum theory. Such axioms require a transformation to be linear, trace non-increasing and completely positive, identifying possible events in a test with *quantum operations*, with *quantum channels* representing deterministic ones.

In quantum information applications, not only quantum states but also transformations can often be considered as carriers of information, e.g. in the context of channel discrimination [9–13], gate programming [14], gate teleportation [15–17], process tomography [18,19,32,20] multi-round quantum games [21], standard

quantum algorithms [22–24], as well as cryptographic protocols [25–27]. This approach suggested to extend the Kraus' axiomatic characterization of quantum operations to the case of *higher order quantum maps*, that is quantum maps that transform other quantum maps. The easiest case of higher order quantum map is the *supermap*, that is a map that transforms quantum operations into quantum operations. As a paradigmatic example, one can consider a supermap that, given a single use of a quantum channel $\mathcal{T}$ as an input, produces as an output channel $\mathcal{T}$ followed by a fixed channel $\mathcal{S}$, namely $\mathcal{S} \circ \mathcal{T}$. It is interesting to realize that *higher order quantum computation*, namely the study of higher order quantum maps, is a subject in which the differences between the quantum and the classical world are evident. In classical computation, the Church–Turing paradigm of *program as data* allows one to compute functions of functions, rather than only functions of bits. In the quantum case quantum data, i.e. states, and quantum functions, i.e. quantum transformations, are intrinsically different objects and the exact programming of unitary transformations via quantum states is impossible with finite resources. Thus the study of the properties of higher order maps achieves a twofold objective: on the one hand their mathematical characterization allows one to address in a systematic way all of the quantum processing tasks, and on the other hand it provides new insights in the distinctive features of quantum theory.

Higher order quantum maps were introduced in Refs. [28,29] and a complete axiomatic characterization of a sub-hierarchy of the higher order quantum maps was presented in Ref. [30]. Such a characterization is based on the generalization of the notion of Choi operator to higher order quantum maps. The subclass of higher order maps studied in Ref. [30], the so-called *quantum*

* Corresponding authors.
E-mail address: alessandro.bisio@unipv.it (A. Bisio).

combs, was therein proved to be in correspondence with the set of adaptive quantum strategies, which are the most general architecture allowed in the quantum circuit model. Such a unified description opened the way to the formulation and optimization of a number of quantum processing tasks [31–36]. However, there exist higher order maps which are admissible, i.e. they preserve the probabilistic structure of quantum theory, but cannot be described as a quantum circuit. For example, as pointed out in Ref. [37] the map which receives one use of channel $\mathcal{C}$ and one use of channel $\mathcal{D}$ as input and outputs the convex combination $\frac{1}{2}(\mathcal{C} \circ \mathcal{D} + \mathcal{D} \circ \mathcal{C})$ is not realizable as a quantum circuit. This issue raises two main questions. The first one, which is still completely open, is which non-circuital higher order maps correspond to physically feasible procedures. The second question asks whether there exist any computational tasks in which this non-circuital higher order map can outperform a circuital strategy. As regards this second question it has indeed been proved that non-circuital maps can enhance non-signaling gate discrimination [38] and the programmability of permutations of N different unitary channels [39].

Here we apply the theory of higher order quantum maps to the optimization of a very general class of quantum information processing tasks that can be sketched as follows. Let $\{\mathcal{U}_g^{(i)}\}$ $g \in G$, $i = 1, \ldots, N$, be a set of unitary channels $\mathcal{U}_g^{(i)}(\rho) = U_g^{(i)} \rho U_g^{\dagger(i)}$, where $U_g^{(i)}$ is a unitary representation of a compact group G for each i. Suppose that an unknown element $g \in G$ was chosen randomly according to the Haar measure on G, and conditionally on the outcome g we had access to a single use of each of the channels $\mathcal{U}_g^{(i)}$ $i = 1, \ldots, N$, in any sequential order. In other words, we can choose to use the channels $\mathcal{U}_g^{(i)}$ in the sequence defined by any permutation $\pi(i)$ of the indices i, and we are also free to use some of the channels in parallel, in a single computational step. Our aim is now to approximate as good as we can the channel $\mathcal{V}_g$ defined by a different representation of G. In simple terms we are considering a higher order map which transforms a single use of the channels $\mathcal{U}_g^{(i)}$ into a single use of a channel $\mathcal{V}_g$. Quantum cloning of a unitary transformation is the special case in which $\{\mathcal{U}_g^{(i)}\} = \{\mathcal{U}_g\}$ and $\mathcal{V}_g = \mathcal{U}_g \otimes \mathcal{U}_g$. Since the input consists of more than a single use of a channel, we should in principle allow for non-circuital maps, like the one that can exchange the sequential order of the unitary channels.

In this paper after a review of the main results in higher order quantum computation in Section 2, we will prove in Section 3.1 that, surprisingly, the optimal strategy for the class of tasks considered above is realizable via a simple three steps procedure: (i) application of a fixed preprocessing channel $\mathcal{C}_1$, (ii) parallel action of the unknown channels $\mathcal{U}_g^{(i)}$ on some part of the output of $\mathcal{C}_1$ and (iii) action of a postprocessing channel $\mathcal{C}_2$. This means that there is no need for any kind of non-circuital quantum maps for the purpose of optimization of this kind of task. Thanks to this result and to the symmetries of the problem in Section 3.2 we will show how the optimization of the circuit is reduced to the problem of finding the set of probability distributions p_K^a, $\sum_K p_K^a = 1$ maximizing the function $\Phi(p_K^a) = \sum_K (\sum_a \sqrt{q_K^a p_K^a})^2$, where q_K^a are a set of coefficients determined by the problem that can be efficiently calculated. Once the parameters p_K^a are found, a realization of the optimal strategy can be found by the method of Ref. [40]. The problem addressed in this paper is very general and allows one to optimize wide variety of problems either analytically or by simple numerical optimization. Some examples of application of our results are presented in Section 4. Finally, Section 5 summarizes our conclusions and possible future extensions of the work.

2. Higher order quantum maps

A *quantum supermap* is a transformation in which the input and the output are quantum transformations themselves. In other words, a quantum supermap describes a transformation that receives a quantum operation as an input and produces another quantum operation as an output, with the condition that channels are mapped to channels. More generally one can consider maps whose input and output are themselves supermaps, and the construction can be brought to arbitrarily high order. In this way one obtains a whole hierarchy of maps, the *higher order quantum maps*. In this section we review the general theory of the higher order quantum maps, as presented in Refs. [28–30,41], which we refer to for an extensive discussion and for the proofs of the results reviewed in this section.

The main issue addressed here is the classification of all the input/output transformations that are admissible in principle according to quantum theory. There are essentially two requirements that an input/output map has to satisfy in order to be admissible. The first one is linearity, which is required to be compatible with the probabilistic structure of the theory. For example, let us consider a supermap $\tilde{\mathcal{R}}$ which transforms channels into channels, i.e. $\tilde{\mathcal{R}} : \mathcal{E} \mapsto \tilde{\mathcal{R}}(\mathcal{E})$. If we apply the map $\tilde{\mathcal{R}}$ to the convex combination $p\mathcal{E}_1 + (1 - p)\mathcal{E}_2$-corresponding to a random choice of the input channel—the output has to be the convex combination $p\tilde{\mathcal{R}}(\mathcal{E}_1) + (1 - p)\tilde{\mathcal{R}}(\mathcal{E}_2)$. For the same reason, we should also have $\tilde{\mathcal{R}}(p\mathcal{E}) = p\tilde{\mathcal{R}}(\mathcal{E})$ for any $0 \leqslant p \leqslant 1$. These two conditions together imply that $\tilde{\mathcal{R}}$ can be extended without loss of generality to a linear map. The same reasoning used for supermaps applies to more general higher order maps, which must then be linear at every order. Actually, it is easy to show by induction that every class of higher order quantum maps corresponds to a convex set. The second requirement is that the map must produce a legitimate output when applied locally on one side of a bipartite input. When the input and the output are quantum states this condition is called complete positivity (CP) and the set of the admissible maps is simply the set of the so-called Quantum Operations [42].

Let us now consider supermaps, whose input and output are quantum operations. In order to simplify the presentation we will restrict ourself to the deterministic case, that is maps $\tilde{\mathcal{R}}$ which transform quantum channels into quantum channels. The generalization to the probabilistic case is possible and we refer to [29, 30] for a comprehensive presentation. If $\tilde{\mathcal{R}}$ is an admissible supermap transforming quantum channels with input (output) space $\mathcal{H}_{\text{in},A}(\mathcal{H}_{\text{out},A})$ then the output of $\tilde{\mathcal{R}}$ is a legitimate quantum channel even when $\tilde{\mathcal{R}}$ is applied locally to a bipartite quantum channel, i.e. a quantum channel $\mathcal{E}$ with bipartite input space $\mathcal{H}_{\text{in}} := \mathcal{H}_{\text{in},A} \otimes \mathcal{H}_{\text{in},B}$ and bipartite output space $\mathcal{H}_{\text{out}} := \mathcal{H}_{\text{out},A} \otimes \mathcal{H}_{\text{out},B}$. This means that $\tilde{\mathcal{R}} \otimes \mathcal{I}_B(\mathcal{E})$ is a CP map for any bipartite quantum channel $\mathcal{E}$, $\mathcal{I}_B$ denoting the identity map on the spaces labeled by B.

When dealing with complete positivity it is convenient to use the Choi isomorphism [43] between $\mathcal{L}(\mathcal{L}(\mathcal{H}_{\text{in}}), \mathcal{L}(\mathcal{H}_{\text{out}}))$ and $\mathcal{L}(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}})$, where $\mathcal{L}(\mathcal{H})$ denotes the space of linear operators on the Hilbert space $\mathcal{H}$ and $\mathcal{L}(\mathcal{L}(\mathcal{H}_{\text{in}}), \mathcal{L}(\mathcal{H}_{\text{out}}))$ denotes the space of linear maps from $\mathcal{L}(\mathcal{H}_{\text{in}})$ to $\mathcal{L}(\mathcal{H}_{\text{out}})$. Before presenting the Choi isomorphism we recall the following one to one correspondence between $\mathcal{L}(\mathcal{H})$ and $\mathcal{H} \otimes \mathcal{H}$:

$$A = \sum_{nm} \langle n|A|m\rangle |n\rangle\langle m| \leftrightarrow |A\rangle\rangle = \sum_{nm} \langle n|A|m\rangle |n\rangle |m\rangle$$

$$A \otimes B |C\rangle\rangle = |ACB^T\rangle\rangle, \tag{1}$$

where $|n\rangle$ denotes a fixed orthonormal basis for $\mathcal{H}$ and A^T denotes transposition of A with respect to the fixed orthonormal basis (A^* denotes complex conjugation with respect to the same basis).

Proposition 1 (*Choi isomorphism*). *Let $\mathfrak{C}$ be a linear map from $\mathcal{L}(\mathcal{L}(\mathcal{H}_{\rm in}), \mathcal{L}(\mathcal{H}_{\rm out}))$ to $\mathcal{L}(\mathcal{H}_{out} \otimes \mathcal{H}_{in})$ defined as follows:*

$$\mathfrak{C}(\mathcal{C}) := \mathcal{C} \otimes \mathcal{I}(|I\rangle\rangle\langle\langle I|), \tag{2}$$

where $|I\rangle\rangle \in \mathcal{H}_{\rm in} \otimes \mathcal{H}_{\rm in}$. Then $\mathfrak{C}$ is invertible and its inverse map is defined as

$$\left[\mathfrak{C}^{-1}(C)\right](\rho) := \mathrm{Tr}_{\rm in}\left[(I_{\rm out} \otimes \rho^T)C\right], \tag{3}$$

where $\mathrm{Tr}_{\rm in}$ denotes the partial trace over $\mathcal{H}_{\rm in}$ and $I_{\rm out}$ denotes the identity operator over $\mathcal{H}_{\rm out}$. The operator $C := \mathfrak{C}(\mathcal{C})$ is called the Choi operator *of the map $\mathcal{C}$.*

For the special case of a unitary channel $\mathcal{Z} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_1)$, $\mathcal{Z}(\rho) := Z\rho Z^\dagger$ Eq. (2) and Eq. (1) give

$$\mathfrak{C}(\mathcal{Z}) = \mathcal{Z} \otimes \mathcal{I}(|I\rangle\rangle\langle\langle I|) = (Z \otimes I)|I\rangle\rangle\langle\langle I|(Z^\dagger \otimes I)$$
$$= |Z\rangle\rangle\langle\langle Z| \quad |Z\rangle\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_0. \tag{4}$$

By means of the Choi isomorphism, for any map $\tilde{\mathcal{R}}$ that transforms linear maps $\mathcal{E} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ to linear maps $\mathcal{E}' : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_3)$ we can introduce the conjugate map $\mathcal{R}$ defined as follows:

$$\mathcal{R} := \mathfrak{C} \circ \tilde{\mathcal{R}} \circ \mathfrak{C}^{-1}, \tag{5}$$

that transforms the Choi operator E of $\mathcal{E}$ into the Choi operator E' of $\mathcal{E}'$. It is possible to show [29] that the admissibility conditions for $\tilde{\mathcal{R}}$ are equivalent to linearity and complete positivity of $\mathcal{R}$. Moreover, since $\mathcal{R}$ is a linear map from $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ to $\mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_3)$ we can apply the Choi isomorphism and introduce its Choi operator R. For the sake of simplicity we will systematically use the map $\mathcal{R}$ instead of $\tilde{\mathcal{R}}$ and denote by R the corresponding Choi operator. Within this framework we can give the following formal definition of a higher order map.

Definition 1. A *1-comb* on $(\mathcal{H}_0, \mathcal{H}_1)$ is the Choi operator of a linear CP map from $\mathcal{L}(\mathcal{H}_0)$ to $\mathcal{L}(\mathcal{H}_1)$. A *probabilistic 1-comb* is a 1-comb corresponding to a quantum operation, and a *deterministic 1-comb* is a 1-comb corresponding to a quantum channel. For $N \geqslant 2$, an *N-comb* $R^{(N)}$ $(\mathcal{H}_0, \dots, \mathcal{H}_{2N-1})$ is the Choi operator of an *admissible N-map*, i.e. a linear completely positive map $\mathcal{R}^{(N)}$ that transforms $(N-1)$-combs on $(\mathcal{H}_1, \dots, \mathcal{H}_{2N-2})$ into 1-combs on $(\mathcal{H}_0, \mathcal{H}_{2N-1})$. A *deterministic N-comb* is an N-comb corresponding to a map that transforms deterministic $(N-1)$-combs to deterministic 1-combs. For $N, M \geqslant 1$ an *(N, M)-comb* is the Choi operator of an *admissible (N, M)-map*, i.e. a linear completely positive map $\mathcal{R}^{(N,M)}$ that transforms N-combs into M-combs. A *deterministic (N, M)-comb* is an (N, M)-comb corresponding to an (N, M)-map that transforms deterministic N-combs to deterministic M-combs. An (N, M)-comb S such that $S \leqslant \bar{S}$ for a deterministic (N, M)-comb $\bar{S}$ is called *probabilistic*.

Notice that $N + 1$-combs can be also denoted as $(N, 1)$-combs. By recursively applying Definition 1 one can define *admissible (x, y)-maps* where x and y are previously defined map types, thus creating the whole hierarchy of higher order maps. Also in this case deterministic and probabilistic (x, y)-combs can be straightforwardly defined.

In Definition 1 we defined N-combs as operators $R^{(N)}$ acting on an ordered sequence of Hilbert spaces $\bigotimes_{k=0}^{2N-1} \mathcal{H}_k$. Such a labeling can be done by exploiting the following diagrammatic representation of quantum combs

$$\tag{6}$$
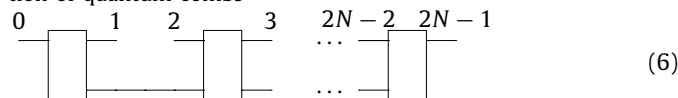
where an N-comb is represented by a comb-like diagram with N teeth.

The following proposition provides an algebraic characterization of the set of deterministic N-combs.

Proposition 2. *A positive operator $R^{(N)}$ on $\bigotimes_{k=0}^{2N-1} \mathcal{H}_k$ is a deterministic N-comb if and only if the following conditions hold:*

$$\mathrm{Tr}_{2j-1}\left[R^{(j)}\right] = I_{2j-2} \otimes R^{(j-1)}, \quad 2 \leqslant j \leqslant N$$
$$\mathrm{Tr}_1\left[R^{(1)}\right] = I_0, \tag{7}$$

*where $R^{(j-1)}, 2 \leqslant j \leqslant N$ are deterministic $(j-1)$-combs.*

Proposition 2 characterizes the set of deterministic N-combs as the set of positive operators subject to the linear constraints of Eq. (7). This implies that the set of deterministic N-combs is a convex set. It is possible to provide a generalization of Proposition 2 to (N, M)-maps and to all the other classes of higher order maps, but this is beyond the main scope of this paper and we will omit it. However, let us remind that each set of deterministic higher order maps is a convex set.

So far we focused our analysis on the mathematical description of the higher order quantum maps which culminated in Proposition 2, which translates the admissibility conditions of linearity and complete positivity in terms of algebraic constraints. However, such a characterization would be just an abstract and rather sterile construction if it was not related to physical achievability of the involved maps. In the following we will show that any admissible deterministic N-map has a physical realization as a concatenation of channels with multipartite input and output.

When considering channels whose input and output spaces are tensor products of Hilbert spaces it is possible to define the composition of these channels only through some of these spaces. For example, if we have $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_2), \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_3))$ and $\mathcal{D} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_5), \mathcal{L}(\mathcal{H}_4 \otimes \mathcal{H}_6))$ it is possible to define the composition

$$\mathcal{D} \star \mathcal{E} := (\mathcal{D} \otimes \mathcal{I}_1) \circ (\mathcal{E} \otimes \mathcal{I}_5), \tag{8}$$

where $\mathcal{D} \star \mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_2 \otimes \mathcal{H}_5), \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_4 \otimes \mathcal{H}_6))$. It can be diagrammatically represented as follows:

$$\tag{9}$$
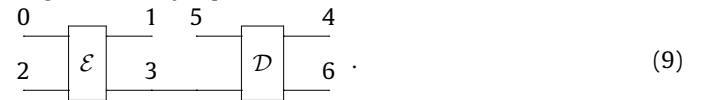
Moreover, here the similarity with Eq. (6) is not a coincidence as it will be clear later. Since the two channels can be represented in terms of their Choi operators one can reasonably wonder how the Choi operator of the composition $\mathcal{D} \star \mathcal{E}$ can be expressed in terms of the Choi operators D and E. For this purpose it is convenient to define the following operation.

Definition 2. Let M be an operator in $\mathcal{L}(\bigotimes_{i \in \mathsf{I}} \mathcal{H}_i)$ and N be an operator in $\mathcal{L}(\bigotimes_{j \in \mathsf{J}} \mathcal{H}_j)$ where I and J are two finite sets of indexes. Then the *link product* $M * N$ is an operator in $\mathcal{L}(\mathcal{H}_{\mathsf{I}\backslash\mathsf{J}} \otimes \mathcal{H}_{\mathsf{J}\backslash\mathsf{I}})$ defined as

$$M * N := \mathrm{Tr}_{\mathsf{I}\cap\mathsf{J}}\left[(I_{\mathsf{J}\backslash\mathsf{I}} \otimes M^{T_{\mathsf{I}\cap\mathsf{J}}})(I_{\mathsf{I}\backslash\mathsf{J}} \otimes N)\right] \tag{10}$$

where $\mathsf{A} \backslash \mathsf{B} := \{i \in \mathsf{A} | i \notin \mathsf{B}\}$ and we introduced the notation $\mathcal{H}_{\mathsf{A}} := \bigotimes_{i \in \mathsf{A}} \mathcal{H}_i$ for any set of indexes A.

It is worth noting that the link product is commutative, i.e. $M * N = N * M$ (here we assume the same ordering of the tensor products of Hilbert spaces). Moreover, the special case $\mathsf{I} \cap \mathsf{J} = \emptyset$ gives $N * M = N \otimes M$ while if $\mathsf{I} = \mathsf{J}$ $N * M = \mathrm{Tr}[M^T N]$. The use of

the link product simplifies the expression for the Choi operator of the composition of two channels, which is the content of the following Lemma.

Lemma 1. *Let* $\text{in}_\mathcal{E}$, $\text{out}_\mathcal{E}$, $\text{in}_\mathcal{D}$, $\text{out}_\mathcal{D}$ *be four sets of indexes such that* $\text{in}_\mathcal{E} \cap \text{out}_\mathcal{D} = \emptyset$. *Let* $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_{\text{in}_\mathcal{E}}), \mathcal{L}(\mathcal{H}_{\text{out}_\mathcal{E}}))$ *and* $\mathcal{D} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_{\text{in}_\mathcal{D}}),$ $\mathcal{L}(\mathcal{H}_{\text{out}_\mathcal{D}}))$ *be a couple of quantum channels. Let E and D be Choi operators of the quantum channels* $\mathcal{E}$ *and* $\mathcal{D}$. *Then the Choi operator of the composition*

$$\mathcal{D} \star \mathcal{E} := (\mathcal{I}_{\text{out}_\mathcal{E} \setminus \text{in}_\mathcal{D}} \otimes \mathcal{D}) \circ (\mathcal{I}_{\text{in}_\mathcal{D} \setminus \text{out}_\mathcal{E}} \otimes \mathcal{E}) \tag{11}$$

is given by

$$\mathfrak{C}(\mathcal{D} \star \mathcal{E}) = D * E \tag{12}$$

where $D * E$ *is the link product of the two operators.*

For sake of clarity, it is useful to apply Lemma 1 to the simple case of two unitary channels

$$\overset{1}{\rule{1em}{0.4pt}}\boxed{\mathcal{U}}\overset{2}{\rule{1em}{0.4pt}}\boxed{\mathcal{V}}\overset{3}{\rule{1em}{0.4pt}}, \tag{13}$$

where $\mathcal{U} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ and $\mathcal{V} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_2), \mathcal{L}(\mathcal{H}_3))$. Reminding Eq. (4) the Choi operators of $\mathcal{U}$ and $\mathcal{V}$ are given by $|U\rangle\!\rangle\langle\!\langle U| \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_1)$ and $|V\rangle\!\rangle\langle\!\langle V| \in \mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_2)$, respectively. By applying Eq. (12) we have

$$\begin{aligned}
\mathfrak{C}(\mathcal{U} \star \mathcal{V}) &= |U\rangle\!\rangle\langle\!\langle U| * |V\rangle\!\rangle\langle\!\langle V| \\
&= \text{Tr}_2\big[\big(|U\rangle\!\rangle\langle\!\langle U| \otimes I_3\big)\big(I_1 \otimes (|V\rangle\!\rangle\langle\!\langle V|)^{T_2}\big)\big] \\
&= \text{Tr}_2\big[\big(|U\rangle\!\rangle\langle\!\langle U| \otimes I_3\big)\big(I_1 \otimes |V^*\rangle\!\rangle\langle\!\langle V^*|\big)\big] \\
&= |UV\rangle\!\rangle\langle\!\langle UV| = \mathfrak{C}(\mathcal{U} \circ \mathcal{V})
\end{aligned} \tag{14}$$

where we used Eq. (2).

Lemma 1 can be applied to the case in which N quantum channels are connected in a sequence, i.e.

$$\mathcal{R} = \mathcal{E}_1 \star \mathcal{E}_2 \star \cdots \star \mathcal{E}_N$$

$$\overset{0}{\rule{0pt}{0pt}}\boxed{\mathcal{E}_1}\overset{1}{\underset{}{}}\boxed{A_1}\overset{2}{\rule{0pt}{0pt}}\boxed{\mathcal{E}_2}\overset{3}{}\boxed{A_2} \cdots \overset{2N-2}{}\boxed{A_{N-1}}\overset{2N-1}{}\boxed{\mathcal{E}_N} \tag{15}$$

where $\mathcal{E}_i : \mathcal{L}(\mathcal{H}_{2i-2} \otimes \mathcal{H}_{A_{i-1}}) \to \mathcal{L}(\mathcal{H}_{2i-1} \otimes \mathcal{H}_{A_i})$, $\mathcal{H}_{A_0} = \mathcal{H}_{A_N} = \mathbb{C}$ and the ordering in which the connections are performed can be proved to be irrelevant. Moreover the Choi operator of the sequence $\mathcal{R} = \mathcal{E}_1 \star \cdots \star \mathcal{E}_N$ becomes

$$\mathfrak{C}(\mathcal{R}) := R = E_1 * \cdots * E_N \tag{16}$$

and also in this case the order in which the link products are performed is not relevant. It is possible to prove [30] that Eq. (16) implies that the Choi operator of a sequence of channels satisfies conditions (7). Moreover, Eq. (7) is a sufficient condition for $R^{(N)}$ to be the Choi operator of a sequence of quantum channels. It is then possible to identify the set of admissible deterministic N-maps with the set of maps that are given by the concatenation of N channels.

Proposition 3. *Let* $\mathcal{R}^{(N)}$ *be a linear map and* $R^{(N)}$ *its Choi operator. Then the following conditions are equivalent:*

- $R^{(N)}$ *is a deterministic N-comb,*
- *there exist N quantum channels* $\mathcal{E}_i : \mathcal{L}(\mathcal{H}_{2i-2} \otimes \mathcal{H}_{A_{i-1}}) \to \mathcal{L}(\mathcal{H}_{2i-1} \otimes \mathcal{H}_{A_i})$, $(\mathcal{H}_{A_0} = \mathcal{H}_{A_N} = \mathbb{C})$, $i = 1, \ldots, N$ *such that* $\mathcal{R}^{(N)} = \mathcal{E}_1 \star \cdots \star \mathcal{E}_N$.

Moreover, for any deterministic $(N-1)$-comb $T^{(N-1)}$ the transformation

$$\mathcal{R}^{(N)} : \mathcal{T}^{(N-1)} \mapsto \mathcal{T}'^{(1)} := \mathcal{R}^{(N)}(\mathcal{T}^{(N-1)})$$

is achieved by connecting the two sequences of channels as follows

$$\mathcal{R}^{(N)}(\mathcal{T}^{(N-1)}) := \mathcal{R}^{(N)} \star \mathcal{T}^{(N-1)}$$

$$= \overbrace{\underbrace{\boxed{\mathcal{E}_1}\boxed{\mathcal{D}_1}\boxed{\mathcal{E}_2}\cdots\boxed{\mathcal{D}_{N-1}}\boxed{\mathcal{E}_N}}_{\mathcal{R}^{(N)}}}^{\mathcal{T}^{(N-1)}} \tag{17}$$

and the Choi operator of the resulting map is given by:

$$\mathfrak{C}(\mathcal{R}^{(N)} \star \mathcal{T}^{(N-1)}) = R^{(N)} * T^{(N-1)}. \tag{18}$$

Proposition 3 shows that any admissible deterministic N-map has a physical realization as a concatenation of quantum channels and tells us through Eq. (18) that the action of an admissible N-map on an $N-1$-map can be expressed by the link product of the corresponding Choi operators.

Unfortunately, the more general case of (N, M)-maps or (x, y)-maps is more involved. Eq. (18) still holds, but it is no longer possible to interpret (N, M)-maps or (x, y)-maps as sequences of channels.

The following lemma can be regarded as a quantum generalization of the uncurrying procedure of the functional calculus and provides some useful insight on the features of the deterministic (N, M) maps.

Lemma 2. *Let* $\mathcal{R}^{(N,M+1)}$ *be an admissible deterministic $(N, M + 1)$-map. Then* $\mathcal{R}^{(N,M+1)}$ *is in one-to-one correspondence with an admissible map* $\mathcal{R}^{(N \otimes M, 1)}$ *that transforms tensor product operators* $S^{(N)} \otimes T^{(M)}$ *of deterministic N and M-combs into deterministic 1-combs.*

Intuitively, the tensor product comb $S^{(N)} \otimes T^{(M)}$ can be seen as couple of combs, one with N teeth and the other with M teeths, which create an $N + M$-comb where the order between two teeth of different comb is not completely fixed, but only restricted by the two orderings of the combs $S^{(N)}$ and $T^{(M)}$. Here follows a pictorial example

$$S^{(2)} \otimes T^{(2)} = \boxed{} \text{ or } \boxed{} \cdots$$

This feature can be rephrased by saying that the tensor product comb $S^{(N)} \otimes T^{(M)}$ is not endowed with a full definite causal order. An admissible map $\mathcal{R}^{(N \otimes M, 1)}$ can in principle exploit this freedom and convex combination or quantum superposition of different causal orderings are allowed, like e.g.

$$C \otimes D = \boxed{\mathcal{C}} \quad \boxed{\mathcal{D}} \text{ or } \boxed{\mathcal{D}} \quad \boxed{\mathcal{C}}$$

$$R^{(1 \otimes 1, 1)} * (C \otimes D) = \frac{1}{2}\boxed{\mathcal{D}}\boxed{\mathcal{C}} + \frac{1}{2}\boxed{\mathcal{C}}\boxed{\mathcal{D}}. \tag{19}$$

It is possible to prove that the admissible $(1 \otimes 1, 1)$-map defined by Eq. (19), cannot be realized as a concatenation of channels. In Ref. [37] the first example of an admissible deterministic (N, M)-map that cannot be realized as a sequence of channels, has been found. Even if an (N, M)-map $R^{(N,M)}$ does not correspond to a sequence of channels this does not imply that $R^{(N,M)}$ is not physically realizable. The $(1 \otimes 1, 1)$-map in Eq. (19) receives in input one use of channel $\mathcal{C}$ and one use of channel $\mathcal{D}$ and

outputs either $\mathcal{C} \circ \mathcal{D}$, or $\mathcal{D} \circ \mathcal{C}$ with probability $\frac{1}{2}$: this is clearly a well defined operational procedure. The characterization of admissible (N, M)-maps that are not a sequence of channels, but nevertheless are physically realizable, is still an open problem. Recently (N, M)-maps have been studied in Refs. [39,38] where it was shown that (N, M)-maps can enhance information processing tasks like controlled permutation of oracle gates or discrimination between no-signaling channels. Also the analysis of quantum correlations without a pre-defined causal order in Ref. [44] can be appropriately phrased in terms of (N, M)-maps.

3. Processing of unitary transformations

An example of a task which one can naturally address in the framework of the higher order quantum maps, is cloning of a transformation. This problem was for the first time introduced in Ref. [31] and can be illustrated as follows. Consider a user who is provided with a single use of an unknown transformation $\mathcal{T}$. Suppose now that he needs to run $\mathcal{T}$ twice in order to accomplish some desired computational task. Then it would be extremely valuable for him to have a "cloner of transformations" available. Such a cloner would be a machine which receives a single use of the transformation $\mathcal{T}$ as an input and outputs two copies of the same transformation, i.e. $\mathcal{T} \otimes \mathcal{T}$. In Ref. [31] a no-cloning theorem for transformations was proved and the optimal cloning map for the case in which the unknown transformation is a unitary in $SU(d)$ was derived. The optimal cloner is an admissible deterministic 2-map (see Definition 1) which thanks to Proposition 3 is realizable as concatenation of channels.

In this section we consider a more general scenario which nevertheless is closely related to the cloning of a unitary transformation. Suppose that a user can have access to N unknown unitary channels $\{\mathcal{U}_g^{(i)}\}_{i=1\ldots N}$, where $\mathcal{U}_g^{(i)} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_{2i-1}), \mathcal{L}(\mathcal{H}_{2i}))$ and we denote by d_k the dimension of $\mathcal{H}_k$.[1] The action of $\mathcal{U}_g^{(i)}$ on a state ρ is described by a unitary representation $U_g^{(i)}$ of a fixed compact group G, i.e. $\mathcal{U}_g^{(i)}(\rho) = U_g^{(i)} \rho U_g^{(i)\dagger}$.

The task is to exploit the uses of the unitary channels $\{\mathcal{U}_g^{(i)}\}_{i=1\ldots N}$ to create a target unitary channel $\mathcal{V}_g : \mathcal{L}(\mathcal{H}_0) \rightarrow \mathcal{L}(\mathcal{H}_{2N+1})$ which is described by a different unitary representation V_g of the same group G. The special case in which $\mathcal{U}_g^{(i)} = \mathcal{U}_g$ $\forall i = 1, \ldots, N$, and $\mathcal{V}_g = \mathcal{U}_g^{\otimes M}$ corresponds to the cloning of a unitary transformation $\mathcal{U}_g$ from N copies to M copies. Since we are dealing with a transformation from a tensor product of N channels to a single channel, the goal is to find the admissible deterministic $(1^{\otimes N}, 1)$-map $\mathcal{R}$ which most faithfully realizes the transformation $\bigotimes_{i=1}^N \mathcal{U}_g^{(i)} \rightarrow \mathcal{V}_g$. This can be expressed in terms of Choi operators as

$$R * \bigotimes_{i=1}^{N} |U_g^{(i)}\rangle\!\rangle\langle\!\langle U_g^{(i)}| \simeq |V_g\rangle\!\rangle\langle\!\langle V_g| \tag{20}$$

where $R \in \mathcal{L}(\bigotimes_{k=0}^{2N+1} \mathcal{H}_k)$ is a deterministic $(1^{\otimes N}, 1)$-comb and we used Eq. (4) and Eq. (18). It is worth stressing that, as we mentioned in Section 2, such $\mathcal{R}$ does not necessarily have a realization as a quantum circuit. We now need a criterion to quantify how close the channel $R * \bigotimes_{i=1}^N |U_g^{(i)}\rangle\!\rangle\langle\!\langle U_g^{(i)}|$ is to the target $|V_g\rangle\!\rangle\langle\!\langle V_g|$. The closeness between two channels $\mathcal{C}, \mathcal{D} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1))$ can

be expressed in terms of the channel fidelity [45], that is defined as follows

$$f(C, D) := \frac{1}{d_0^2} \left(\text{Tr}\left[\sqrt{\sqrt{C} D \sqrt{C}} \right] \right)^2, \tag{21}$$

where C and D are the Choi operators of the channels. As a figure of merit for our task we use the channel fidelity between $R * \bigotimes_{i=1}^N |U_g^{(i)}\rangle\!\rangle\langle\!\langle U_g^{(i)}|$ and $|V_g\rangle\!\rangle\langle\!\langle V_g|$ uniformly averaged over the unknown unitaries,[2] that is

$$F(R) = \int \mathrm{d}g \, f\left(R * \bigotimes_{i=1}^{N} |U_g^{(i)}\rangle\!\rangle\langle\!\langle U_g^{(i)}|, |V_g\rangle\!\rangle\langle\!\langle V_g| \right)$$

$$= \frac{1}{d_0^2} \int \mathrm{d}g \, \text{Tr}\left[R\left(|U_g^*\rangle\!\rangle\langle\!\langle U_g^*| \otimes |V_g\rangle\!\rangle\langle\!\langle V_g| \right) \right]$$

$$U_g := \bigotimes_{i=1}^{N} U_g^{(i)}. \tag{22}$$

The problem we address consists of finding a deterministic $(1^{\otimes N}, 1)$-comb R that maximizes the function F in Eq. (22) i.e.

$$\underset{R}{\text{maximize}} \quad F(R)$$

subject to $\quad R$ is a deterministic $(1^{\otimes N}, 1)$-comb. $\tag{23}$

Eq. (23) can be formulated as a *semidefinite program*, namely a problem that can be phrased as

$$\underset{\rho}{\text{maximize}} \quad \text{Tr}[\rho X]$$

subject to $\quad \mathcal{F}(\rho) \leqslant Y$

$$\rho \geqslant 0, \tag{24}$$

where $X \in \mathcal{L}_{\mathsf{H}}(\mathcal{H})$, $Y \in \mathcal{L}_{\mathsf{H}}(\mathcal{K})$, $\mathcal{F} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$, and $\mathcal{L}_{\mathsf{H}}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ denotes the space of Hermitian operators on $\mathcal{H}$, and the map $\mathcal{F}$ is required to be Hermitian-preserving. The fact that the constraint "R is a deterministic $(1^{\otimes N}, 1)$-comb" in Eq. (23) involves equalities while the constraint in Eq. (24) is given by the inequality $\mathcal{F}(\rho) \leqslant Y$ does not represent a problem. Indeed, one can easily see that for any probabilistic $(1^{\otimes N}, 1)$-comb R there exists a deterministic one $\bar{R}$ such that $F(\bar{R}) \geqslant F(R)$. For this reason, we can replace the optimization problem of Eq. (23) with the following one

$$\underset{R}{\text{maximize}} \quad F(R)$$

subject to $\quad R$ is a probabilistic $(1^{\otimes N}, 1)$-comb

that is equivalent to a semidefinite programming in the form of Eq. (24).

In the next subsections we will see that by exploiting symmetries it is possible to radically simplify the problem, reducing it to a much simpler semidefinite program.

3.1. Optimality of the parallel strategy

As we discussed at the end of Section 2 the set of admissible $(1^{\otimes N}, 1)$-maps includes mathematical objects that currently lack a physical interpretation. Before dealing with the optimization problem it is good to know whether the map which maximizes Eq. (22) is known to be realizable in the physical world. In this subsection we prove that the symmetries of the problem allow us to choose the optimal map $\mathcal{R}$ to be a deterministic 2-comb. This fact by Proposition 3 implies that $\mathcal{R}$ can be realized as a concatenation of channels and the task can be optimally accomplished using quantum circuits. We start by proving the following lemma.

---

[1] Since $\mathcal{H}_{2i-1}$ and $\mathcal{H}_{2i}$ must have the same dimension, there exists a natural isomorphism $T^{(i)} : \mathcal{H}_{(2i-1)} \rightarrow \mathcal{H}_{2i}$ connecting their canonical bases $T^{(i)}|n\rangle_{2i-1} = |n\rangle_{2i}$. For the sake of simplicity, in the following for every operator $A : \mathcal{H}_{(2i-1)} \rightarrow \mathcal{H}_{2i-1}$ we will denote the four operators A, $T^{(i)} A T^{(i)\dagger}$, $A T(i)^\dagger$ and $T^{(i)} A$ by the same symbol A whenever the actual input and output spaces of the operators will be clear from the context.

[2] The hypothesis that G is a compact group guarantees the existence of an invariant Haar measure $\mathrm{d}g$.

Lemma 3. *The optimal deterministic $(1^{\otimes N}, 1)$-comb $R \in \mathcal{L}(\bigotimes_{k=0}^{2N+1} \mathcal{H}_k)$ which maximizes Eq. (22) can be assumed without loss of generality to satisfy the commutation relation*

$$\left[R, W_h^* \otimes W_g\right] = 0 \quad \forall g, h \in G, \tag{25}$$

where $W_h = U_h^* \otimes V_h$ with $U_h \in \bigotimes_{i=1}^N \mathcal{L}(\mathcal{H}_{2i-1})$, and $V_h \in \mathcal{L}(\mathcal{H}_0)$, and $W_g = U_g^* \otimes V_g$ with $U_g \in \bigotimes_{i=1}^N \mathcal{L}(\mathcal{H}_{2i})$, $V_g \in \mathcal{L}(\mathcal{H}_{2N+1})$.[3]

Proof. The proof follows the Holevo averaging argument for covariant estimation [46]. Let R be optimal. Then consider the operator

$$\tilde{R} := \int dg\, dh \left(W_h^* \otimes W_g\right) R \left(W_h^T \otimes W_g^\dagger\right).$$

The set of deterministic $(1^{\otimes N}, 1)$-combs is a convex set, hence $\tilde{R}$ is a well defined deterministic $(1^{\otimes N}, 1)$-comb. One can easily verify that $\tilde{R}$ satisfies Eq. (25) and $F(R) = F(\tilde{R})$. □

Lemma 3 is the key ingredient for proving the following proposition.

Proposition 4. *Let R be a $(1^{\otimes N}, 1)$-comb in $\mathcal{L}(\bigotimes_{k=0}^{2N+1} \mathcal{H}_k)$ which obeys the commutation relation (25). Then there exist a deterministic 2-comb R' formed by channels $\mathcal{C}_1 : \mathcal{L}(\mathcal{H}_0) \rightarrow \mathcal{L}(\bigotimes_{i=1}^N \mathcal{H}_{2i-1} \otimes \mathcal{H}_M)$, and $\mathcal{C}_2 : \mathcal{L}(\bigotimes_{i=1}^N \mathcal{H}_{2i} \otimes \mathcal{H}_M) \rightarrow \mathcal{L}(\mathcal{H}_{2N+1})$, such that*

$$R * |U_g\rangle\!\rangle\langle\!\langle U_g| = R' * |U_g\rangle\!\rangle\langle\!\langle U_g|$$
$$= \mathcal{C}_1 * |U_g\rangle\!\rangle\langle\!\langle U_g| * \mathcal{C}_2 \quad \forall g \in G, \tag{26}$$

where $R' = \mathcal{C}_1 * \mathcal{C}_2$ and the link is performed on $\mathcal{H}_M$.

Proof. Let R be a $(1^{\otimes N}, 1)$-comb in $\mathcal{L}(\bigotimes_{k=0}^{2N+1} \mathcal{H}_k)$ and let us define $\mathcal{H}_A = \bigotimes_{i=1}^N \mathcal{H}_{2i-1}$ and $\mathcal{H}_B = \bigotimes_{i=1}^N \mathcal{H}_{2i}$. With this notation we have $R \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_{2N+1})$ and $U_g \in \mathcal{L}(\mathcal{H}_B)$. Let us consider the operator $S \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$ defined as $S := \text{Tr}_{2N+1}[R]$. Upon introducing auxiliary Hilbert spaces $\mathcal{H}_{0'} \equiv \mathcal{H}_0$, $\mathcal{H}_{A'} \equiv \mathcal{H}_A$ and $\mathcal{H}_{B'} \equiv \mathcal{H}_B$, it is possible to define the rank one operator $|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$, where $|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| = (S^{\frac{1}{2}} \otimes I_E)|I\rangle\!\rangle\langle\!\langle I|(S^{\frac{1}{2}} \otimes I_E)$ and we also defined $\mathcal{H}_E := \mathcal{H}_{0'} \otimes \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$, and identity I_E on $\mathcal{H}_E$. The commutation (25) implies $[S^{\frac{1}{2}}, I_0 \otimes I_A \otimes U_g] = 0$, and together with Eqs. (1), (4) and (10) we have

$$|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * |U_g\rangle\!\rangle\langle\!\langle U_g|$$
$$= U_g \left(|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * |I\rangle\!\rangle\langle\!\langle I|\right) U_g^\dagger, \tag{27}$$

where $|U_g\rangle\!\rangle, |I\rangle\!\rangle \in \mathcal{H}_B \otimes \mathcal{H}_A$ and in the last line $U_g \in \mathcal{L}(\mathcal{H}_{B'})$. From the definition of a $(1^{\otimes N}, 1)$-comb we have that $R * |I\rangle\!\rangle\langle\!\langle I|$ is a channel from $\mathcal{L}(\mathcal{H}_0)$ to $\mathcal{L}(\mathcal{H}_{2N+1})$ and then $\text{Tr}_{2N+1}[R * |I\rangle\!\rangle\langle\!\langle I|] = I_0$. From this relation, from $\text{Tr}_E[|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}|] = S$ and from the definition of S we have

$$\text{Tr}_E\left[|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * |I\rangle\!\rangle\langle\!\langle I|\right]$$
$$= S * |I\rangle\!\rangle\langle\!\langle I| = \text{Tr}_{2N+1}\left[R * |I\rangle\!\rangle\langle\!\langle I|\right] = I_0. \tag{28}$$

Denoting by C_1 the CP map with Choi operator $C_1 := |S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * |I\rangle\!\rangle\langle\!\langle I|$, one can easily realize that by virtue of Eq. (28) C_1 is a

channel from $\mathcal{L}(\mathcal{H}_0)$ to $\mathcal{L}(\mathcal{H}_E)$. Eq. (27) can be diagrammatically represented as

$$U_g C_1 U_g^\dagger = C_1 * |U_g\rangle\!\rangle\langle\!\langle U_g| = \boxed{\begin{array}{c} 0 \\ C_1 \end{array} \quad \begin{array}{c} B' \\ \boxed{\mathcal{U}_g} \\ M \end{array} \quad B'}, \tag{29}$$

where we defined $\mathcal{H}_M := \mathcal{H}_{0'} \otimes \mathcal{H}_{A'}$.

Let us now introduce the operator $C_2 := T S^{-\frac{1}{2}} R S^{-\frac{1}{2}} T^\dagger + T(I - \Pi_S) T^\dagger \otimes \frac{1}{d_{2N+1}} I_{2N+1}$, where T is the isomorphism between $\mathcal{H}_0 \otimes \mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{H}_{0'} \otimes \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$, and $\Pi_S := S^{-\frac{1}{2}} S S^{-\frac{1}{2}}$ is the projection on the support of S. Since we have that $\text{Tr}_{2N+1}[T S^{-\frac{1}{2}} R S^{-\frac{1}{2}} T^\dagger] = T \Pi_S T^\dagger$, it is easily verified that C_2 is a channel from $\mathcal{H}_E$ to $\mathcal{H}_{2N+1}$. By direct computation one can verify that

$$|S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * C_2 = R, \tag{30}$$

where the link is performed on the Hilbert space $\mathcal{H}_E$. Combining Eqs. (27), (29) and (30) and exploiting commutativity and associativity of the link product we get

$$R * |U_g\rangle\!\rangle\langle\!\langle U_g| = |S^{\frac{1}{2}}\rangle\!\rangle\langle\!\langle S^{\frac{1}{2}}| * C_2 * |U_g\rangle\!\rangle\langle\!\langle U_g|$$
$$= C_1 * |U_g\rangle\!\rangle\langle\!\langle U_g| * C_2.$$

which is Eq. (26) up to relabeling of Hilbert spaces. □

Proposition 4 tells us that the optimal transformation from a set of unitary transformations $\{\mathcal{U}_g^{(i)}\}_{i=1\dots N}$ to a target unitary channels $\mathcal{V}_g$ is physically realizable with the following scheme:

i. application of a preprocessing channel $\mathcal{C}_1$ from $\mathcal{H}_0$ to $(\bigotimes_{i=1}^N \mathcal{H}_{2i-1}) \otimes \mathcal{H}_M$;

ii. parallel application of the unitary channels $\mathcal{U}_g^{(i)}$ on $\mathcal{H}_{2i-1}$;

iii. final application of a postprocessing channel $\mathcal{C}_2$ from $(\bigotimes_{i=1}^N \mathcal{H}_{2i}) \otimes \mathcal{H}_M$ to $\mathcal{H}_{2N+1}$.

This means that the problem of finding an optimal $(1^{\otimes N}, 1)$-comb mapping the set of unitary channels $\{\mathcal{U}_g^{(i)}\}_{i=1\dots N}$ to $\mathcal{V}_g$ is equivalent to the problem of finding an optimal 2-comb that maps a single channel $\mathcal{U}_g = \bigotimes_{i=1}^N \mathcal{U}_g^{(i)}$ to $\mathcal{V}_g$,
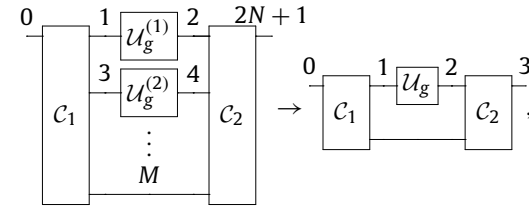
where from now on we refer to $\bigotimes_{i=1}^N \mathcal{H}_{2i-1}$ as the new Hilbert space $\mathcal{H}_1$, and we refer to $\bigotimes_{i=1}^N \mathcal{H}_{2i}$ as the new Hilbert space $\mathcal{H}_2$ to be coherent with the notation used in Section 2.

3.2. The optimal circuit

Thanks to the results of the previous section, the optimization problem (23) can be restated as follows:

$$\underset{R}{\text{maximize}} \quad F(R) = \frac{1}{d_0^2} \int dg \langle\!\langle V_g|_{30} \langle\!\langle U_g^*|_{21} R |U_g^*\rangle\!\rangle_{21} |V_g\rangle\!\rangle_{30}$$

$$\text{subject to} \quad \text{Tr}_3[R] = I_2 \otimes S_{10}, \quad \text{Tr}_1[S] = I_0, \; R, S \geqslant 0 \tag{31}$$

where we used the notation $|A\rangle\!\rangle_{ij} \in \mathcal{H}_i \otimes \mathcal{H}_j$. The constraints on R translate the condition that R is a deterministic 2-comb (see Eq. (7)).

---

[3] We stress that, since g and h are independent indices, W_h and W_g are two independent representations of the same group G. We will use this notation whenever the same representation acts independently on two different Hilbert spaces.

As a consequence of Lemma 3 it is not restrictive to search for the optimal comb R for the problem (31) among those having the following symmetry

$$\left[R, \left(V_h^* \otimes U_h\right)_{01} \otimes \left(U_g^* \otimes V_g\right)_{23}\right] = 0 \quad \forall g, h \in G, \tag{32}$$

where the two independent unitary representations of group G act on Hilbert spaces $\mathcal{H}_0 \otimes \mathcal{H}_1$ and $\mathcal{H}_2 \otimes \mathcal{H}_3$, respectively. It is now useful to consider the decompositions of U and V into irreducible representations as follows

$$U_h = \bigoplus_\beta U_h^{[\beta]} \otimes I_{m_\beta} \qquad V_h = \bigoplus_a V_h^{[a]} \otimes I_{m_a}$$

$$U_g = \bigoplus_\gamma U_g^{[\gamma]} \otimes I_{m_\gamma} \qquad V_g = \bigoplus_d V_g^{[d]} \otimes I_{m_d},$$

where for $\forall f \in G$, $x \in \{a, d\}$, $\xi \in \{\beta, \gamma\}$

$$U_f^{[\xi]} \in \mathcal{L}(\mathcal{H}_\xi) \qquad V_f^{[x]} \in \mathcal{L}(\mathcal{H}_x) \tag{33}$$

are unitary irreducible representations (irreps) of G and I_ξ, I_{m_x} are the identity operators on the multiplicity spaces $\mathcal{H}_{m_\xi}$ and $\mathcal{H}_{m_x}$. As we prove in Appendix A, we can without loss of generality restrict ourselves to the case in which the multiplicity spaces $\mathcal{H}_{m_\xi}$ are one dimensional for all ξ i.e. $U_f = \bigoplus_\xi U_f^{[\xi]}$ and

$$V_f^{[x]} \otimes U_f^{[\xi]*} = \bigoplus_Y W_f^{[Y]} \otimes I_{m_Y^{x,\xi}}. \tag{34}$$

Eq. (34) induces the following decomposition of Hilbert spaces:

$$\mathcal{H}_0 = \bigoplus_a \mathcal{H}_a \otimes \mathcal{H}_{m_a} \qquad \mathcal{H}_1 = \bigoplus_\beta \mathcal{H}_\beta$$

$$\mathcal{H}_2 = \bigoplus_\gamma \mathcal{H}_\gamma \qquad \mathcal{H}_3 = \bigoplus_d \mathcal{H}_d \otimes \mathcal{H}_{m_d}$$

$$\mathcal{H}_0 \otimes \mathcal{H}_1 = \bigoplus_K \mathcal{H}_K \otimes \mathcal{H}_{m_K}$$

$$\mathcal{H}_2 \otimes \mathcal{H}_3 = \bigoplus_L \mathcal{H}_L \otimes \mathcal{H}_{m_L}$$

$$\mathcal{H}_{m_K} = \bigoplus_{a,\beta} \mathcal{H}_{m_K^{a,\beta}} \qquad \mathcal{H}_{m_L} = \bigoplus_{\gamma,d} \mathcal{H}_{m_L^{\gamma,d}}. \tag{35}$$

We notice that with the notation of Eq. (35) one has $\mathcal{H}_{m_a} \subset \mathcal{H}_{m_K^{a,\beta}}$ and $\mathcal{H}_{m_d} \subset \mathcal{H}_{m_L^{\gamma,d}}$.

The commutation relation (32) can be rewritten as

$$\left[R, \left(\bigoplus_K W_h^{[K]} \otimes I_{m_K}\right) \otimes \left(\bigoplus_L W_g^{*[L]} \otimes I_{m_L}\right)\right] = 0 \tag{36}$$

which thanks to the Schur's lemma implies

$$R = \sum_{K,L} \Pi^K \otimes \Pi^L \otimes R^{KL} \tag{37}$$

where $R^{KL} \in \mathcal{L}(\mathcal{H}_{m_K} \otimes \mathcal{H}_{m_L})$ and Π^Y for $Y \in \{K, L\}$ is a projector onto $\mathcal{H}_Y$. It is convenient to define the projectors $P_Y^{x,\xi}$ on the multiplicity space $\mathcal{H}_{m_Y^{x,\xi}}$, P_x on the multiplicity space $\mathcal{H}_{m_x}$ and the Π_x on the representation spaces $\mathcal{H}_x$. We also define projector $P_Y^x := \sum_\xi P_Y^{x,\xi}$ onto a subspace $\bigoplus_\xi \mathcal{H}_{m_Y^{x,\xi}}$. In the following m_x will denote the dimension of the multiplicity space $\mathcal{H}_{m_x}$, d_x will denote the dimension of the representation space $\mathcal{H}_x$ and $m_Y^{x,\xi}$ will denote the dimension of the multiplicity space $\mathcal{H}_{m_Y^{x,\xi}}$.

The main result of this section, stated in the following proposition, is that the optimization problem (31) can be transformed into an optimization problem defined by a set of quadratic expressions for a probability distribution vector.

Proposition 5. *Let us consider the following optimization problem*

$$\underset{p_K^a}{\text{maximize}} \quad \Phi\left(p_K^a\right) = \sum_K \left(\sum_a \sqrt{q_K^a p_K^a}\right)^2$$

$$\text{subject to} \quad \sum_K p_K^a = 1 \quad \forall a$$

$$p_K^a \geqslant 0. \tag{38}$$

where $q_K^a = \frac{m_a d_a}{d_K d_0^2} \sum_\beta m_K^{a\beta} d_\beta$ and let $\check{R} = \check{R}(p_K^a)$ be defined as follows:

$$\check{R} := \sum_{KL} \Pi_K \otimes \Pi_L \otimes \check{R}^{KL}$$

$$\check{R}^{KL} = \left(\delta_{KL} |\psi_K\rangle\langle\psi_K| + \sum_\beta D_K^\beta \otimes \sum_{\gamma \neq \beta} \Delta_L^\gamma\right)$$

$$D_K^\beta = d_K d_\beta \sum_a p_K^a \frac{P_K^{a\beta}}{h_K^a} \qquad \Delta_L^\gamma = \frac{d_\gamma P_L^\gamma}{\text{Tr}[P_L^\gamma] d_L k_L^\gamma}$$

$$|\psi_K\rangle = \sum_{a,\beta} \sqrt{\frac{p_K^a d_\beta^2}{h_K^a}} |I_{m_K^{a,\beta}}\rangle\rangle$$

$$h_K^a = \frac{d_K^2}{m_a d_a} \sum_\beta m_K^{a\beta} d_\beta, \tag{39}$$

where k_L^γ denotes for how many L's $W_g^{[L]}$ is in the decomposition of $U_g^{*[\gamma]} \otimes V_g^{[d]}$ for some d.

If $\tilde{p}_K^a$ is a solution of the optimization problem (38) then $\check{R}(\tilde{p}_K^a)$ is a solution of the optimization problem of (31) and $F(\check{R}) = \Phi(\tilde{p}_K^a)$.

We split the proof of Proposition 5 into two parts. In the first lemma we prove that the operator defined through the ansatz of Eq. (39) is a well defined deterministic 2-comb.

Lemma 4. *Let $\check{R}$ have the form as in Eq. (39). Then $\check{R}$ satisfies the constraints of Eq. (31) if and only if $\sum_K p_K^a = 1$ $\forall a$ and $p_K^a \geqslant 0$.*

Proof. From the decompositions of Hilbert spaces in Eq. (35) we have

$$(\mathcal{H}_a \otimes \mathcal{H}_{m_a})_0 \otimes \mathcal{H}_1 = \bigoplus_{K,\beta} \mathcal{H}_K \otimes \mathcal{H}_{m_K^{a,\beta}}$$

$$(\mathcal{H}_\gamma)_2 \otimes \mathcal{H}_3 = \bigoplus_{L,d} \mathcal{H}_L \otimes \mathcal{H}_{m_L^{d,\gamma}} \tag{40}$$

which lead to the following identities

$$(\Pi_a \otimes P_a)_0 \otimes I_1 = \sum_{K,\beta} \Pi^K \otimes P_K^{a,\beta}$$

$$(\Pi_\gamma)_2 \otimes I_3 = \sum_{L,d} \Pi^L \otimes P_L^{\gamma,d}. \tag{41}$$

We now recall the normalization constraints for a 2-comb R: $\text{Tr}_3[R] = I_2 \otimes S_{10}$ and $\text{Tr}_1[S] = I_0$. If R obeys Eq. (32) we have $[\text{Tr}_3[R], (V_h^* \otimes U_h)_{01} \otimes (U_g^*)_2] = 0$ $\forall g, h \in G$, which gives, by using the Schur lemma $\text{Tr}_3[R] = \sum_{\gamma, K} \Pi_\gamma \otimes \Pi^K \otimes Q^{K,\gamma}$. Moreover, since we have also $[S_{10}, (V_h^* \otimes U_h)_{01}]$ which implies $I_2 \otimes S_{10} = \sum_\gamma \Pi^\gamma \otimes \sum_K \Pi^K \otimes S^K$, the condition $\text{Tr}_3[R] = I_2 \otimes S_{10}$ is equivalent to $Q^{K,\gamma} = S^K$ independently on γ. If we write $Q^{K,\gamma}$ in terms of R we have then

$$\sum_{L,d} \frac{d_L}{d_\gamma} \text{Tr}_{m_L}\left(R^{KL} P_L^{\gamma,d}\right) = S^K \quad \forall \gamma, \forall K, \tag{42}$$

where Tr_{m_L} indicates the trace over $\mathcal{H}_{m_L}$. Similarly, the condition $\mathrm{Tr}_1[S] = I_0$ can be rewritten as

$$d_a P_a = \sum_{K,\beta} \mathrm{Tr}_{\mathcal{H}_a} \mathrm{Tr}_1\big[S\big(\Pi^K \otimes P_K^{a,\beta}\big)\big] \quad \forall a. \tag{43}$$

We notice that by construction operator $\check{R}$ defined in Eq. (39) obeys the symmetry from Eq. (32). In addition $\check{R}$ obeys $\mathrm{Tr}_{23}[\check{R}]/\dim \mathcal{H}_2 = \check{S}_{10} = \sum_{K,a,\beta} s_K^{a,\beta} \Pi^K \otimes P_K^{a,\beta}$, which allows us to rewrite Eqs. (42) and (43) for $\check{R}$ in more convenient form:

$$\sum_{L,d} \frac{d_L}{d_\gamma} \mathrm{Tr}_{m_L}\big(\check{R}^{KL} P_L^{\gamma,d}\big) = \check{S}^K \quad \forall \gamma, \ \forall K, \tag{44}$$

$$\sum_{K,\beta} \frac{d_K}{d_a m_a} \mathrm{Tr}_{m_K}\big(\check{S}^K P_K^{a,\beta}\big) = 1 \quad \forall a, \tag{45}$$

where we used Eq. (41). In the following we demonstrate that the above two equations are fulfilled, i.e. $\check{R}$ is a properly normalized 2-comb. We notice that $\mathrm{Tr}_{m_L}[|\psi_K\rangle\langle\psi_K|P_K^\gamma] = d_\gamma D_K^\gamma/d_K$, $\mathrm{Tr}_{m_L}[\Delta_L^{\gamma'} P_L^\gamma] = \frac{\delta_{\gamma\gamma'} d_\gamma}{(k_L^\gamma d_L)}$. This implies $(\frac{d_L}{d_\gamma}) \sum_L \mathrm{Tr}_{m_L}[\check{R}^{KL} P_L^\gamma] = \sum_\beta D_K^\beta \equiv \check{S}^K$ for each K and independently on γ. Thus, the first normalization condition is satisfied. Inserting $\check{S}^K$ into Eq. (45) we obtain the condition $\sum_K p_K^a = 1 \ \forall a$. The positivity of the p_K^a guarantees the positivity of $\check{R}$ and $\check{S}$. $\quad\square$

In the next lemma we prove that the deterministic 2-comb that solves the optimization problem (31) can be assumed without loss of generality to be of the form of Eq. (39).

Lemma 5. *For any deterministic 2-comb R there exist a set of positive coefficients p_K^a, $\sum_K p_K^a = 1 \ \forall a$ such that for the 2-comb $\check{R}(p_K^a)$ defined by Eq. (39) we have $F(R) \leqslant F(\check{R}) = \Phi(p_K^a)$.*

Proof. From Eqs. (36) and (37) we have

$$F(R) = \frac{1}{d_0^2} \int dg \langle\!\langle V_g|_{30} \langle\!\langle U_g^*|_{21} R | U_g^*\rangle\!\rangle_{21} |V_g\rangle\!\rangle_{30}$$

$$= \frac{1}{(d_0)^2} \langle\!\langle I|_{30} \langle\!\langle I|_{21} R |I\rangle\!\rangle_{21} |I\rangle\!\rangle_{30}$$

$$= \sum_K \frac{d_K}{(d_0)^2} \sum_{a,\beta,\gamma,d} \langle\!\langle I_{m_K^{a,\beta}}| R^{KK} | I_{m_K^{\gamma,d}}\rangle\!\rangle, \tag{46}$$

where $|I_{m_K^{a,\beta}}\rangle\!\rangle \in \mathcal{H}_{m_K^{a,\beta}} \otimes \mathcal{H}_{m_K^{a,\beta}}$ and we used $|I\rangle\!\rangle_{03}|I\rangle\!\rangle_{12} = \sum_K |I_K\rangle\!\rangle \sum_{a,\beta} |I_{m_K^{a,\beta}}\rangle\!\rangle$ with $|I_K\rangle\!\rangle \in \mathcal{H}_K \otimes \mathcal{H}_K$.

For a positive operator X and arbitrary vectors $|\psi\rangle$ and $|\phi\rangle$ we have: $|\langle\psi|X|\varphi\rangle| \leqslant \sqrt{\langle\psi|X|\psi\rangle}\sqrt{\langle\varphi|X|\varphi\rangle}$, $\langle\psi|X|\psi\rangle \leqslant \langle\psi|\psi\rangle \mathrm{Tr}[X]$. Moreover we have $\langle\!\langle I_{m_K^{a,\beta}}|I_{m_K^{a,\beta}}\rangle\!\rangle = m_K^{a,\beta}$ and $|I_{m_K^{a,\beta}}\rangle\!\rangle = P_K^{a,\beta} \otimes P_K^{a,\beta}|I_{m_K^{a,\beta}}\rangle\!\rangle$. Applying the above two inequalities to Eq. (46) we obtain

$$F(R) \leqslant \sum_K \frac{d_K}{(d_0)^2} \left(\sum_{a,\beta} \sqrt{m_K^{a,\beta} d_\beta} \sqrt{\frac{R_{KK}^{a\beta\beta a}}{d_\beta}}\right)^2$$

$$\leqslant \sum_K \frac{d_K}{(d_0)^2} \left(\sum_a \sqrt{\sum_\beta m_K^{a,\beta} d_\beta \sum_{\beta'} \frac{R_{KK}^{a\beta'\beta'a}}{d_{\beta'}}}\right)^2$$

$$\leqslant \sum_K \frac{d_K}{(d_0)^2} \left(\sum_a \sqrt{\sum_\beta m_K^{a,\beta} d_\beta \sum_{\beta'Ld} \frac{d_L R_{KL}^{a\beta'\beta'd}}{d_{\beta'} d_K}}\right)^2 \tag{47}$$

where we used Schwarz inequality again in the second step and we defined $R_{KL}^{a\beta\gamma d} = \mathrm{Tr}[R^{KL} P_K^{a,\beta} \otimes P_L^{\gamma,d}]$. Let us now define $p_K^a = \sum_{\beta,L,d}(d_K d_L R_{KL}^{a\beta\beta d})/(m_a d_a d_\beta)$ and $\check{R}(p_K^a)$ with the ansatz of Eq. (39). We notice that the positivity of $R_{KL}^{a\beta\beta d}$ implies that $p_K^a \geqslant 0$. By substituting the above definition into Eq. (47) we have

$$F(R) \leqslant \sum_K \left(\sum_a \sqrt{q_K^a p_K^a}\right)^2 = F(\check{R})$$

where we inserted the definition of q_K^a given in Proposition 5.

It only remains to prove that $\sum_K p_K^a = 1 \ \forall a$. Since Eq. (44) holds for any γ we can insert Eq. (44) into Eq. (45) in such a way that for every term we choose $\gamma = \beta$ and obtain

$$\sum_{K,\beta,L,d} \frac{d_L d_K}{d_a m_a d_\beta} \mathrm{Tr}\big[R^{KL} P_K^{a,\beta} \otimes P_L^{\beta,d}\big] = 1 \quad \forall a$$

which completes the proof. $\quad\square$

One can now easily prove that the problem in Eq. (38) can be expressed as a semidefinite program of Eq. (24). Indeed, one can take the spaces $\mathcal{H} := \mathrm{span}(|a\rangle \otimes |K\rangle)$ and $\mathcal{K} := \mathrm{span}(|a\rangle)$, with $X := \sum_K |\varphi_K\rangle\langle\varphi_K| \otimes |K\rangle\langle K|$, $|\varphi_K\rangle := \sum_a \sqrt{q_K^a}|a\rangle$, $Y := \sum_a |a\rangle\langle a|$, the map $\mathcal{F}$ being just given by

$$\mathcal{F}(\rho) := \sum_{K,a} \mathrm{Tr}\big[\rho\big(|a\rangle\langle a| \otimes |K\rangle\langle K|\big)\big]|a\rangle\langle a|. \tag{48}$$

Finally, notice that the constraint in Eq. (38) involves an equal sign, namely $\mathcal{F}(\rho) = Y$. However, we can without loss of generality consider the looser constraint $\mathcal{F}(\rho) \leqslant Y$ because for any ρ satisfying $\mathcal{F}(\rho) < Y$ one can find ρ' such that $\mathcal{F}(\rho') = Y$ and $\mathrm{Tr}[\rho'X] \geqslant \mathrm{Tr}[\rho X]$. This implies that the final formulation corresponds to a much simpler semidefinite program than the original one in Eq. (23).

4. Examples

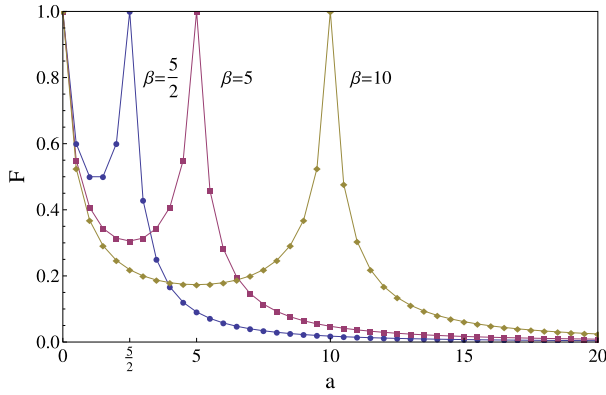### 4.1. Transformations between irreducible representations

The simplest problem that falls into our general setting is the transformation of unitary channels from an irreducible representation β of group $SU(2)$ into channels from a different irreducible representation a of the same group. Since we have only one irrep a the figure of merit (38) simplifies to $F(p_K^a) = \sum_K q_K^a p_K^a$. It is clear that the maximum $F = \max_K q_K^a$ is achieved by a probability distribution p_K^a with just one non-zero entry. Let us remind that the irreps of $SU(2)$ are defined by a half-integer called *spin*, and the generators of the representation with spin l are the usual quantum angular momentum components $J_x^{(l)}, J_y^{(l)}, J_z^{(l)}$. Notice also that for the group $SU(2)$ the complex conjugate representation of spin l is equivalent to the l representation, and is obtained by conjugating the l representation with the unitary $\exp(-i\pi J_y)$. Moreover, the irreps of $SU(2)$ obey a simple composition rule, when they are tensorized

$$U_a \otimes U_\beta = \bigoplus_{K=|a-\beta|}^{a+\beta} U_K \tag{49}$$

This implies $m_K^{a\beta} = 1$ and in our case for each K, a there exists exactly one β, which leads with a to irrep K. Since the dimension of the spin j irrep is $d_j = 2j + 1$ we have $q_K^a = d_\beta/(d_K d_a) = (2\beta + 1)(2a + 1)^{-1}(2K + 1)^{-1}$ and

$$F_{max} = \frac{2\beta + 1}{(2a + 1)(2|a - \beta| + 1)}. \tag{50}$$

Fig. 1. Average process fidelity F of an optimal transformation between irreducible representations β and a for the group $SU(2)$. The three lines correspond to different choice of the starting representation β, while the x axes represents the choice of the target irrep a.

As one might expect we can mimic reasonably only the irreps that have spin number a very close to β, the irrep that we have at disposal, or irreps that are very close to the trivial representation, namely those having a very small a. For illustration of the achievable process fidelities see Fig. 1.

4.2. $1 \rightarrow 2$ cloning of $SU(d)$ transformations

The results of Section 3.2 enable us to simplify the optimization of the $1 \rightarrow 2$ cloning of an $SU(d)$ transformation originally done in Ref. [31]. From our current perspective the problem might be viewed as a transformation of the defining representation U of $SU(d)$ into the reducible representation $U \otimes U$. The $d_0 = d^2$ dimensional representation $U \otimes U$ decomposes into two irreps acting in symmetric and antisymmetric subspaces $\mathcal{H}_\pm$. Let us distinguish these irreps by the index $a = \pm$. Their dimensions are $d_\pm = d(d \pm 1)/2$. On the other hand, the starting representation U is irreducible, which implies that the indices β and γ range over a single value. The representation $U \otimes U \otimes U^*$ defining the symmetries in Eq. (32) decomposes into three irreps, which we denote $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$. The d-dimensional representation $\hat{\alpha}$ appears with multiplicity two, whereas $\hat{\beta}$ and $\hat{\gamma}$ have multiplicity one and dimensions $d(d_+ - 1)$, $d(d_- - 1)$, respectively.

The following table summarizes all the parameters q_K^a that are used for the optimization in Proposition 5. For the sake of simplicity we actually report the expressions for $d^4 q_K^a$.

	$K = \hat{\alpha}$	$K = \hat{\beta}$	$K = \hat{\gamma}$
$a = +$	d_+	$d_+/(d_+ - 1)$	0
$a = -$	d_-	0	$d_-/(d_- - 1)$

The figure of merit (38) for this problem then takes the following form

$$F = \left(\sqrt{q_{\hat{\alpha}}^+ p_{\hat{\alpha}}^+} + \sqrt{q_{\hat{\alpha}}^- p_{\hat{\alpha}}^-}\right)^2 + q_{\hat{\beta}}^+ (1 - p_{\hat{\alpha}}^+) + q_{\hat{\gamma}}^- (1 - p_{\hat{\alpha}}^-)$$

where we also used the constraint $\sum_K p_K^a = 1$ $\forall a$. Under the constraints $0 \leqslant p_K^a \leqslant 1$ the maximization of F yields $p_{\hat{\alpha}}^+ = p_{\hat{\alpha}}^- = 1$ and $F = (\sqrt{d_+} + \sqrt{d_-})^2/d^4$ in agreement with [31].

4.3. $1 \rightarrow N$ cloning of $SU(2)$ transformations

Cloning of qubit unitary gates might be viewed as a transformation of the defining representation U of $SU(2)$ into the reducible
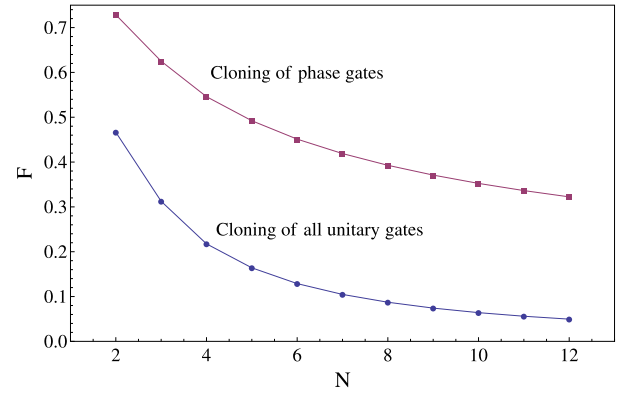
Fig. 2. Average process fidelity of optimal $1 \rightarrow N$ cloning of a qubit channel. The bottom line shows optimal process fidelity for cloning of all qubit unitary channels, whereas the top line corresponds to optimal cloning of only qubit phase gates.

representation $U^{\otimes N}$. The representation $U^{\otimes N}$ decomposes into irreps as:

$$U^{\otimes N} = \bigoplus_{a = \langle\langle N/2 \rangle\rangle}^{N/2} U_a \otimes I_{m_a}, \tag{51}$$

where $\langle\langle x \rangle\rangle$ denotes the fractional part of x (i.e. $\langle\langle N/2 \rangle\rangle$ is 0 for N even and $1/2$ for N odd) and $m_a = \frac{2a+1}{N/2+a+1}\binom{N}{N/2+a}$ [47]. Since the input representation has $\beta = 1/2$, the irreps in $U^{\otimes N} \otimes U^*$ are labeled by K ranging from $\langle\langle (N+1)/2 \rangle\rangle$ to $(N+1)/2$. In particular, each value of K derives either from $a = K - 1/2$ or from $a = K + 1/2$. The only exceptions to this rule are the maximum K and $K = 0$ for odd N, which derive from a single value of a. This simplifies the problem and we can rewrite it as the maximization of

$$F = \sum_{K = \langle\langle \frac{N}{2} \rangle\rangle + \frac{1}{2}}^{\frac{N-1}{2}} \left(\sqrt{q_K^{K - \frac{1}{2}} x_K} + \sqrt{q_K^{K + \frac{1}{2}} (1 - x_{K+1})}\right)^2$$

$$+ q_{\frac{N+1}{2}}^{\frac{N}{2}} x_{\frac{N+1}{2}} + 2\langle\langle \tfrac{N}{2} \rangle\rangle q_0^{\frac{1}{2}} (1 - x_0) \tag{52}$$

with respect to $0 \leqslant x_K \leqslant 1$, where we denoted $x_K \equiv p_K^{K - 1/2}$ and consequently $p_K^{K+1/2} \equiv 1 - x_{K+1}$ due to the normalization constraints (38). Thus, for a given N we need to optimize roughly $N/2$ parameters x_K. This can be done analytically by symbolic calculus for small values of N or numerically. In Fig. 2 the optimal fidelity is plotted for N up to 12.

4.4. $1 \rightarrow N$ cloning of qubit phase gates

The third application of the general method that we show here is cloning of qubit phase gates, i.e. unitary transformations $U = diag(1, e^{i\phi})$ that are diagonal in the computational basis $\{|0\rangle, |1\rangle\}$. In this case the input representation U of $U(1)$ is reducible, and it is transformed into the different reducible representation $U^{\otimes N}$. Since $U(1)$ has only 1-dimensional irreps we have $d_a = d_\beta = d_K = 1$. We can decompose $U^{\otimes N}$ as $U^{\otimes N} = \bigoplus_{a=0}^{N} e^{ia\phi} \pi_a$, where π_a denotes the projection on the subspace spanned by tensor products of vectors in the computational basis with a factors equal to $|1\rangle$. Consequently, $U^{\otimes N} \otimes U^*$ contains representations $e^{iK\phi}$ $K = -1, \ldots, N$ and each $K \in \{0, \ldots, N-1\}$ can be obtained either from $(a, \beta) = (K, 0)$ or from $(a, \beta) = (K - 1, 1)$. The irreps $K = -1$ and $K = N$ can derive only from one pair (a, β). This allows us to

rewrite the problem as the maximization of the following expression

$$F = \sum_{K=0}^{N-1}\left(\binom{N}{K}\sqrt{x_K} + \binom{N}{K+1}\sqrt{1-x_{K+1}}\right)^2$$
$$+ (1-x_0) + x_N \tag{53}$$

with respect to $0 \leqslant x_K \leqslant 1$, where we denoted $x_K \equiv p_K^K$ and $p_K^{K+1} \equiv 1 - x_{K+1}$ thanks to the normalization constraints (38). We performed the optimization for small values of N by symbolic calculus. As one could expect, the optimal fidelity for $1 \to N$ cloning of phase gates is much better than the one for cloning of arbitrary qubit unitary channels, as illustrated in Fig. 2.

4.5. Realization of $1 \to 2$ cloning of qubit phase gates

In this section we discuss physical schemes for the realization of optimal $1 \to 2$ cloning of qubit phase gates. Before describing our proposals, let us summarize the results implied by the previous sections. In the case of $N = 2$ the maximization of Eq. (53) yields $x_0 \equiv p_0^0 = 1$, $x_1 \equiv p_1^1 = 1/2$, $x_2 \equiv p_2^2 = 0$, which using Eq. (39) gives

$$R = |\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|$$
$$+ \left(\frac{1}{2}P_1^{1,0} + P_0^{0,0}\right) \otimes \frac{1}{3}\left(P_{-1}^{0,1} + \frac{1}{2}P_0^{1,1} + P_1^{2,1}\right)$$
$$+ \left(P_1^{2,1} + \frac{1}{2}P_0^{1,1}\right) \otimes \frac{1}{3}\left(P_0^{0,0} + \frac{1}{2}P_1^{1,0} + P_2^{2,0}\right), \tag{54}$$

where we defined

$$|\psi_0\rangle = |000000\rangle + \frac{1}{\sqrt{2}}|011101\rangle + \frac{1}{\sqrt{2}}|101110\rangle$$
$$|\psi_1\rangle = |111111\rangle + \frac{1}{\sqrt{2}}|010001\rangle + \frac{1}{\sqrt{2}}|100010\rangle,$$

we used notation $|a\bar{a}\beta\gamma d\bar{d}\rangle \equiv |a\bar{a}\rangle_0|\beta\rangle_1|\gamma\rangle_2|d\bar{d}\rangle_3$ and the tensor products are ordered as $X \otimes Y \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3)$. Let us evaluate $R * |U\rangle\rangle\langle\langle U|$, which corresponds to an overall channel between $\mathcal{H}_0$ and $\mathcal{H}_3$ that is created after the unitary gate $U = diag(1, e^{i\phi})$ is inserted into the cloning circuit. All the terms $P_K^{a,\beta} \otimes P_L^{d,\gamma}$ in Eq. (54) do not contribute, since they have $\beta \neq \gamma$ and $|U\rangle\rangle$ contains only the terms $|\beta\rangle_1|\beta\rangle_2$. Thus, we obtain

$$R * |U\rangle\rangle\langle\langle U| = |\varphi_0\rangle\langle\varphi_0| + |\varphi_1\rangle\langle\varphi_1|$$
$$|\varphi_0\rangle = |0000\rangle + \frac{e^{i\phi}}{\sqrt{2}}(|0101\rangle + |1010\rangle)$$
$$|\varphi_1\rangle = e^{i\phi}|1111\rangle + \frac{1}{\sqrt{2}}(|0101\rangle + |1010\rangle). \tag{55}$$

One can check by direct calculation that this channel is achieved by the following quantum circuit
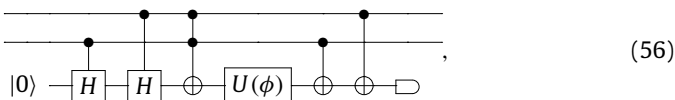
where H denotes the Hadamard gate and the ancillary qubit is prepared in state $|0\rangle$. The dimension of quantum system that is used in parallel with the action of the cloned gate is called quantum memory in the context of quantum protocols [48]. In the proposed circuit the memory is four dimensional (2 qubits). In order to make the memory smaller, one can employ the techniques from Ref. [48] that are based on the covariance of the problem. In this way one

can trade a four dimensional quantum memory for a three dimensional memory and one bit of classical communication. We were able to describe such a memory efficient realization of the optimal cloning of a phase gate in terms of isometries (see Eq. (57)),
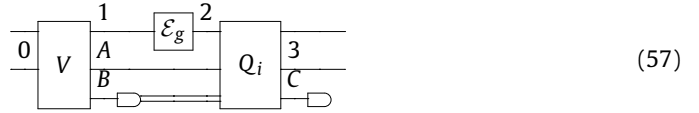
however synthesizing a corresponding quantum circuit goes beyond the scope of this manuscript. The isometry V in Eq. (57) is the following

$$V = |0\rangle_B\left(|1\rangle|0\rangle\langle00| + \frac{1}{\sqrt{2}}|2\rangle|1\rangle\langle01| + \frac{1}{\sqrt{2}}|3\rangle|1\rangle\langle10|\right)$$
$$+ |1\rangle_B\left(|1\rangle|1\rangle\langle11| + \frac{1}{\sqrt{2}}|2\rangle|0\rangle\langle01| + \frac{1}{\sqrt{2}}|3\rangle|0\rangle\langle10|\right)$$

where the shortened expressions $|1\rangle|0\rangle\langle00|$ stand for $|1\rangle_A|0\rangle_1\langle00|_0$ and the subsystems A, B are a qutrit, and a qubit, respectively. The result of the measurement in the $\{|0\rangle, |1\rangle\}$ basis determines whether Q_0 or Q_1 will be used after the action of the input gate. The isometries Q_0 and Q_1 are defined as follows

$$Q_0 = |1\rangle_C\left(|00\rangle\langle0|\langle1| + |01\rangle\langle1|\langle2| + |10\rangle\langle1|\langle3|\right)$$
$$+ |2\rangle_C|11\rangle\langle1|\langle1| + |3\rangle_C|00\rangle\langle0|\langle2| + |4\rangle_C|00\rangle\langle0|\langle3|$$
$$Q_1 = |1\rangle_C\left(|11\rangle\langle1|\langle1| + |01\rangle\langle0|\langle2| + |10\rangle\langle0|\langle3|\right)$$
$$+ |2\rangle_C|00\rangle\langle0|\langle1| + |3\rangle_C|11\rangle\langle1|\langle2| + |4\rangle_C|11\rangle\langle1|\langle3|,$$

where we shortened $|00\rangle_3\langle0|_2\langle1|_A$ as $|00\rangle\langle0|\langle1|$ and the ancillary quantum system C is four dimensional.

5. Conclusions

In this paper we reviewed the general theory of higher order quantum maps and within this framework we addressed a general class of quantum computational tasks involving the processing of unitary channels. We considered the scenario in which one has access to a single use of N unknown unitary channels $\{\mathcal{U}_g^{(i)}\}_{i=1...N}$ in an arbitrary sequential order, where the action of each $\mathcal{U}_g^{(i)}$ on a state ρ is described by a unitary representation $U_g^{(i)}$ of a fixed compact group G, i.e. $\mathcal{U}_g^{(i)}(\rho) = U_g^{(i)}\rho U_g^{(i)\dagger}$. The task we considered is to exploit the uses of the unitary channels $\{\mathcal{U}_g^{(i)}\}_{i=1...N}$ to create a target unitary channel $\mathcal{V}_g$ which is described by a different unitary representation V_g of the same group G. As a figure of merit we chose the group average of the channel fidelity between the output channel and the ideal one. We proved that the optimal scheme does not require any non-circuital higher order map, but it can be realized by a three-step protocol: (i) application of a preprocessing channel $\mathcal{C}_1$, (ii) parallel application of the unitary channels $\mathcal{U}_g^{(i)}$ and (iii) final application of a postprocessing channel $\mathcal{C}_2$.

Moreover, we rephrased the circuit optimization problem as simplified semidefinite programming that significantly reduces the number of variables involved in the optimization, as can be appreciated by comparing the original formulation of the problem in Eq. (23) and the simplified one in Eq. (38). One can see, for example, that in the case of $1 \to N$ cloning of an $SU(2)$ gate (see Section 4.3) the number of parameters D in the semidefinite program exponentially reduces from $D \sim 2^{2N}$ to $D \sim N^2$. Remarkably, the results of Proposition 5 along with the results of Ref. [48] allow us to assess an upper bound to the amount of quantum memory which must be kept coherent from the optimal preprocessing

to the postprocessing phase through the parameter $\max_K m_K$, the maximal multiplicity in the decomposition of Eq. (35).

The quantum processing task that we consider in this paper is very general and includes a number of interesting scenarios as special cases. Indeed, in Section 4, besides recovering the results of $1 \to 2$ cloning of $SU(d)$ unitaries, we provided the optimal solution for the task of transforming an $SU(2)$ irrep into a different one, and for the $1 \to N$ cloning of $SU(2)$ and $SU(1)$. The last two cases illustrate how a stronger prior knowledge about the unknown unitaries enables a higher fidelity (see Fig. 2) in the same way as it happens for phase covariant [49] versus universal state cloning [50,51].

An alternative way to achieve the transformation from $\{\mathcal{U}_g^{(i)}\}_{i=1...N}$ to $\mathcal{V}_g$ is to estimate g and then to prepare the estimated unitary. This measure and prepare strategy can be generally more easily implemented than the pre- and post-processing one and has the advantage that it could be applied even in the case in which the uses $\{\mathcal{U}_g^{(i)}\}_{i=1...N}$ and the quantum state ρ which $\mathcal{V}_g$ will be applied to, are not available at the same time. Because of that, there can be situations in which one could prefer to apply the measure-and-prepare strategy if the consequent performance loss is below a given threshold. Within this perspective it would be useful to characterize under which conditions this two strategies achieve similar fidelity. Especially interesting would be the study of the asymptotic scaling of the optimal $N \to M$ cloning of unitaries and to verify whether the two strategies exhibit the same scaling for $M \to \infty$. This would be a generalization of the known result of the asymptotic convergence of optimal state cloning to state estimation [52]. The results of the current paper provide versatile tools for the study of this problem and this investigation will be the subject of future works.

Acknowledgements

Appendix A. Irrelevance of the multiplicity spaces

Our aim is to show that two sets of channels $\{\mathcal{U}_g : g \in G\}$, $\{\mathcal{U}_g' : g \in G\}$ defined by two representations of group G that differ only in the multiplicities of the irrep's are mutually perfectly transformable. This statement is made precise in the following lemma.

Lemma 6. *A set of unitary channels $\mathcal{U}_g$ defined by a representation $U_g = \bigoplus_\beta U_g^{[\beta]} \otimes I_{m_\beta}$ and a set of unitary channels $\mathcal{U}_g'$ defined by a representation $U_g' = \bigoplus_\beta U_g^{[\beta]}$ are perfectly mutually transformable, i.e. there exist two deterministic 2-combs R and $\widetilde{R}$ such that*

$$|U_g'\rangle\!\rangle\langle\!\langle U_g'| = R * |U_g\rangle\!\rangle\langle\!\langle U_g| \quad \forall g \in G$$
$$|U_g\rangle\!\rangle\langle\!\langle U_g| = \widetilde{R} * |U_g'\rangle\!\rangle\langle\!\langle U_g'| \quad \forall g \in G. \tag{A.1}$$

Proof. The proof is constructive. For the construction of R we define two channels $\mathcal{X}, \mathcal{Y}$ such that

$$-\boxed{\mathcal{U}_g'}- = -\boxed{\mathcal{X}}-\boxed{\mathcal{U}_g}-\boxed{\mathcal{Y}}- . \tag{A.2}$$

The channel $\mathcal{X}$ is an isometry that embeds the Hilbert space $\bigoplus_\beta \mathcal{H}_\beta$ in a subspace π of $\bigoplus_\beta \mathcal{H}_\beta \otimes \mathbb{C}^{m_\beta}$ defined by the choice of a single vector $|\phi_\beta\rangle$ in each multiplicity space. The channel $\mathcal{Y}$ has Kraus operators the inverse isometry $V^\dagger := \sum_\beta (I_\beta \otimes \langle\phi_\beta|)\Pi_\beta$ and $K_i\sqrt{I - VV^\dagger}$ where Π_β represents the projection on the subspace

$\mathcal{H}_\beta \otimes \mathbb{C}^{m_\beta}$, while K_i are Kraus operators of an arbitrary trace-preserving map from the support of $I - VV^\dagger$ to $\bigoplus_\beta \mathcal{H}_\beta$. Finally, $R := X \otimes Y$, with X, Y being Choi matrices of $\mathcal{X}, \mathcal{Y}$.

For the construction of $\tilde{R}$ we define an ancillary system M and two channels $\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}$ as follows

$$-\boxed{\mathcal{U}_g}- = -\boxed{\tilde{\mathcal{X}}}\overset{M}{}\boxed{\mathcal{U}_g'}\boxed{\tilde{\mathcal{Y}}}- . \tag{A.3}$$

We set the dimension of $\mathcal{H}_M$ to be $\max_\beta m_\beta$. The channel $\tilde{\mathcal{X}}$ is just an isometric embedding of $\bigoplus_\beta \mathcal{H}_\beta \otimes \mathbb{C}^{m_\beta}$ into $\bigoplus_\beta \mathcal{H}_\beta \otimes \mathbb{C}^M$. The channel $\tilde{\mathcal{Y}}$ is now analogous to $\mathcal{Y}$, with the only difference that one projects $\mathcal{H}_\beta \otimes \mathbb{C}^M$ into $\mathcal{H}_\beta \otimes \mathbb{C}^{m_\beta}$ and its orthogonal complement. $\square$

References

[1] G. Ludwig, Foundations of Quantum Mechanics, Springer, New York, 1983 and 1985, vols. 1 and 2.
[2] L. Hardy, arXiv:quant-ph/0101012.
[3] C.A. Fuchs, arXiv:quant-ph/0205039, 2002.
[4] G.M. D'Ariano, in: A. Bokulich, G. Jaeger (Eds.), Philosophy of Quantum Information and Entanglement, Cambridge University Press, Cambridge, UK, 2010.
[5] G. Brassard, Nat. Phys. 1 (2005) 2.
[6] L. Masanes, M.P. Mueller, New J. Phys. 13 (2011) 063001.
[7] B. Dakic, C. Brukner, in: H. Halvorson (Ed.), Deep Beauty: Understanding the Quantum World through Mathematical Innovation, Cambridge University Press, 2011, pp. 365–392.
[8] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A 84 (2011) 012311.
[9] A. Acín, Phys. Rev. Lett. 87 (2001) 177901.
[10] G.M. D'Ariano, P. LoPresti, M.G.A. Paris, Phys. Rev. Lett. 87 (2001) 270404.
[11] M.F. Sacchi, J. Opt. B 7 (2005) S333.
[12] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. Lett. 101 (2008) 180501.
[13] A.W. Harrow, A. Hassidim, D.W. Leung, J. Watrous, Phys. Rev. A 81 (2010) 032339.
[14] M.A. Nielsen, I.L. Chuang, Phys. Rev. Lett. 79 (1997) 321.
[15] S.F. Huelga, J.A. Vaccaro, A. Chefles, M.B. Plenio, Phys. Rev. A 63 (2001) 042303.
[16] S.D. Bartlett, W.J. Munro, Phys. Rev. Lett. 90 (2003) 117901.
[17] Y.-F. Huang, X.-F. Ren, Y.-S. Zhang, L.-M. Duan, G.-C. Guo, Phys. Rev. Lett. 93 (2004) 240501.
[18] Y.S. Weinstein, T.F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, D.G. Cory, J. Chem. Phys. 121 (13) (2004) 6117.
[19] J.L. O'Brien, G.J. Pryde, A. Gilchrist, D.F.V. James, N.K. Langford, T.C. Ralph, A.G. White, Phys. Rev. Lett. 93 (2004) 080502.
[20] M. Ziman, Phys. Rev. A 77 (2008) 062112.
[21] G. Gutoski, J. Watrous, in: Proc. of the 39th Annual ACM Symposium on Theory of Computation, 2007, p. 565.
[22] D. Deutsch, R. Jozsa, Proc. R. Soc. Lond. A 439 (1992) 553.
[23] L.K. Grover, in: Proceedings of the 28th Annual ACM Symposium on the Theory of Computing, 1996, p. 212.
[24] P. Shor, SIAM Rev. 41 (1999) 303.
[25] H.P. Yuen, arXiv:quant-ph/0207089.
[26] G.M. D'Ariano, D. Kretschmann, D.M. Schlingemann, R.F. Werner, Phys. Rev. A 76 (2007) 032328.
[27] S. Pirandola, S. Mancini, S. Lloyd, S.L. Braunstein, Nat. Phys. 4 (2008) 726.
[28] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. Lett. 101 (2008) 060401.
[29] G. Chiribella, G.M. D'Ariano, P. Perinotti, Europhys. Lett. 83 (2008) 30004.
[30] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A 80 (2009) 022339.
[31] G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. Lett. 101 (2008) 180504.
[32] A. Bisio, G. Chiribella, G.M. D'Ariano, S. Facchini, P. Perinotti, Phys. Rev. Lett. 102 (2009) 010404.
[33] A. Bisio, G. Chiribella, G.M. D'Ariano, S. Facchini, P. Perinotti, Phys. Rev. A 81 (2010) 032324.
[34] A. Bisio, G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A 82 (2010) 062305.
[35] A. Bisio, G.M. D'Ariano, P. Perinotti, M. Sedlak, Phys. Rev. A 84 (2011) 042330.
[36] A. Bisio, G.M. D'Ariano, P. Perinotti, M. Sedlak, Phys. Lett. A 375 (2011) 3425.
[37] G. Chiribella, G.M. D'Ariano, P. Perinotti, B. Valiron, arXiv:0912.0195.
[38] G. Chiribella, Phys. Rev. A 86 (2012) 040301.
[39] T. Colnaghi, G. Mauro D'Ariano, P. Perinotti, S. Facchini, Phys. Lett. A 376 (2012) 2940.
[40] A. Bisio, G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A 83 (2011) 022325.
[41] A. Bisio, G. Chiribella, G.M. D'Ariano, P. Perinotti, Acta Phys. Slovaca 61 (3) (2011) 273–390.
[42] K. Kraus, States, Effects, and Operations, Springer-Verlag, Berlin, 1983.
[43] M.-D. Choi, Linear Algebra Appl. 10 (1975) 285.
[44] O. Oreshkov, F. Costa, C. Brukner, Nat. Commun. 3 (2012) 1092.

[45] M. Raginsky, Phys. Lett. A 290 (2001) 11.
[46] A.S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, North-Holland, Amsterdam, 1982.
[47] J.I. Cirac, A.K. Ekert, C. Macchiavello, Phys. Rev. Lett. 82 (1999) 4344.
[48] A. Bisio, G.M. D'Ariano, P. Perinotti, M. Sedlák, Phys. Rev. A 85 (2012) 032333.
[49] D. Bruss, M. Cinchetti, G.M. D'Ariano, C. Macchiavello, Phys. Rev. A 62 (2000) 12302.
[50] V. Bužek, M. Hillery, Phys. Rev. A 54 (1996) 1844–1852.
[51] R. Werner, Phys. Rev. A 58 (1998) 1827.
[52] G. Chiribella, G.M. D'Ariano, Phys. Rev. Lett. 97 (2006) 250503.